Sensemaking with Shared Visualizations:
Investigating the Effects of Visualizations in Remote Collaborative Analysis

Aruna D. Balakrishnan
CMU-HCII-11-104
August 2011

Human-Computer Interaction Institute
School of Computer Science
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, Pennsylvania 15213

Thesis Committee:
Sara Kiesler (co-chair), Carnegie Mellon University
Susan Fussell (co-chair), Cornell University
Aniket Kittur, Carnegie Mellon University
Jason Hong, Carnegie Mellon University
Peter Pirolli, PARC Research

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

Keywords: computer-supported cooperative work, distributed work, remote collaborations, complex problem solving, information sharing, decision making, information overload, empirical studies, information visualization, collaboration, communication, computer-mediated communication, information access.

ABSTRACT

Increasingly, collaborators are separated geographically and are also faced with large quantities of information that can complicate collaboration. Visualizing information can help collaborators sort through large quantities of data, but visualizations help only when they promote effective problem-solving behaviors such as division of labor and open communication. This dissertation explores the impact of network visualizations on collaborative problem solving by examining three laboratory studies.

Using the "detective mystery" as an experimental paradigm, remote collaborators worked synchronously via instant messenger to identify a serial killer who was hidden within multiple crime reports. In the first study, the evidence was divided between the pair of collaborators. When collaborators were given a network visualization tool that showed them how the evidence was linked, they performed better than those without a visualization. The visualization also fostered discussion between partners.

The second study looked at whether the visualization would help collaborators if they already had full access to all the evidence. Whereas the visualization tool improved performance for collaborators with half the evidence, the same visualization tool did *not* improve performance when each collaborator had access to all the evidence. Collaborators seemed to be overwhelmed and did not approach the task systematically. Unlike their counterparts, who each had half the evidence, collaborators with all the evidence talked less, discussed fewer hypotheses, and did not divide the labor.

The final study asked whether interpersonal and task-oriented discussion-prompt interventions encourage collaborators to adopt problem-solving strategies necessary for success. Discussion-prompt interventions helped pairs improve their search and analysis process.

This dissertation suggests that visualization tools may prompt collaborations to be more systematic, but this depends on collaborators effectively using the visualization, finding relevant patterns within the visualization, and ultimately using these findings to direct their analysis.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

1.1 MOTIVATION

In 1998, Detective Dave Dickson of the Vancouver Police noticed that a number of cases involving missing women had gone unsolved. Dickson assembled a team of investigators, called the British Columbia Missing Women Task Force, that found clues about these missing women (Newton, 2005). One member of this task force, Detective Kim Rossmo, looked for similarities between the information Dickson shared and his own cases. Rossmo found a pattern of missing women concentrated in one area of Vancouver, the Downtown Eastside district. The more they searched, the more clear that pattern became. Ultimately, the task force found forty related but unsolved cases that dated back to 1971.

Detectives tracked down whatever clues they could, interviewing family members of the missing women. By 2000, the number of unsolved missing women cases reached 53. But in 2002 they caught a break that led to the arrest of Robert Pickton. Pickton was found guilty of six murders in 2007, but still claims responsibility for forty-nine murders within the same city. The investigation continues, and twenty more charges have been filed.

There are several important aspects of this story. First, two detectives worked together toward a shared goal. Second, to successfully solve the cases, the detectives had to search through a large volume of information and find a pattern that connected each case together. Their situation is similar to situations in many domains of analysis—in science, business, criminology, epidemiology, government, and intelligence. The amount of data that must be collected, perused, and analyzed to solve problems in such situations is often huge. Analysts must search vast datasets for patterns or anomalies. At times, the sheer body of information can exceed the unaided capacity of individual analysts.

To combat this "information overload," two main approaches have been taken. *Social* approaches make analysis a collaborative process. A long-held vision in CSCW is to improve distributed access to data for collaboration (Greif & Sarin, 1987). By having multiple perspectives on the data and considering the evidence together, collaborators can derive better

conclusions from the data. Two heads are better than one. *Cognitive* approaches, on the other hand, focus on enhancing an individual's cognitive capacity through visualization and other analytical tools that help people process more data, more rapidly (e.g., Billman, Convertino, Pirolli, Massar, & Shrager, 2005).

The premise of this dissertation is that a combination of social and cognitive approaches can improve the opportunities for success during complex analytical tasks. In particular, I explore collaborative analysis and how visualization tools affect collaboration and analysis. In the remainder of chapter one, I describe the particular task studied, and review prior literature and relevant theories of how collaborations and visualizations affect analysis. I also discuss the potential benefits and drawbacks of the social and cognitive approaches as well as the outline and chief contributions of this dissertation.

1.2   THE ANALYST'S TASK

Analysis is the "separation of a whole into its component parts" (analysis, 2011). Analysis is not a new method; rather, analysis is a process of study that extends back to the ancient Greeks (Beaney, 2011). What this process of analysis entails depends on whom you ask. Philosophers have debated, and continue to debate, between three main approaches to analysis: (1) regressive analysis, based on Euclidian geometry, wherein one works back towards the first of fundamental principles; (2) decompositional analysis, based on the ideas of Descartes, wherein one breaks things down into its logical components; and (3) interpretive or transformative analysis, based on the work of Frege and Russell, wherein one maintains that before decomposition is possible the material to be analyzed must first be translated into its logical forms (Beaney, 2011; Jahnke, 2003). In contrast to a philosopher, a business leader might describe analysis as competitive intelligence, or the process of collecting qualitative and quantitative information on competitors' activities in order to inform business decisions (Kahaner, 1997). While these two domains, philosophy and business, maintain different ideas about the process of analysis, the overall concept and goal of analysis remains the same. Both domains insist that exploring individual components of a whole is necessary in order to achieve a better understanding of that whole.

This dissertation focuses on types of analysis most often found in crime, intelligence, and business. Here, "crime analysis" specifically refers to law enforcement. This type of analysis has been popularized and fictionalized through detective crime novels and television dramas such as

Law and Order and *CSI*. Within the law enforcement community, there are several definitions of analysis. The most frequently cited definition is:

> a set of systematic, analytical processes directed at providing timely and pertinent information relative to crime patterns and trend correlations to assist the operational and administrative personnel in planning the deployment of resources for the prevention and suppression of criminal activities, aiding the investigative process, and increasing apprehensions and the clearance of cases. (Gottlieb, Arenberg, & Singh, 1994, p. 13)

Crime analysts look for patterns and anomalies in crime reports, and integrate this information to help policy makers create effective policing strategies, help find criminals, and prevent future crime. This systematic process involves taking into account many types of information and situational facts, "including sociodemographic, spatial, and temporal factors–to assist the police in criminal apprehension, crime and disorder reduction, and evaluation" (Boba, 2005, p. 6). Because this process is systematic, "crime analysis is not haphazard or anecdotal; rather, it involves the application of social science data collection procedures, analytic methods, and statistical techniques" (Boba, 2005, p. 6). Crime analysts rely on extensive qualitative interviews with many individuals, including witnesses and possible suspects, in addition to quantitative data logs regarding crime statistics and location-based information.

Similar to crime analysis, intelligence analysis is performed by intelligence agencies such as the Central Intelligence Agency. Intelligence analysis involves collecting information about situations and entities critical to national security and analyzing the numerical and factual information with the eventual outcome of an interpretation of events; it is designed to guide decision making (Kight, 1996). As with crime analysis, intelligence analysis requires a wide variety of information, and qualitative and quantitative techniques.

Business intelligence is very similar to intelligence and crime analysis, albeit with a different focus. For example, a competitive intelligence analyst pours over information and trends with respect to a business's competition (e.g., Kahaner, 1997). The business analyst will constantly search the web for relevant news that might hint at a competitor's next move. The business analyst will interpret this information, formulate a cohesive narrative about the competitor, and

try to predict what the competitor might do next in order to guide the appropriate decision makers.

The terms "analysis," "intelligence," and "analytics" are sometimes used interchangeably. I define "analysis" in a general manner to encompass many domains of analysis, only one of which is intelligence analysis. "Analytics" refers to a specific method of analysis commonly used in the intelligence and business communities when large volumes of quantitative data are involved. Analytics applies statistical methods and uses computer science technology and techniques to aid the analysis process, sifting through large datasets to help guide decision making (Kohavi, Rothleder, & Simoudis, 2002). For example, many companies save sales records from previous years; analytics creates trend reports from this information to help build forecasting models that may improve future decision making and business strategies.

Successful analysis can be extremely difficult. In the case of the Vancouver serial killer, it took detectives years to connect disconnected information into a cohesive story that described a potential serial killer. One reason it is difficult to connect disconnected information is that there *is* a correct answer, but the *form* of the correct answer is unknown (McGrath, 1984). Thus this type of analysis is an intellective, conceptual "choose task" because the correct answer must be invented, selected, or computed; the facts point to a correct conclusion, but the logic necessary to prove the solution may be difficult and not as intuitive (McGrath, 1984). In contrast, "Eureka-type tasks" have a correct answer that, once found, seems obvious in retrospect. While simple arithmetic problems have a solution of a known form (e.g., the answer to the equation $2 + 2$ is known to be a rational, real number), solutions to crime or intelligence analysis tasks may be in the form of identifying a suspicious person (e.g., Mr. Pickton), or a suspicious activity (e.g., prior arrests for violent activities, visiting radical mosques), or a collection of people and activities (e.g., missing women, a group of terrorists). The *form* of the solution that connects these pieces of distributed evidence may remain unclear. The facts may point to a correct conclusion, but the logic necessary to prove the solution may be difficult and not intuitive.

Additionally, analysts in domains like crime, intelligence, and business have limited time and vast amounts of information that comes from a variety of sources, in a variety of modes, such as text, video, and images (Johnston, 2005). For a single intelligence report, an analyst may comb through thousands of pieces of information. A recent two-year investigation by *The Washington*

Post found that "analysts who make sense of documents and conversations obtained by foreign and domestic spying share their judgment by publishing 50,000 intelligence reports each year - a volume so large that many are routinely ignored" (Priest & Arkin, 2010). While these intelligence reports are the results of analysis, they also become another information resource that analysts must use when performing future work. Even this database alone may be overwhelming for the analyst.

The crime, intelligence, and business domains have developed systematic approaches to help analysts handle such high volumes of information. In the next section (1.2.1), I describe a general systematic approach of analysis within these three domains, outline the steps involved, and detail difficulties that must be overcome to be successful.

1.2.1 ANALYSIS APPROACHES

It is important to understand how an analyst thinks and the general methods an analyst may employ when conducting crime, intelligence, or business analyses. There are several methods of rational thought relevant to performing intelligence analysis. Many analysts engage in some combination of four "types" of rational thinking when analyzing information: deductive, inductive, intuitive (or abductive), and scientific (Krizan, 1999; Clauser & Weir, 1975). First, analysts may use *deductive* reasoning when starting from general principles and applying them to a specific situation. Arthur Conan Doyle's Sherlock Holmes and Agatha Christie's Hercule Poirot provide examples of deductive reasoning. In contrast, *inductive* reasoning starts with facts and moves towards specific hypotheses. *Intuitive* reasoning relies on intuition and past experience to interpret information. And the *scientific* approach to analysis, first proposed by Sherman Kent in the 1930s, combines both deductive and inductive reasoning: inductive reasoning is used to develop hypotheses, while deductive reasoning is used to test those hypotheses.

There are also competing methodologies that analysts may use to help systematically explore vast quantities of information. The four main methodologies of traditional intelligence analysis include: opportunity analysis, analogy analysis, lynchpin analysis, and the analysis of competing hypotheses (Krizan, 1999). Opportunity analysis focuses on potential opportunities and vulnerabilities to an organization's agenda, and how to exploit opportunities while protecting against vulnerabilities (Davis, 1992). Analogy analysis, on the other hand, examines a specific

item in order to understand and recreate it with existing information or raw materials (Krizan, 1999). More relevant to this dissertation, however, are linchpin analysis and analysis of competing hypothesis.

Linchpin analysis structures the analysis process in accordance with critical intelligence failures. Championed by Douglas MacEachin (1994), Linchpin analysis focuses on managing assumptions, uncertainty, and misconceptions, by focusing on a highly-structured exploration of information. Richards Heuer (1999) later developed a popular method known as the analysis of competing hypotheses (ACH). Drawing on Simon's concept of "bounded rationality," Heuer maintains that analysts are intrinsically biased and that these biases produce poor analyses. He develops a systematic approach that focuses on understanding assumptions and uncertainties while gathering many competing or different perspectives and hypotheses. ACH involves developing a set of plausible hypotheses and listing the evidence and arguments for and against each hypothesis; it focuses on disproving rather than proving a hypothesis, and on understanding possible outcomes or consequences if the available evidence is incorrect or misleading.

While these four approaches to rational thinking and four methods for performing analysis shed light on the increasingly complex task of analysis, these are also very high-level, theoretical notions of the analyst's task. In the following section, I outline specific, practical steps that are common to the analyst's task.

1.2.2 BASIC TASK PROCESS

No matter what type of rational thinking or analysis methodology an analyst chooses, there is a series of recommended steps to structure the analysis process (Heuer, 1999; Krizan, 1999; Kight, 1996). Any process of analysis is a multi-sequence process, iterative at times, especially due to the constant collection and development of new information (Poole, 1981; Dearth, 1995). There is no one right path, but there are critical tasks involved in analysis that can improve overall performance.

Traditional views of intelligence analysis process, for example, place analysts within the context of their organization and the policy implications of that organization. Krizan (1999) presents an overview process of intelligence creation and use, illustrated in Figure 1.1. Treverton (2001) describes a similar process that illustrates how policy implications influence the search for data
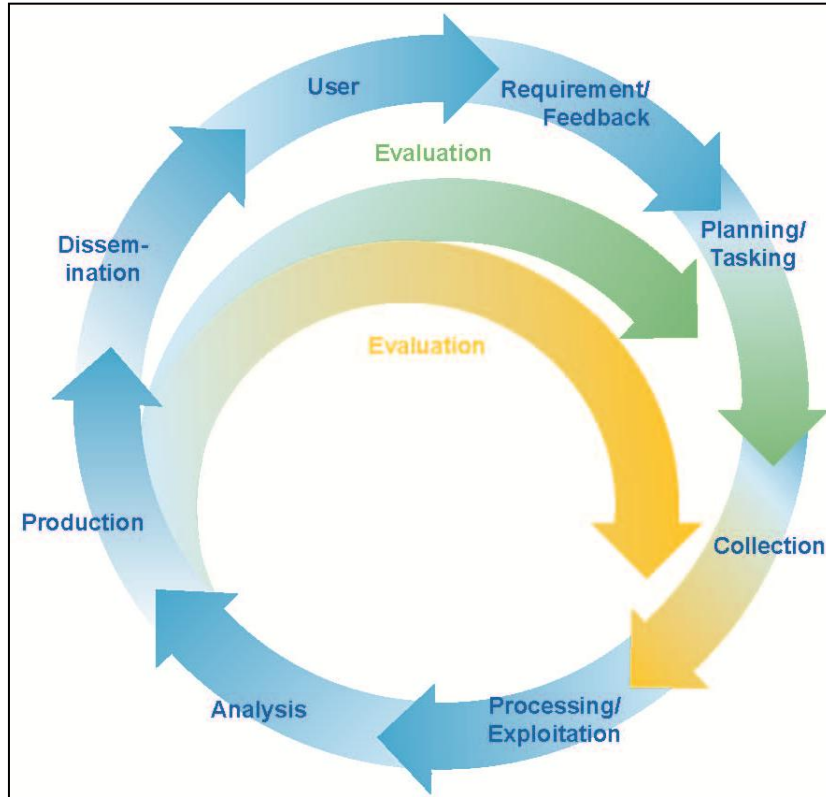
Figure 1.1. Process of Intelligence Creation and Use. (Krizan, 1999, p. 8)
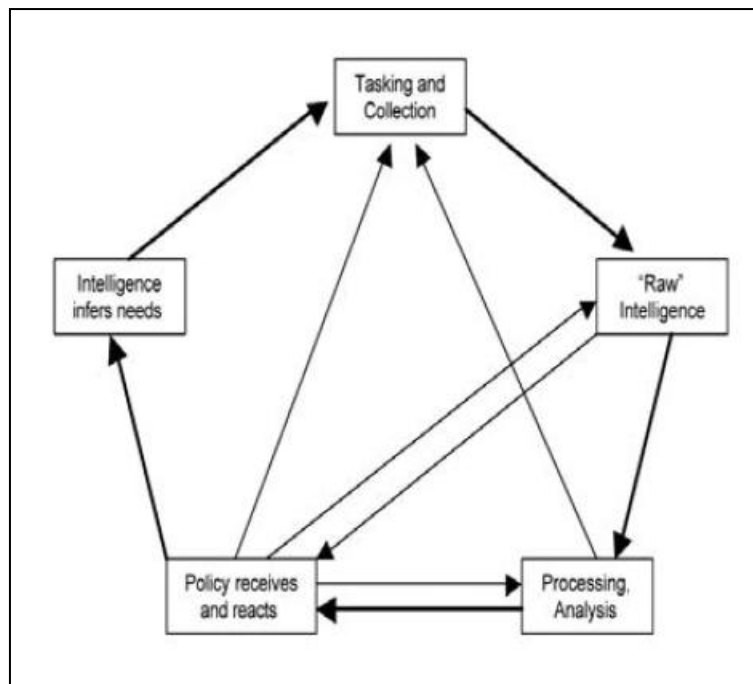
Figure 1.2. Treverton's intelligence cycle

and the need for analysis (See Figure 1.2). Both process models emphasize the iterative nature of intelligence analysis. The user obtains requirements from a client or management and begins to plan the task. The analyst then starts to collect data or information necessary to perform the task. After data are collected, the analyst can begin to process data and analyze them. Once the information has been broken down into its logical meaning, the analyst can synthesize those data, producing intelligence based on whatever information has been gathered and analyzed. Finally, once the narrative or intelligence has been created, the analyst must disseminate the product, most often in the form of a briefing or report, to relevant users.

However, it has been argued that such models of analysis are inadequate and do not capture the complexity of the intelligence process (Johnston, 1997 & 2005). After an extensive ethnographic study of intelligence analysts, Johnston developed a systems model that depicts the complex interactions and dependencies within the intelligence process (See Figure 1.3; for a detailed walk-through of this model, please visit chapter four of Johnston, 2005). Johnston takes into
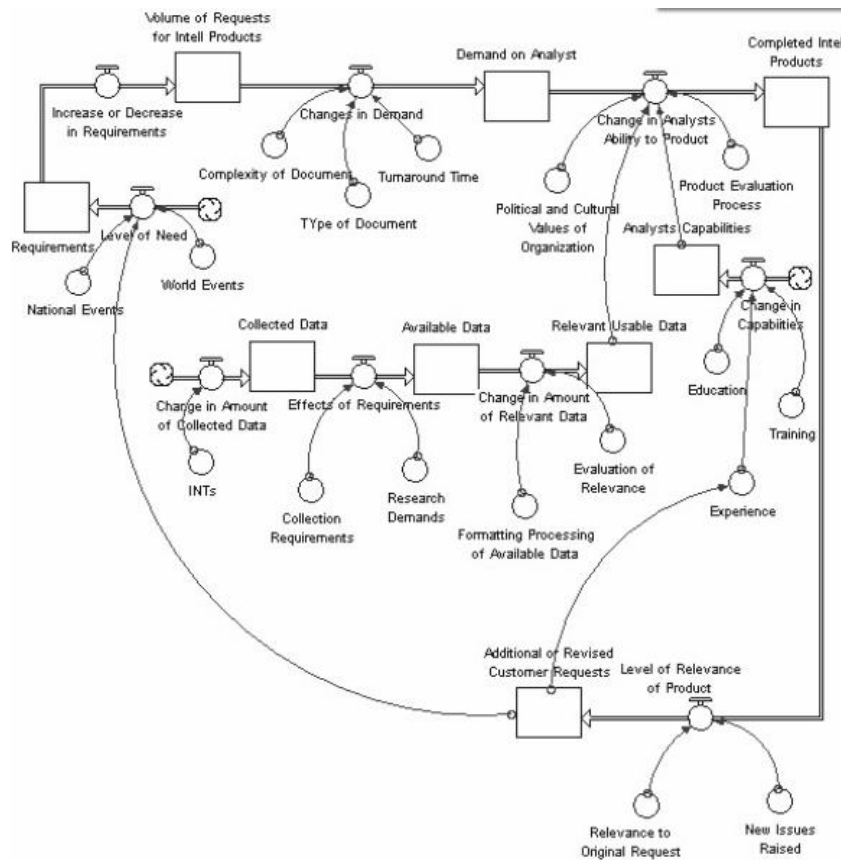
Figure 1.3. Johnston's systems model of the intelligence cycle (Johnston, 2005).

account how time constraints (called "stocks") and the need for information change based on current events or needs (called "converters"). These link together (via "connectors") to form dependencies and assumptions, which shape the "actions" or "flows" that control the nature of stocks and dependencies. Such a system emphasizes the constantly-changing environment and constraints that analysts must adapt to.

These models account for the fact that an analyst works within the context of a complex organization, and draw attention to factors that greatly impact the types of questions they seek to answer, as well as how they plan to disseminate their analyses. In this dissertation, I acknowledge the importance of such models, but I focus on components central to the search, processing, and production of actionable intelligence. With respect to the Krizan model, then, my focus pertains to only the lower half of the process.

Collaborative analysis task model

I have developed a model to describe the basic collaborative analysis task process (See Figure 1.4). The process in this model moves from left to right. On the left-hand side, the model illustrates two analysts working together at the start of the task, leading to an eventual decision on the right-hand side. The path to a final decision, however, is not straightforward.

There are two main components to this task model. The first component is *sensemaking* (Pirolli
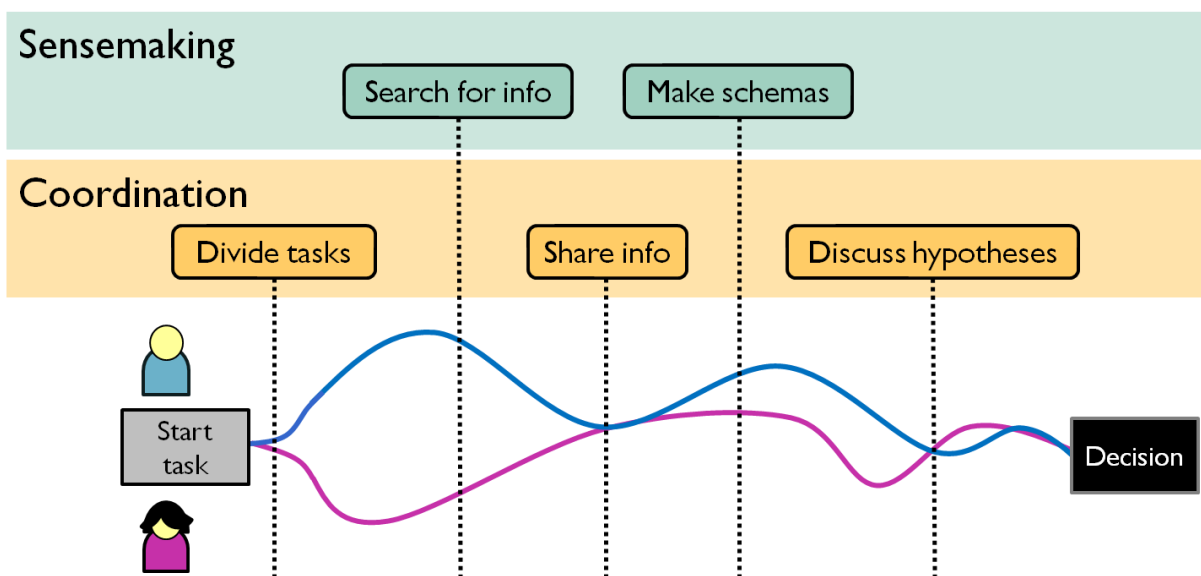
Figure 1.4 Collaborative analysis task model

& Card, 2005). Sensemaking is the process of finding meaning from information (Weick, 1995). Pirolli and Card describe several steps of the sensemaking process, including information foraging, or searching for information, as well as information schematization, or creating schemas that combine information into a cohesive narrative. This is an iterative process. Remember the case of the Vancouver serial killer? Detectives Dickson and Rossmo's sensemaking process involved searching for facts, then finding a pattern that fit all the information they had.

In addition to sensemaking, collaborative analysis requires an additional task: coordination. Coordination is dividing up tasks to help distribute labor, sharing information, and then discussing hypotheses to come to a consensus. This process is difficult for most collaborators and there are many ways to fail. In the following sections, I describe sensemaking and coordination in greater detail.

1.3 SENSEMAKING

An analyst's sensemaking process can be categorized into three main phases: information foraging, information schematization, and decision making and dissemination (Pirolli & Card, 2005; Heer & Agrawala, 2008). Figure 1.5 illustrates the sensemaking loop; it highlights the
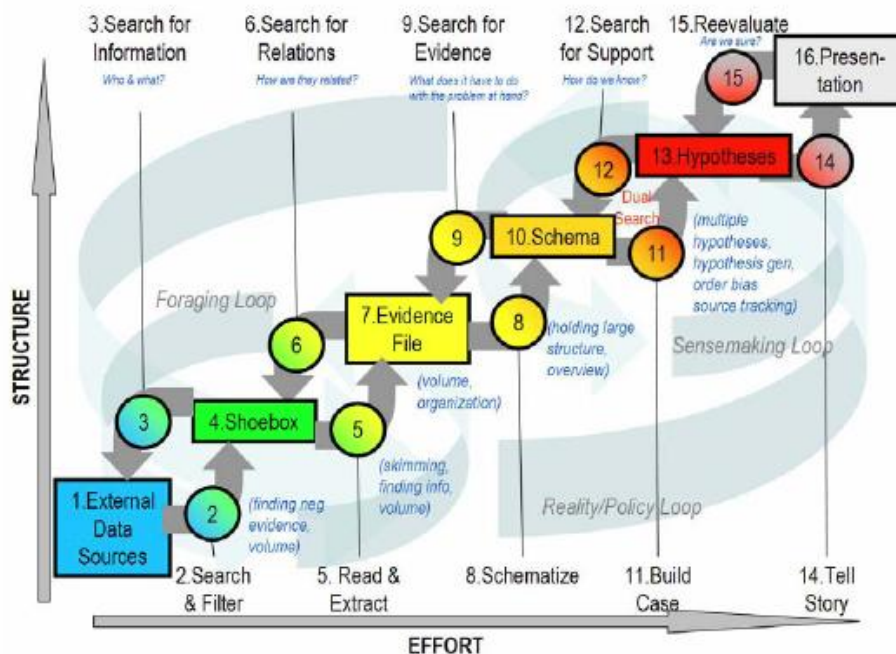
Figure 1.5. Sensemaking loop for intelligence analysis (Pirolli & Card, 2005).

iterative nature of the sensemaking process. For example, oftentimes business analysts are assigned a particular corporate competitor to follow (Kahaner, 1997; Balakrishnan, Matthews, & Moran, 2010). The analyst will maintain a basic level of general awareness and understanding of that corporate competitor. This process involves searching through large amounts of information via publicly-available resources, such as company websites, business magazines, and newspapers. Once the analyst has collected enough information, he will then analyze the information and try to construct "meaning" from the data. If he has been tasked to answer a specific question (e.g., how does the competitor market share appear to be growing?), he will then generate a series of hypotheses based on information currently available to him. Next, he will evaluate each hypothesis, which may involve returning to the information search stage, wherein he tries to confirm or disconfirm evidence for those hypotheses. Finally, he will decide among his remaining hypotheses and disseminate at least one as an answer to the question at hand. This may *seem* straightforward, but this process is extremely complex and deserves the more detailed treatment that follows.

1.3.1 INFORMATION FORAGING

The first major phase of sensemaking is to forage for information (Pirolli & Card, 1995, 1999). Information foraging involves searching for information and filtering it for relevancy, reading and extracting critical pieces of information, and forming a basic schema that organizes information into some cohesive structure. Information foraging theory oftentimes likens the human who seeks out information to an animal who seeks out food. Just as there are constraints on the animal and their abilities to find the right food, there are also constraints on the human and their abilities to find the right information. These factors include: time and the amount of information available, along with limiting forces, such as how much information an individual needs to come to a decision. The information forager selects relevant pieces in order to maximize the overall rate of information gain with respect to their specific task. The forager moves on to the next patch of information only once the available new information has been gathered, exhausted, or become "old."

When an analyst locates a new piece of information and consumes it, it impacts his next inquiry in an iterative way. Another way to imagine this is to think of "berrypicking." Bates' (1989) "berrypicking" approach to information seeking emphasizes the haphazard and dynamic nature
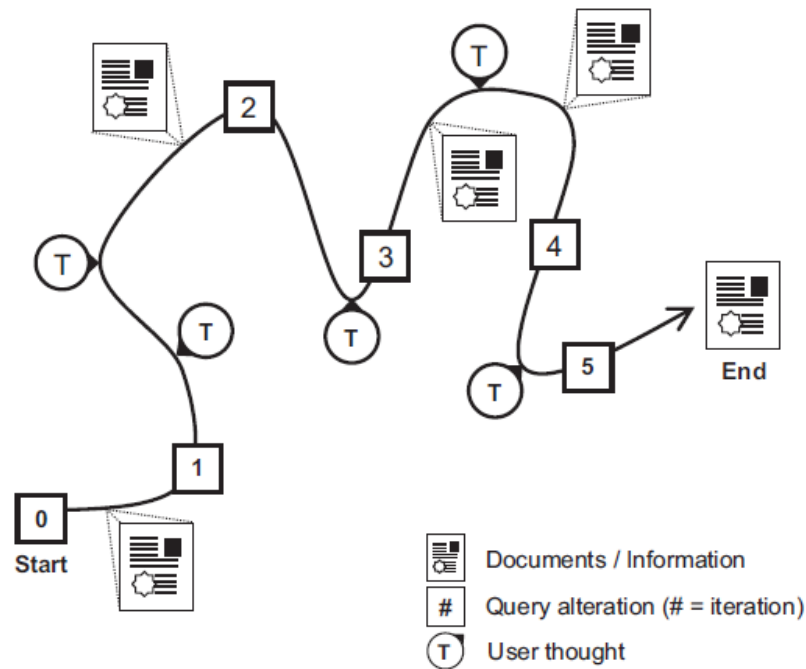
Figure 1.6. Berrypicking search process (Bates, 1989).

of information search. Berrypicking imagines the information seeker within an information space. As the analyst begins a search, he naturally has an initial inquiry that guides his search. An analyst of competitive intelligence, for example, might begin with the question "What is the next major product competitor X plans to release?" Information seekers may then change the direction of their search based on new information (See Figure 1.6). If the competitive intelligence analyst finds out that competitor X recently acquired a new company, he may refocus or narrow his search to only that acquired company. His new line of inquiry might go on to ask what this newly acquired company specializes in, etc. The analyst begins his search in an exploratory mode, hence the reason for his search in the first place. But the information the analyst desires may be unknown, and the direction of his search may change drastically over the course of the search (White & Roth, 2009).

1.3.2 INFORMATION SCHEMATIZATION

The second phase of the sensemaking process is information schematization. Information schematization is the construction of meaning based on discovered information (Pirolli & Card, 1995; Russell, Stefik, Pirolli, & Card, 1993). Like information foraging, the development of information schema is an iterative process. A schema is a mental model that weaves together

22

relevant pieces of information into a cohesive structure or narrative for the analyst. Analysts must generate representations of whatever information they seek, so that they know how to "capture" or "find" the right information. For example, an analyst may look for patterns and similarities between different entities (what is the common weapon used in a series of bank robberies?), or focus on anomalies and outliers within a dataset (what to make of an individual with ties to two different yet competing corporations?). As the analyst seeks out more and more information, he will try to incorporate or encode relevant new pieces within an already instantiated representation. But not all new pieces of information will be able to fit into his current representation. In order to accommodate information that does not fit, the analyst must alter or shift their representations. Additionally, once a representation has been instantiated, the analyst can still identify "missing pieces" or links of evidence needed to support their current representation, a process that naturally involves data-driven induction, and top-down deduction.

An analyst may also develop several information schematizations at the same time. Several different schemas are nothing more than several different hypotheses about what the available information *means*. Heuer's analysis of competing hypothesis, for example, encourages multiple representations in order to avoid common failures of analysis such as confirmation bias (Nickerson, 1998). Analysts have confirmation bias when they too easily find evidence that fits well into an existing hypothesis or schema, rather than trying to find disconfirming evidence that suggests a change or shift of representation. By analyzing information through multiple schemas, analysts have the advantage of competing hypotheses, which forces analysts to account for and incorporate many other types of information. In the case of our competitive intelligence analyst, he may find that competitor X's newly acquired company makes very small, very long-lasting batteries. The analyst could work from one schema or hypothesis wherein competitor X might want to reduce the size of its leading mobile device. But the analyst could also create a different schema, for example that the competitor might want to break into the tablet market.

1.3.3 DECISION MAKING AND DISSEMINATION
The final stage of the sensemaking process is decision making (i.e., what will become of the possible hypotheses?) and finally the dissemination of the analyst's results. The analyst must decide upon an appropriate schema or representation for the information and then disseminate it in a manner appropriate to his relevant audience. In decision making, an individual evaluates and

chooses a preferred option among alternatives (Jarboe, 1996). Decision making involves a choice between two or more alternatives, and so an evaluation of possible consequences on future events (Hastie & Dawes, 2001). Much like a jury member would, the analyst weighs evidence and information to confirm or deny possible hypotheses.

Because analysts also rely on intuitive thinking, their prior experiences directly impact their decision-making process. Prior experiences *can* potentially bring about better quality decisions, but that is not always the case. Individuals, for example, may value their prior experiences with features of an old situation more than they value facts, evidence, or information about a new situation. Because the analyst naturally has a salient memory and preconceived notions about what "makes sense," this sort of availability bias is difficult for an analyst to avoid (Hastie & Dawes, 2001; Tversky & Kahneman, 1983; Tversky & Kahnemen, 1973; Combs & Slovic, 1979). But Heuer's ACH method of analysis does try to mitigate such biases. If analysts can evaluate multiple hypotheses and find both confirming and disconfirming evidence for those hypotheses, then they might be able to better combat the influence of salient memories and the preconceived notions about what "makes sense" to them; in so doing, analysts should come to a better decision

Once he has formed a decision, the analyst usually presents his findings to his relevant audience. The form the analyst's output takes depends on the context he is imbedded in *as* analyst. In certain work environments, for example, analysts function as advisors; they are oftentimes not the ultimate decision makers. Analysts must distill their analysis into a form that is easily digested by ultimate decision makers. In most business and intelligence settings, then, analysis comes in the form of text briefings and presentations (Priest & Arkin, 2010). Another example of how output is mediated by context can be found in this dissertation itself. While this dissertation does not focus on the dissemination of analytical findings, it is important to be aware of the *form* these findings take. In general, presentations and briefings require the distillation of an analyst's research into a cohesive structure or narrative. All the while, the analyst crafts a story and offers information or evidence that helps bolster the believability and perceived accuracy of him and his story. The decision making and dissemination phase of sensemaking thus culminates in story building, wherein the analyst echoes the structures, narratives, schemas, and forms he foraged through to find his evidence and information in the first place.

1.3.4 BREAKDOWNS

Most analysts of crime, intelligence, and business undergo extensive training and this extensive training costs a lot, both in terms of money and time. The proliferation of knowledge in the "digital age," accompanied with its specialized domains, results in a shortage of analysts within these domains generally. Experts in crime, intelligence, and business are usually trained in a specific domain (i.e. East Asian economics, Yemini Islam extremism) and may have difficulty interpreting data outside the lens of their domain (Johnston, 2005). Confirmation bias and availability bias have already been discussed, but what else can lead an analyst astray during the sensemaking process?

Consider the case of the Vancouver serial killer. How was Mr. Pickton able to go on killing women for so many years? First, there may have been failures in information foraging. Detectives deal with information of all types, including cellphone call logs, still and video images, statistical breakdowns of crime patterns or trends, and text briefings or summaries of previous analyses from other analysts. As we make advances with technology, we have greater and greater access to more and more information. And yet it remains difficult to find useful information when needed (Edmunds & Morris, 2000; Eppler & Mengis, 2004). How can an analyst know what pieces of information to explore first if they are simultaneously unaware of *all* available research? This kind of "information overload" is difficult to overcome, even for analysts trained in a systematic methodology. Additionally, when an analyst finds an interesting piece of evidence, or several pieces of evidence that *seem* to be related, they could just as well be a false positive, and so lead the analyst on an irrelevant tangent, resulting in wasted time and effort.

Another reason Mr. Pickton might have been able to go on killing for so long is because of failures in information schematization. It is difficult to see patterns within diverse, seemingly disparate bodies of information, even for experts. Was the large number of missing women in Vancouver indicative of a real pattern, or was it just a statistical anomaly that *seemed* indicative of a real pattern? How many women needed to go missing before a pattern could be shown? This sort of fine combing of data requires patience. Analysts must construct new narratives that may not have been seen before; these narratives may resemble previous cases, but the pieces of evidence and the timeline are unique to the current narrative. The analysts responsible for

tracking missing women maybe did not see a persuasive narrative connecting the evidence available.

Finally, there may have been a failure in decision making, that is, a failure to correctly evaluate possible scenarios and hypotheses. If a single analyst cannot have access to all the necessary pieces of evidence, the analyst can still brainstorm a complete set of plausible hypotheses and systematically confirm or disconfirm them. But without the necessary information, an analyst may not have known to hypothesize that there was a serial killer on the loose in the first place. Without a plausible narrative, decision makers could not establish appropriate preventative measures against the attacks because it never occurred to them that any such preventive measure were needed. In this scenario, it seems impossible for any one analyst to be aware of all the potentially relevant information, and so impossible to represent the full multitude of possibilities. One hope of this dissertation, then, is to show that pairs or teams of analysts—working with visual analytic tools—can improve their opportunities for success with collaboration and coordination.

1.4 COORDINATION

Coordination takes collaboration, and vice-versa. Collaborative sensemaking already occurs in many domains, including healthcare, military, emergency services, and intelligence analysis (Paul & Reddy, 2010; Jenson, 2007; Weick, 1993; Landgren & Nulden, 2007; Johnston, 2005). What makes collaborative analysis process different from an individual analysis process? While the major tasks involved in the sensemaking process remain the same, collaboration requires additional coordination as well as communication-centered activities such as information sharing.

Collaboration here refers to two or more individuals working together to achieve a common goal. In particular, this dissertation focuses on collaborative analysis and the collaborative sensemaking process. It asks how do individuals working together search for information, build schemas, and decide upon a coherent narrative together? There is a wide variety of research on groups because there are a wide variety of groups (e.g., Shaw, 1973; Hackman, 1976; Steiner, & Rajaratnam, 1961; Davis, Laughlin, & Komorita, 1976; Levine & Moreland, 1990). But this dissertation focuses on concocted groups. The hope is that this research will have implications for natural, real-world groups (McGrath, 1984). In particular, I studied remote synchronous

pairs, or groups of two. These two individuals were brought together to perform a specific intellective task; they had a goal, worked synchronously (or in real-time as opposed to asynchronous shift work) and were separated by some geographic distance.

I argue that the sheer volume of information makes it impossible for even the best individual to perform well under time constraints, at least for the type of intellective tasks I study. In order to succeed, groups must take advantage of heterogeneity of subtasks required within a larger goal and also apply all necessary resources to reach their goal. One obvious advantage of working with a group is that the aggregated group-level knowledge is greater than the knowledge of any single individual (Argote, Gruenfeld, & Naquin, 2001). This idea has been made recently popular by Malcolm Gladwell's book, *The Wisdom of Crowds*, which extols the merits of aggregated group-level knowledge. Working with a collaborator *should* help analysts along the various sections of the analysis process. However, it should be pointed out that Steiner's (1972) models for group task performance predict that a group's potential productivity is higher than a group's *actual* productivity. Steiner concludes that this difference between potentiality and actuality is because of process losses incurred by motivation and coordination costs intrinsic to group work. Why, then, would I look to collaboration as a potential solution if in fact groups do less well when compared to an idealized potential productivity level? It is sometimes true that, in certain cases, the best member of a group will outperform that group if he had worked alone. But as McGrath (1984) points out, "we need to keep in mind that most or all of the rest of the members would have done worse if working alone" (p. 72). So what factors steer collaborations toward successful problem solving? And what factors steer collaborations toward failure?

Coordination is one of those factors. Coordination is "the act of working together" (Malone & Crowston, 1991) and "managing dependencies between activities (Malone & Crowston, 1994). Coordination can be thought of as an invisible force that makes an activity run smoothly. Successful coordination may go unnoticed, but unsuccessful coordination is obvious. For example, actors in a play coordinate their lines, delivery, blocking, and so on. When done correctly, an audience can forget that the events unfolding on stage are fiction. When done poorly, the audience clearly recognizes the staged presence. Successful teamwork depends on effective coordination. Successful coordination involves communication among partners, shared resources, a shared understanding of the group's goal, an agreed upon overall strategy or plan to

perform the task, assigned tasks, and an understanding of the dependencies between tasks (Malone & Crowston, 1994; Van de Ven , Delbecq, & Koenig, 1976; Cartwright & Zander, 1960; Fandt, Richardson, & Conner, 1990).

1.4.1 DIVISION OF LABOR

In particular, having multiple partners could aid analysis because it reduces the burden on a single individual through a division of labor and a shared awareness of parallel lines of work. With respect to information foraging, many people can forage through more information than a single person can. Collaborative information foraging, or collaborative information seeking, occurs when individuals come together "during the seeking, searching, and retrieval of information" (Foster, 2006, p. 330). During collaborative information seeking a group or team has a shared information need; it requires they find and share information between or among collaborators (Poltrock, Dumais, Fidel, Bruce, & Pejtersen, 2003). Collaborators can divide the search process between themselves. One strategy a group of analysts could use, for example, is to agree that each analyst will focus on a unique category (e.g., one analyst is assigned car bomb attacks and another attacks from suicide bombers). Or each analyst could be given a unique corpus of data and the team, as a whole, could agree to focus on a single category of corpus (e.g., on male suspects or on a specific type of weaponry).

Both these strategies depend on effective coordination and awareness of partner activities. If an analyst is unaware of specific searches his partner has performed, he may duplicate work. Additionally, awareness or insight into a collaborator's search process may help individual searchers refine their search process or provide guidance on search tactics (Morris, 2008; Morris & Horvitz, 2007). Knowing the progress of task collaborators can help avoid duplications of effort, identify whether collaborators are in need of additional help, and foster feelings of team cohesion (Carroll, Rosson, Convertino, & Ganoe, 2005; Mullen & Copper, 1994). Analysts must develop a shared mental model or "knowledge-in-common" as a helpful guide for coordination of their activities, which is important for group-consensus activities generally (Mohammed & Dumville, 2001).

A group's division of labor between tasks highlights the interdependent nature of collaborations. The success of individual depends on the success of the team and vice versa (Johnson & Johnson, 1985, 1981; Slavin, 1989; Deutsch, 1960). Individual workload is reduced only if all

team members perform their tasks and do not succumb to "social loafing," which is the inclination for individuals within a group to take advantage of the efforts of others by doing less (Harkins & Petty, 1982). When groups divide labor and directly assign tasks to individuals, they can also motivate group members with feelings of accountability, responsibility, and team cohesion (e.g., Weldon & Gargano, 1988; Seashore, 1954).

1.4.2 INFORMATION SHARING

Information sharing is not merely communication about coordination activities; rather, information sharing *is* communication pertinent to performing the task. For example, an analyst may share a clue he found on a potential suspect, or hypothesize that two cases are connected due to similarities between crimes. This sort of information sharing contributes to the development of a shared mental model, or collective knowledge, and a shared understanding of the problem at hand, which in turn improves group consensus activities (Mohammed & Dumville, 2001). Consequently, information sharing is critical to the success of intelligence analysis. When success hinges on serendipitously finding important facts, two collaborators may independently find discrete information whose value only becomes apparent when both those facts are shared between the two collaborators (Fraidin, 2004; Simonton, 2003).

Collective knowledge between collaborators depends on an individual's ability to recall information and whether or not the individual shares the information with their group (Larson & Christensen, 1993). Even the anticipation of working in a group affects individual recall; those who think they will be working alone remember more information than those who believe they will be making a decision with a group (Henningsen, Cruz, & Miller, 2000). Once collective knowledge has been created, the group must be able to recall and discuss the information as needed. Groups have an advantage over individuals in this as well. In comparison to individuals, groups recall more, are more accurate, and have fewer reconstructive errors (Hill, 1982; Clark & Stephenson, 1989).

Even if every member of the group has perfect recall, information sharing is still difficult to execute in practice. For example, each member of a team may have access to a unique set of facts or information, and this information may be distributed among various team members rather than to all the team members. Consider that each member of a competitive intelligence team may have information on a unique competitor, or that each member of an aviation security

group may be responsible for collecting evidence based on a unique location. If a member of a group has crucial knowledge or formulates the correct hypothesis or hypotheses, that knowledge is useless to the group until the member has shared it. In a series of seminal research studies on information sharing, Stasser and Titus (1985; 1987) found that, most of the time, information sharing is far from ideal. In these studies, four-member groups discussed fictional political candidates for student body president. Information on each political candidate was either shared equally *or* distributed among group members. Their results show that in distributed information situations, where group members have different pieces of information, groups tend to focus on shared or common information that supports consensus rather than on unshared information, which oftentimes leads to faulty decisions. Essentially, groups have a tendency towards confirmation bias, which is naturally brought about by their patterns of information sharing. These results have been replicated in various settings, including medical decision-making teams (Larson, Christenson, Abbot, & Franz, 1996; Larson, Christensen, Franz, & Abbot, 1998; Stewart, Billings, & Stasser, 1998; Stasser, Taylor, & Hanna, 1989).

When group members fail to exchange information with one another, the group may succumb to "groupthink." Groupthink describes a tendency in members of groups, who seek out and pressure other members into agreement rather than considering alternative opinions (Janis, 1982). One potential outcome of groupthink is "polarization." Polarization occurs when groups chose extreme or risky decisions that they would not have chosen if they were choosing alone. Groupthink may occur because of group discussion or because of a persuasive argument from a more vocal member or group leader. Another theory, however, the *social comparison theory*, maintains that groups become more extreme due to normative pressures from learning about other members' opinions and positions (Isenberg, 1986).

Introducing diversity of opinion may help avoid groupthink. In a laboratory study, three participants discussed a murder mystery case; those who were primed to have a counterfactual mind-set discussed more unshared information, which improved the group's ability to correctly identify the suspect (Galinsky & Kray, 2004). Motivating individuals to focus on epistemic reasoning also mitigates groupthink. A recent study on motivation showed that groups who had to be accountable for their process chose a more information-driven process for decision making;

this led group members to focus on unshared information and, eventually, to reach the correct decision (Scholten, van Knippenberg, Nijstad & De Dreu, 2007).

1.4.3 HYPOTHESES DISCUSSION

As an analyst gets closer to a final decision regarding the outcome of his query, he may begin to focus on several information schemas or hypotheses. Discussing hypotheses is the sharing of facts or clues; it involves the discussion and analysis of created schemas. In a series of studies, Laughlin and colleagues (1985; 1986) showed that groups were better than individuals at identifying the correct hypothesis once someone shared it with the group (i.e. they were better at spotting the correct hypothesis, and knowing it was correct). However, groups did not generate more hypotheses in comparison to individuals. The power of collective induction, then, is in the evaluation of a hypothesis rather than in the generation of distinct hypotheses.

Additionally, collaborators can strategize together to minimize this information-sharing component; for example, they can rely on members to share their own interpretations or hypotheses that members create on their own. Such a strategy does reduce the burden on individuals to share *every* piece of found information. However, relying on individuals to filter data with their own interpretation or hypotheses may result in a distortion of the actual facts (Hansen & Jarvelin, 2005; Harper & Sellen, 1995).

1.4.4 REMOTE COLLABORATIONS

Another difficulty with collaboration is that oftentimes analysts work in remote teams. Remote collaborations occur when two or more individuals, who are separated by some distance, work together towards a common goal. While remote collaborations are becoming increasingly popular, they do not always offer the same levels of productivity and success as collocated teams, who work side-by-side (Olson & Olson, 2000; Cummings & Kiesler, 2005; Cummings & Kiesler, 2007; Kiesler & Cummings, 2002).

Group awareness and communication patterns of teams also suffer as a result of distance (Kraut, Egido, & Galegher, 1988). In comparison to collocated teams, distributed teams have reduced opportunity for serendipitous "cross talk" and decreased awareness of task progress. Technology can help overcome some of these barriers by supporting information sharing, task structuring, and general communication (McGrath & Berdahl, 1998; McGrath & Hollingshead, 1994). For example, a study I performed with colleagues at IBM Research – Almaden found that a shared

information repository can help foster team cohesion, increase team awareness of task progress, and help reduce the burden of communication (Balakrishnan, Matthews, & Moran, 2010). One of the failures exposed by the Vancouver serial killer incident was the ineffectiveness of shared repositories within the police community. Other technologies, such as group decision-making systems (Poole & DeSanctis, 1989 & 1990; Poole, Holmes, & DeSanctis, 1991), can help teams manage conflict during decision making.

Many distributed and collocated teams rely extensively on communication tools to share information and coordinate tasks. Both synchronous communication tools, such as phone, video conferencing and instant messaging (IM), and asynchronous tools, such as email, have come to dominate work practices. While these tools are less rich than face-to-face conversations, IM and email are effective modes of communication for distributed teams (Walsh, Kucker, Maloney, & Gabbay, 2000; Bradner & Mark, 2002). However, this effectiveness often depends on the *type* of task. For example, Hollingshead and colleagues found that computer-mediated groups performed as well as face-to-face groups during decision making and task generation, but not as well during negotiation and intellective tasking (Hollingshead, McGrath, & O'Connor, 1993). Such findings suggest that present communication tools do not adequately support remote collaborators, and that remote collaborators may need additional support on highly cognitive tasks.

1.4.5 HELPING TEAMS BE SUCCESSFUL: AN OPPORTUNITY FOR TECHNOLOGY
The advantages of collaboration on intelligence analysis, coupled with the enormity of intelligence tasks, necessitate the use of teams for collaborative analysis. Again, effective teams require coordination, open communication and information sharing, common goals, and shared mental models (Cannon-Bowers, Salas, & Converse, 1983; Orasanu, 1990). This effectiveness is not so easy to develop and maintain, especially when dealing with remote collaborators. Deciding how to divide tasks, or coming to a common mental model of the situation, may be easier for teams with repetitive tasks and clearly defined roles. One example of this is surgical unit, wherein a surgeon, anesthesiologist, and various nurses work together to perform many surgeries in a single day. But this level of effectiveness is extremely difficult to achieve in fields such as intelligence analysis because oftentimes the goals or final solution in intelligence analysis is vague and unknown, while the amount of information to be analyzed is large, possibly

disorganized, and under constant growth. Without these organizing and structural factors, team performance will suffer. As Johnston (2005) describes:

> Without specific processes, organizing principles, and operational structures, interdisciplinary teams will quickly revert to being simply a room full of experts who ultimately drift back to their previous work patterns. That is, the experts will not be a team at all; they will be a group of experts individually working in some general problem space. (p. 70)

But the situation is not hopeless. Technology can help collaborations be more effective. For example, remote collaborations can benefit from technologies such as basic email and instant messaging, but also from more sophisticated video-conferencing applications. There are also tools that help groups navigate more difficult, decision-making tasks (Karacapilidis & Papadias, 2001). In this dissertation, I explore technologies with a potential to aid the information foraging and schematization processes of group analysis; here, I focus specifically on the potential for visualizations to improve collaborative analytics.

1.5 VISUALIZATIONS AND VISUAL ANALYTICS

Visualization techniques are representations of complex numerical and textual data in pictorial or graphical form; ideally, they allow for easy exploration of data (Andrews & Heidegger, 1998; Shneiderman, 1996; Wattenberg, 1999). By removing the burden of mentally consolidating disparate information, holistic representations of large amounts of data can help individuals spot anomalies, perceive patterns, and thus improve their problem-solving success (Larkin & Simon, 1987). Information visualization tools reduce task completion time and increase productivity on many information retrieval tasks in data analysis (Hendrix, Cross, James, Maghsoodloo, & McKinney, 2000; Stasko, Catrambone, Guzdial, & McDonald, 2000; Veerasamy & Belkin, 1996).

Although considerable research on visualization tools has been carried out in the academic community *and* in the private sector, most of that research has been directed towards building new types of visualizations. It still remains unclear how tools help collaborators during specific analytical tasks. This dissertation explores when visualizations aid coordination and sensemaking by exploring the behavioral components of using visualizations during collaborative analysis.

1.5.1 VISUALIZATIONS FOR SENSEMAKING

The phrase "a picture is worth a thousand words" has become cliché because of a simple visualization can make it possible to understand, or make sense of, a very large amount of data very quickly. Visualizations help facilitate sensemaking or the interpretation of large amounts of data (Edelson, Pea, & Gomez, 1996; Wood, Wright, & Brodlie, 1997). Visualizations aid the subprocesses of the sensemaking process, including collaborative data exploration (Pang & Wittenbrink, 1997; Brewer, MacEachren, Abdo, Gundrum, & Otto, 2000; Börner, 2001; Lascara, Wheless, Cox, Patterson, Levy, Johnson, & Leigh, 1999; Sawant, Scharver, & Leigh, 2000). Popular websites such as Many Eyes, sense.us, and Swivel attempt to make visual data exploration a social process. These distributed, asynchronous, collaborative visualizations can encourage knowledge discovery (Heer, Viégas, & Wattenberg, 2007). In the cases of sense.us and Many Eyes, which are websites designed for asynchronous social-data exploration, users are motivated by data-driven exploration *and* social-data exploration via comments from other members. These collaborative visualizations also can promote feelings of community and foster discussion in "wiki" websites (Viégas, Wattenberg, & Dave, 2004). Having access to collaborator comments encourages individuals to view multiple perspectives and build upon a diverse set of insights (Heer, Viégas, & Wattenberg, 2009). Novak and Wurst (2005) explore a collaborative community visualization that allows users to view visual states created by computer algorithms, fosters self-exploration, and presents relationships between concepts by different users. Exploring new perspectives may, in turn, lead analysts to develop a more numerous and more diverse set of hypotheses, and to avoid cognitive pitfalls such a confirmation bias (Billman, Convertino, Shrager, Pirolli, & Massar, 2006).

Additionally, certain tools allow analysts to save versions or "states" of analysis. This enables them to explore data without losing previous analysis (Ellis & Groth, 2004; Gotz, Zhou, & Aggarwal, 2006; Palantir, n.d.). Recording paths of analysis or "action trails" may encourage reuse of data, which may lead to higher levels of analysis and a more thoroughgoing construction of meaning (Shrinivasan & van Wijk, 2008; Shrinivasan & van Wijk, 2009). Collaborators can share static, annotated images of their data exploration process with their group members. Such shared annotated visuals can offer insight onto what collaborators are doing, and how they are doing it. Additionally, collaborators build a shared understanding via shared external representations (Qu & Hansen, 2008). Having a shared object that both partners can easily

reference promotes joint attention and may help in the establishment of common ground (Kraut, Fussell, Brennan, & Siegel, 2002; Mohammed & Dumville, 2001; Monk, 2003). Similarly, sharing annotations with partners gives individuals access to a new set of perspectives, allowing collaborators to leverage a diversity of opinions and analysis tactics (Ellis & Groth, 2004). Indeed, many visual analytic tools support the sharing of annotated visual states (e.g. General Dynamic's CoMotion Discovery; Tibco's Spotfire, Visual Analytic's VisualLinks, and DataClarity).

Another type of collaborative visualization are those that support information schematization (Brennan, Mueller, Zelinsky, Ramakrishnan, Warren, & Kaufman, 2006). For example, Sandbox allows users to create concept maps based on an evidence search process (Wright, Schroh, Proulx, Skaburskis, & Cort, 2006). In GeoTime, users are encouraged to create narratives around the evidence they have extracted (Eccles, Kapler, Harper & Wright, 2007). Creating a narrative or story enables individuals to more easily communicate and share their findings with collaborators. Finally, EWall allows team members to individually explore shared datasets and then combine relevant findings in a common space (Keel, 2006 & 2007). This tool aims to integrate the findings of many group members and highlights relevant information that may have been found by multiple users but otherwise overlooked.

Studies show that visualization tools can improve collaborative analysis (Mark, Carpenter, & Kobsa, 2003b; Mark, Kobsa, & Gonzalez, 2002). While sophisticated tools that aid collaborative, investigative analysis exist (Stasko, Görg, & Liu, 2008; General Dynamics's CoMotion), little research has been done examining how they affect the very collaborative processes they seek to improve (Billman, Convertino, Pirolli, Massar & Shrager, 2005; Johnston, 2005; Tolcott, Marvin, & Bresoick, 1989). Even fewer researchers offer any systematic evaluation of the effectiveness of these tools against common, collaborative, problem-solving pitfalls.

A notable exception is Convertino et al.'s (2008) evaluation of the CACHE system. The CACHE system was designed to lead analysts through a systematic analysis process, and to help analysts avoid common decision biases, such as confirmation bias. While the tool helped teams that possessed members with an initially diverse set of beliefs (i.e. heterogeneous groups), groups that already held similar beliefs (homogeneous groups) performed poorly, maintaining and strengthening their initial biases. In another example, Cook and Smallman (2007) found that

graphical layouts of evidence could mitigate confirmation bias and improve performance for collaborative analysts but this improvement depended on each member of the group reviewing pieces of information individually. This dissertation aims to build upon such work in an effort to further understand when and how such tools offer improvements or detriments to collaborative problem solving.

1.5.2 VISUALIZATIONS FOR COORDINATION

Visualizations can also serve as a mechanism to facilitate coordination during collaboration. Visual cues can increase awareness of collaborator activities and encourage a sense of team cohesion. For example, Scupelli and colleagues examined Project View IM, an instant messenger tool that provides visual information regarding a task and a partner. They found that such visual information could improve the subjective feelings of workload by reminding members working in distributed teams of tasks, as well as an increased awareness of their partner's activities (Scupelli, Kiesler, Fussell, & Chen, 2005). Hill et al. (1992a) shows how visual representations of a document's "wear" can aid collaborative document processing without any additional work from group members. Visualizing what segments of a document have been edited more than others, for example, helped collaborators understand how a document evolved, and what sections of the document required the most effort.

In a set of studies, Joan DiMicco and her colleagues (2004, 2007) examined how shared representations of group participation affect group dynamics and group decision making. They found a simple display that depicts how much each individual participated in the group conversation could help over-participators reduce their contributions, and also minimize the amount of non-critical information shared between the group members. CoSense, a collaborative web search tool, allows co-searchers to view histories of member searches (Paul & Morris, 2009). Individuals were then influenced by the behavior of their collaborators. Some group members were inspired to try new avenues of inquiry, or avoided particular search options altogether because they knew what already been attempted unsuccessfully.

While such tools are no doubt a step in the right direction, questions still remain unanswered as to how basic collaboration components are altered by the introduction of visualization tools into a group intelligence task. Insights by researchers studying network intrusion analysts highlight potential limitations of visualization tools. Goodall et al. (2004) documents how visualizations

tools used by network intrusion analysis fail to support the collaborative needs of network analysts. While exploring the task process of network intrusion analysts, Thompson and her colleagues (2006) lament the limitations of visualizations for analytical tasks:

> Our research suggests that network security engineers will continue to use the textual resources despite advances in data visualization. Textual resources are often rich with detailed information critical to understanding the context of the attack, whereas visualization tools tend to present an overview of the data.

Further, the Thompson and her colleagues warn "visualization tools have merely added to the plethora of existing resources that engineers need to search and sift through each day." It is critical, then, that we understand how such tools interact with both the sensemaking and coordination processes during analysis tasks so that we can ultimately design more effective visualization tools.

1.6 DISSERTATION CONTRIBUTIONS

This dissertation contributes to a fundamental understanding of collaborative problem solving in domains such as criminal and intelligence investigations and business intelligence analysis. Fundamentally, it explores how network visualizations improve complex collaborative analysis, and whether or not they do in fact improve them. This research also impacts studies in Human Computer Interaction (HCI), as well as small group research theory and investigative analysis. It also provides design implications for collaborative visualization developers.

Specifically, this dissertation's contributions are:

1. A better understanding of how and when network visualizations improve remote collaborative problem solving. Specifically, I highlight external factors in collaborative problem solving, such as information access, which influences the effectiveness of visualizations.

2. A better understanding of how visualizations impact both cognitive and social processes during remote collaborative problem solving. I offer in-depth analysis of the problem solving process for paired groups in controlled laboratory experiments, which adds to the understanding of how collaborators develop problem solving strategies, make sense of different types of information, and deal with "information overload." I develop a model

of the collaborative problem solving process and build upon that model to uncover
failures and possibilities for intervention.

3. A set of measures and analysis techniques that assess the impact of visualization tools on problem solving collaborations. I explore features of visualization use, communication, and solutions that uncover critical insights into the underlying mechanisms of the collaborative problem solving process.

4. A set of design implications that support visualization tools for collaborative analysis.

1.7 DISSERTATION OVERVIEW

In the chapters that follow, I describe three laboratory studies that explore collaborative problem solving with visualizations and reflect upon the implications and contributions of this research. Chapter two details my first laboratory study, which explores the effect of network visualizations, with varying degrees of sharing capabilities, on collaborative problem solving. I find that visualizations *do* improve collaborative performance. But this improvement depends on having access to a shared visualization because it encourages discussion between partners. Chapter two also introduces the "detective mystery paradigm" used in all three of my laboratory studies. In Chapter three, I describe my second study that examines the effect of shared and distributed information and visualizations on collaborative problem solving. This second study finds that the visualization was unable to improve performance when partners were overloaded with information; simply put, it does not encourage or introduce positive collaborative behavior. Chapter four describes my third study, wherein I explore the impact of discussion prompts as interventions for the collaborative process. I find that if collaborative processes are properly facilitated, visualizations can once again improve performance—even with information overload. Finally, chapter five offers a summary of my main findings, discusses future directions for research, and lists this dissertation's contributions.

2 DO VISUALIZATION TOOLS HELP COLLABORATIVE ANALYSIS?[1]

This first study explores the impact of network visualizations on collaborative problem solving. Using a detective mystery experimental paradigm developed by Scupelli et al. (2005), remote pairs worked synchronously to identify a serial killer hidden within multiple crime reports. They discussed disparate evidence distributed across the pair using IM. Four conditions, respectively, offered (a) spreadsheet only (controls), (b) individual unshared visualizations, (c) view-only shared visualizations, and (d) a full-access shared visualization of all evidence. I examined collaborative performance, use of the visualization tool, and communication as a function of condition. *All* visualization conditions improved remote collaborators' performance over the control condition. Full access to a shared visualization best facilitated remote collaboration because it encouraged tool use and fostered discussion between the partners. Shared visualization without full access impaired performance somewhat because it made communication even more vital to identifying the serial killer. This study provides direct evidence that visualization tool features and partner behavior promote collaboration.

2.1 INTRODUCTION

The focus of this study is problem solving, wherein successful task performance, as in the case of the Vancouver serial killer, depends on whether individuals share information crucial to a group's ability to "connect the dots." In the case of the Vancouver serial killer, a critical step that helped lead Detectives Dickson and Rossmo find a pattern for missing women in the Downtown Eastside district was the sharing of information and case files among the British Columbia Missing Women task force. In many domains, such as intelligence analysis (Heuer, 1999), business innovation (Baron, 2006) and scientific research (Klahr & Simon, 1999; Simonton, 2003), success may hinge on whether collaborators share information. In the following chapter, I

---

[1] The material presented in this chapter has been previously published as Balakrishnan, A.D., Fussell, S., Kiesler, S. (2008). Do visualizations improve synchronous remote collaboration? *Proc. CHI 2008.* NY: ACM.

argue that advances in computing that allow collaborators to visualize information create new opportunities for collaborative problem solving that have failed in the past.

2.1.1 ANALYSIS AS A COLLABORATIVE TASK

Two defining attributes of real-world complex analysis are that it is ill structured (in the sense that the problem definition is unclear) and that it often involves knowledge or information dispersed across people and groups (Klahr & Simon, 1999; Simonton, 2003). For instance, a detective in Vancouver investigating the possibility of a serial killer may sift through local "cold cases" looking for linkages but, unknown to this detective, the relevant cases may exist in other nearby cities such as Seattle (Safarik, Jarvis, & Nussbaum, 2000). Because criminal investigations need insight, and because relevant information is sometimes widely dispersed, the success of criminal and intelligence investigations, scientific discovery, medical problem solving, and other important real world problems often depends on opportunistic cross-talk between information sources and serendipity (Fraidin, 2004; Simonton, 2003). Collaboration can increase the likelihood that such cross talk and serendipity will occur, and increase group performance over that of individual's performance in these situations (Hill, 1982), but effective collaboration may also depend on the free flow of information among partners (Lavery, Franz, Winquist, & Larson, 1999; Stasser & Titus, 1987).

Computer-based visualization tools that support scanning for hidden linkages and the sharing dispersed information now exist. The question is: do these tools in fact change analysis strategies such as information sharing, collaborative relationships among partners, and ultimate collaborative task performance. I studied a network visualization application similar to those used in intelligence analysis and criminal investigations (for example, Analyst's Notebook, 2010). In preliminary studies, participants either worked on an analysis task individually or with a partner through Instant Messaging (IM). A network visualization tool improved analysis overall, but collaborative analysis was less successful than individual analysis. This result suggests that we need to learn more about the process of analysis when collaborators use visualization tools so that tools can be improved to overcome coordination costs and cognitive biases. The experimental design tested whether the network visualization tool improved collaborative task performance of remote partners who were synchronously solving a complex

analytic problem, and the extent that sharing features in the tool affected the overall success of the collaboration.

2.1.2 INFORMATION VISUALIZATION IN COLLABORATION

Previous studies have shown that visualizations can facilitate information sharing in collocated groups (Edelson et al., 1996; Ryall, Forlines, Shen, & Morris, 2004). Mark, Carpenter, and Kobsa (2003a), have studied visualization in collaboration, and showed that collocated pairs' and remote pairs' use of visualization tools for making bar graphs of statistical data improved their analysis performance over that of participants using the tools alone. This work builds on their promising results, examining *how* visualizations aid collaboration.

Visualization tools could aid collaborative analysis in at least two ways. First, if each member of a group has a visualization of his or her own data, then the individual member's insight into the problem may improve; this would raise the probability of better group performance. If this were the case, visualization tools might not need to provide for jointly viewable or manipulated data, or even promote discussion, as long as they improve the problem solving of individuals in the group. This idea leads to the following general hypothesis:

Hypothesis 1: Access to a visualization tool will increase remote pair performance in synchronous complex problem solving.

Second, prior research suggests that visualization tools may improve collaborative performance because they allow for shared access to data, and encourage information sharing and discussion. In their evaluation of CACHE—a system that supports intelligence analysis via visual data presentation—Billman et al. (2005) reports that distributed pairs using CACHE collaboratively did overcome *a priori* biases, resulting in more effective data analysis. Mark et al.'s (2003b; 2002) video analyses of experimental data showed that remote pairs using a visualization communicated more intensively than collocated pairs. Their results suggest that communication is necessary to take best advantage of visualization tools. This idea leads to a second hypothesis.

Hypothesis 2: Access to a visualization tool will increase remote pair performance in complex problem solving when this access increases information sharing and discussion by the pair.

But how fully does a visualization tool need to support shared information and communication? Visualization tools can support different levels of sharing. At the most basic level, each member can visualize his or her own data but cannot see other collaborator's visualizations (Unshared Visualizations). Alternatively, collaborators might be able to view their own and others' visualizations but could only directly manipulate their own (Shared View-Only Visualizations). Many applications can already be shared in this manner. A third possibility, however, is that collaborators have *full access* to a shared visualization application, which allows them to view everyone's data and to jointly manipulate that data (Shared Full-Access Visualization) (Pang & Wittenbrink, 1997). Full access would support shared information sharing automatically; this might be especially important when collaborators perform analysis without the same data (Billman et al., 2005). Full access could also promote joint attention, which may help establish a common ground more so than applications that allow for only shared views (Kraut et al., 2002; Kraut, Fussell, & Siegel, 2003; Monk, 2003). This last idea leads to a third and final hypothesis.

Hypothesis 3: Access to a shared full-access visualization tool will encourage discussion between partners, and it will increase remote pair performance in complex problem solving more so than the performance of partners using a visualization tool that only supports unshared visualizations or shared but view-only visualizations.

2.2 METHOD
This study is a single-level factorial, with four visualization conditions. Participants worked in pairs. Pairs were randomly assigned to one of the four visualization conditions. The pairs were told they were members of the homicide unit of a local police department, and that they had been assigned to a task force responsible for the capture of a serial killer.

2.2.1 PARTICIPANTS
Ninety-four participants were recruited for a "Detective Mystery Study" using an online-participant recruiting website (54 female, 40 male; 81% U.S. born; age range 18-64, median age approximately 22). Eighty percent of the participants were undergraduate or graduate students. Participants were paid $15. They were told the experiment would last 1.5 hours.

2.2.2 PROCEDURE
For the duration of the experiment, participants were seated separately, such that they could not see or hear their partner or their partner's workstation. Participants assumed the roles of a pair of
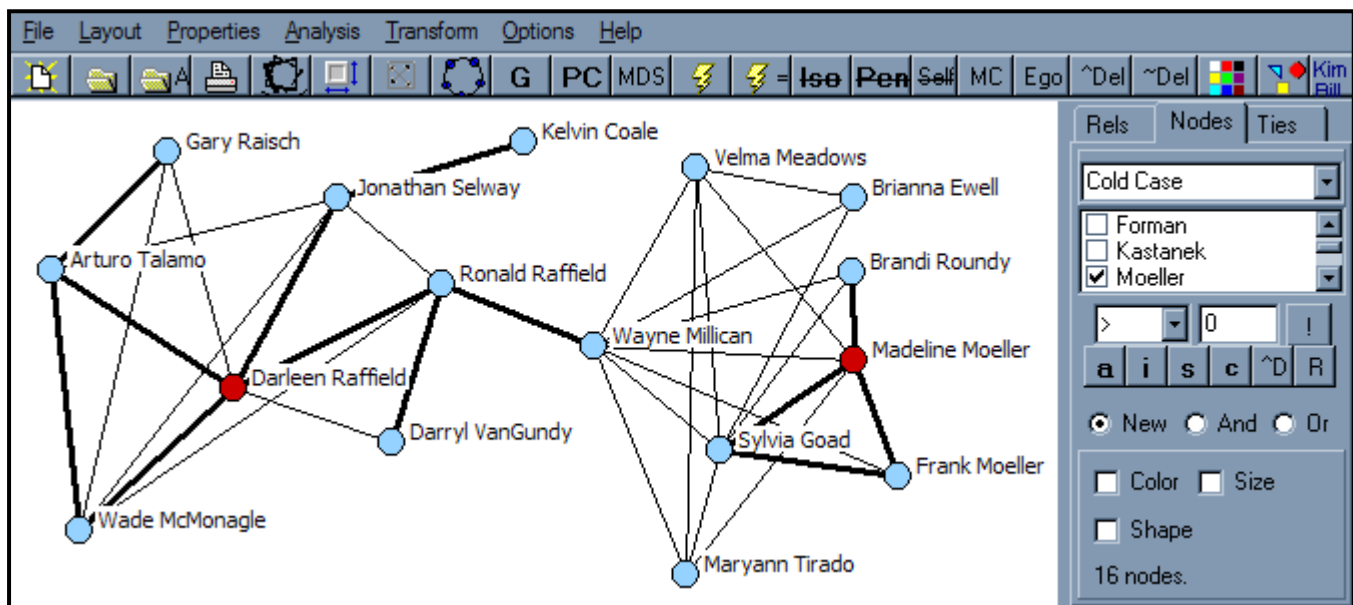
Figure 2.1 Screenshot of NetDraw, the network diagram tool used by participants.

detectives, working in "Zone 5" of a fictional police department. Pairs worked together remotely to identify a possible serial killer in Zone 5, and completed a report on their findings.

The participants were trained to perform detective work, to use the visualization tool, given their task assignment *as* detectives, and left to work on the task assignment for one hour. After an hour, or when the participants had completed their report, they completed an online survey to elicit their memory of the evidence about the serial killer. The experimenter then debriefed the participants.

Training

The participants practiced first on a comparatively simple problem involving the theft of a laptop computer from a college locker room. They read documents containing evidence relevant to four suspects in the theft and were asked to organize the data using a template. The template organized evidence into the motive, opportunity, and alibi of each suspect. Then they practiced on a more complex problem involving a rash of electronic equipment thefts. The case was constructed to give participants experience scanning and organizing information across crimes. Participants were also shown how to use a timeline and geographic worksheet.

Additionally, participants were trained how to use NetDraw (see Figure 2.1), the visualization tool adapted for this study, but only if they were assigned to one of the three visualization

43

conditions. Controls were trained on a spreadsheet that contained the same data. In the second practice case, a sample network diagram depicted the connections among the crimes. Participants were familiarized with the concepts of nodes and relationships, and they practiced searching and manipulating the diagram by location, time, and type of theft to give different perspectives on the crimes. Participants were encouraged to ask questions throughout the training. The average training session took 30 minutes.

2.2.3 COMPLEX PROBLEM SOLVING TASK

The pairs' task was to identify a possible serial killer in Zone 5. Each participant was instructed to report any other important information that might help their department solve the murders.

Documents and reporting forms

Evidence related to the serial killer was scattered in 15 assorted documents summarizing 6 "cold cases" and one open homicide, which also functioned as a simple problem-solving control task. The documents included witness and suspect interviews as well as coroner's reports. There were additional documents available on crime statistics by police district zone (e.g., Zone 1, Zone 2, etc.), a map of Zone 5 and adjacent zones, a map of Zone bus routes, and a police department organizational chart. Participants also could use an MO (modus operandi) worksheet for recording dates, weapons, and other relevant evidence for each case, as well as a suspect worksheet for recording different suspects, the suspect's connection to the victim, given alibis, and a timeline worksheet for recording when and where each crime took place, which was intended to support inter-case linkages. Finally, participants were asked to complete two online reports on the results of their investigation, one on their analysis of the serial killer, and another to report any other criminal activity they wanted to convey to the Zone 5 department.

All of the evidentiary documents and reports were available online and could be opened, searched, put in new or different folders, and manipulated freely. To ensure that sufficient screen space was available, participants had access to two 17" display monitors placed side by side. Also, participants were given paper versions of the instructions and worksheets.

Dispersed Evidence

The serial killer was responsible for four of the six homicides in the cold cases folder. Eight pieces of evidence—six pieces of evidence within the cold case files and two pieces of evidence

in the open homicide case file—could be linked to the serial killer: (1) similar blunt force trauma injuries to the victims; (2) victims killed in the evening after they returned from work; (3) victims rode the same bus route; (4) victims lived near the same bus route; (5) offender had been identified as a bus rider along with one of the victims (6) offender worked at a local hospital on the bus route; (7) offender had been identified on the bus (alibi for a witness in the open homicide case file); and (7) offender had been seen carrying a tool box on the bus. Identifying the serial killer required conceptually linking these disparate pieces of evidence from different cases rather than simply eliminating a defined group of suspects in one current case folder.

The caseload and evidence for the serial killer were distributed evenly between each member of the pair. To accomplish this, the six unique cold cases and the current open homicide case were divided between the pair such that each member received 3 distinct cold cases and half of the documentation for the current homicide case.

The open homicide case concerned the murder of a woman named Darlene Raffield. To solve this murder, participants had only to examine the documents in one folder, review the alibis of witnesses, and evaluate their motives and opportunities to commit the crime. If a pair spent too much time on this case, they would have less time to focus on the more complex task of finding the serial killer. In pretesting, we did find that individuals who spent more time on the Raffield homicide were less likely to identify the serial killer.

Communication

Participants were given an MSN Instant Messenger (IM) client and encouraged to use the IM to talk with their partner. All IM conversations between partners were recorded.

2.2.4 VISUALIZATION INDEPENDENT VARIABLE

Each pair was randomly assigned to one of four conditions, where each of the four conditions varied with respect to their access to a visualization tool. The tool enabled participants to see social and information network relationships in the data that linked names, places, events and objects, thereby providing a visual analysis perspective to help identify the serial killer.

Visual analysis was provided by NetDraw v.2, a tool used to create a social network diagram of all persons mentioned in the police documents. NetDraw v.2 is a software application for drawing 2D social network diagrams available online from Analytic Technologies. Social

network diagrams are aptly suited for complex problem solving. The evidence documents contained over 50 unique names and a diagram that represented how each person was connected to various other persons helped participants categorize and group people, and to view how people might be connected across cases. At the start of a session, each participant (except for those in the control condition) received the software, which showed a predetermined social network diagram that reflected the relationships between the documents they already had.

Figure 2.1 is a screen shot of the application. Within the diagram, each circle (or "node") represented a person from one of the crimes, and each line represented a relationship between two people. Victims were represented in red and other persons (such as witnesses and suspects) in blue. If printed in black and white, victims are black and others are grey. Thick lines denoted a strong tie (e.g., married people or coworkers). Thin lines denoted a weak tie (e.g., two people in the same place at the same time; for example, a waiter who served a customer in a restaurant, or two people who rode the same bus route).

Participants could freely manipulate and move the nodes on the screen, but they could not change the underlying relationships. Participants also could search or filter the diagrams based on a set of attributes to reveal people with common characteristics. Searchable attributes included: police district zone affiliation; case affiliation; occupation; mode of transportation; time of crime; location of crime; the weapon used to injure or kill the victim; and the injured body part of the victim. For example, within the attribute weapon, the three options were handgun, blunt instrument, and poison. If participants selected "handgun," all victims injured by a handgun would be visible onscreen.

The four experimental conditions—no visualization, unshared visualization, shared view-only visualization, and shared full access visualization—varied the degree of access that each participant had to NetDraw.

No Visualization

No visualization functioned as our control condition, wherein pairs did not have access to NetDraw. To ensure that they received the same information as the other three experimental conditions, participants were given Microsoft Excel spreadsheets that contained the same information about relationships between characters (See Figure 2.2). The names of these people

Spreadsheet   Case file

IM
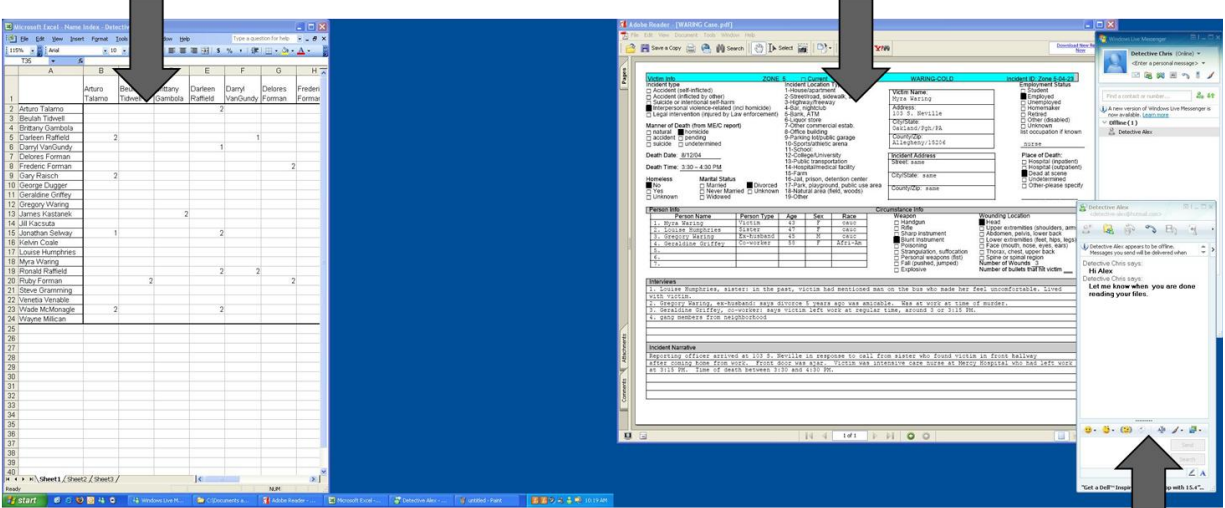
Figure 2.2. Screenshot of what participants in the No Visualizations condition saw on their dual screens.

Visualization   Case file

IM
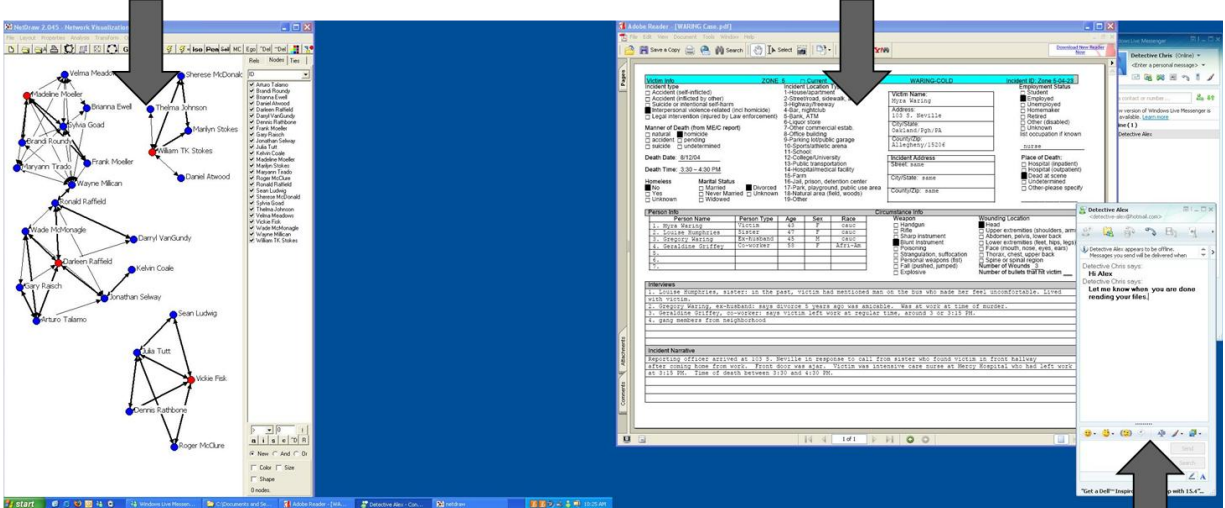
Figure 2.3. Screenshot of what participants in the Unshared Visualization condition saw on their dual screens.

47
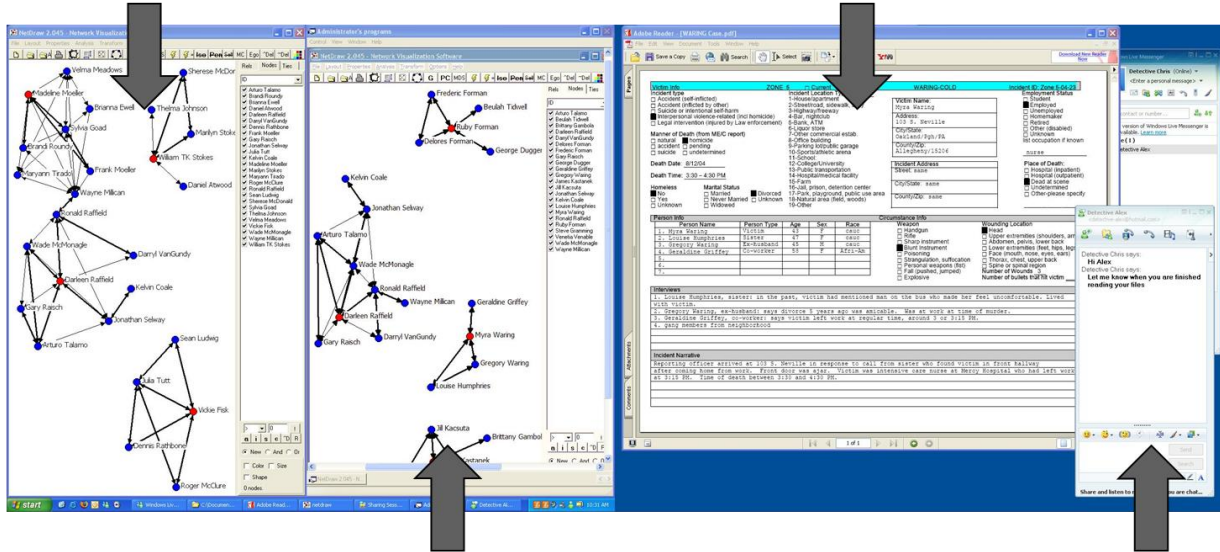
Figure 2.4. Screenshot of what participants in the Shared View-Only Visualizations condition saw on their dual screens.
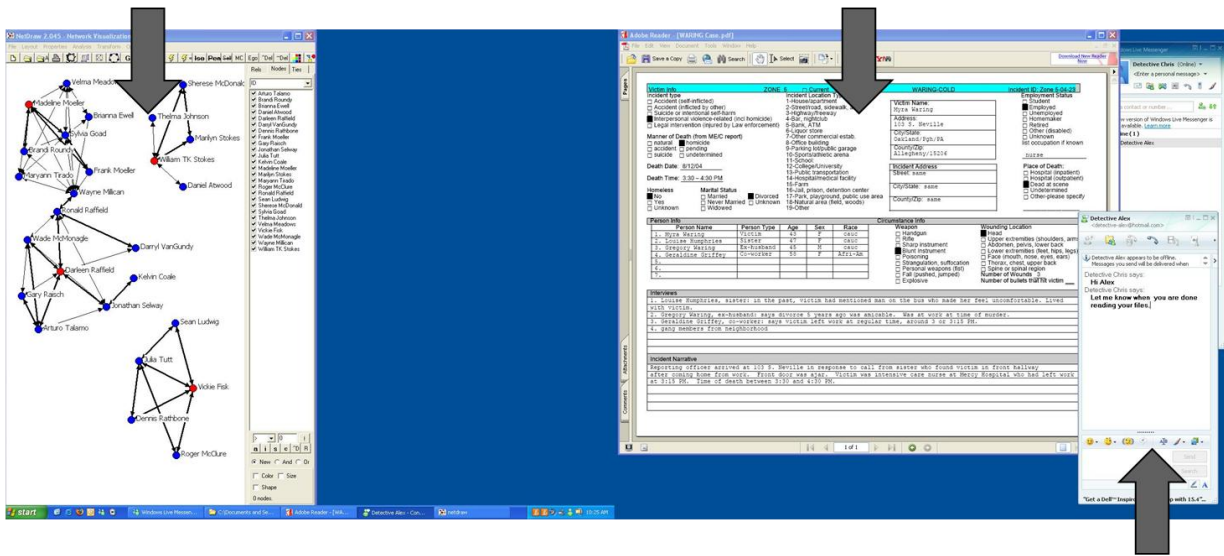
Figure 2.5. Screenshot of what participants in the Shared Full Access Visualization condition saw on their dual screens.

were arranged to form a matrix. Relationships in the matrix were represented by 0, 1, or 2 scores, which reflected no relationship, a weak relationship (such as a witness), or a strong relationship (such as a family member), respectively. Each participant received a spreadsheet that contained the relationship data only for their own cases. The experimenter explained how the spreadsheet could be used as well as the meaning of the numerical data.

Unshared Visualizations

Each member of the pair had access to NetDraw and a manipulable and searchable social network diagram of the data for the cases that they were given (See Figure 2.3). They could not view their partner's visualization.

Shared View-Only Visualizations

As in unshared visualization, each member of the pair had access to NetDraw and a manipulable and searchable social network diagram of the data for the seven cases. Each participant also had a window on their computer monitor that displayed their partner's social network diagram (See Figure 2.4). The participant could not search or manipulate this diagram, but they could view how their partner acted upon the diagram. The diagrams were shared via the Share Applications feature in MSN's Messenger client.

Shared Full Access Visualization

Each member of the pair had access to NetDraw and a manipulable and searchable social network diagram of data, but unlike the first three conditions participants also shared access to a network diagram that contained data from *all* the cases (See Figure 2.5). This diagram could be manipulated and searched by both participants, and was shared via a third computer using TightVNC, an open-source remote desktop software application.

2.2.5 MEASURES

We had four main sources of data: the final reports participants completed when their hour was up (or earlier if they had completed their analysis); an online posttest survey; IM logs; and WinWhatWhere files that recorded participant's use of the visualization tool.

	1	2	3	4	5	6	7
Collaborative Performance							
1. Pair identified serial killer (0 - 1)							
2. Time spent problem solving (minutes)	- .77 **						
Visualization Tool Use							
3. Visualization selected (min.)	.11	- .10					
4. Visualization tool active (min.) [a]	- .07	.14	.94 **				
Communication							
5. Total IM (# IM lines)	.21	- .26 [t]	.00	- .34			
6. Discuss serial killer (# IM lines)	.27 [t]	- .39 *	.14	- .30	.83 **		
7. Discuss Raffield homicide (# IM lines)	.08	- .01	- .01	- .04	.67 **	.21 [t]	
8. Discuss visualization (# IM lines)	.31 *	- .26	.41 **	- .10	.49 **	.62 **	.27 [t]

[t] $p < .10$, *$p < .05$, ** $p < .01$

[a]Visualization conditions only

Table 2.1. Correlation of measures of pair performance, use of the visualization tool, and communication (N = 47).

Individual and Collaborative Performance

Participants' were responsible for generating a written report that identified the serial killer. The reports were coded for whether or not they named the serial killer, and for whether or not they named the killer in the discrete Raffield case. Mentioning the name of the correct offender as either guilty or as a primary suspect who should be arrested counted as successful identification. We scored individuals members, but we were mainly interested in the success of collaboration. Hence both members of the pair had to have named the killer for the pair to have a successful collaborative performance.

Visualization Tool Use

Online activities were recorded via WinWhatWhere. WinWhatWhere is a software tool that records the application a participant is using, the time a participant spent with that window as the selected window, all keystrokes, and screenshots of the selected window. Due to resource constraints, only one randomly selected participant within each pair was recorded with WinWhatWhere. To estimate tool use, we calculated the total amount of time these participants had the visualization tool *as* the selected window. Active use was highly correlated with visualization window selection (see Table 2.1). In analyses, the total minutes the tool was selected and also active were log transformed to adjust for skewness.

Topic	Definition	Example
Serial killer task	Pertains to solving the serial killer task or evidence pointing to the serial killer.	"I see a connection between 2 of my cold cases; they both involve a blunt object."
Raffield homicide	Discussion pertaining to solving the Raffield homicide.	"I think the person who poisoned Darlene is Wade McMonagle."
Visualization	References the visualization tool or the visualization.	"My diagram says that Wayne Millican is somehow involved in the Darleen Raffield case." "Move those two out of the way."

Table 2.2. Conversational coding scheme.

Communication

We calculated how much conversation occurred between members of a pair by counting the total number of IM lines they exchanged during their session. One IM line refers to each new line within the recorded IM logs.

Participants' attention to different topics was also coded in IM conversations. IM logs were coded by line for whether or not the participants discussed the serial killer task, whether or not they were discussed the Raffield homicide, and whether or not they were referred to the social network diagram (See Table 2.2 for our coding scheme). An IM line was coded as a serial killer task if the line clearly showed that participants talked about or worked on searching for patterns of the serial killer; for example, "Here we have another blunt instrument incident," or "How do we connect these cases?" Discussion of the Raffield homicide was coded if the IM line referenced any person related to the Raffield homicide, or if the line clearly showed that participants thought about facts relating to the case; for example, "what did Darlene Raffield's boss say?" Because a single IM conversation line could be affiliated with both the Raffield homicide and the serial killer task, these counts were not mutually exclusive. For example, some participants discussed whether the Raffield homicide was connected to the serial killer task.

References to the social network diagram were coded if the IM line directly referenced the diagram; for example, if participants used words such as "diagram," "visualization," and "picture," or if participants discussed their active search within the diagram, such as "Watch this" and "See how these pop out?" The percentage of total IM lines during which IM lines referenced the visualization was calculated and log transformed. Over 5,000 lines of IM were coded using the scheme. An independent coder coded 7% of the data (Kappa = .76).

2.3 RESULTS

We obtained data from 47 pairs (94 participants), 13 pairs in the No Visualization condition, 10 pairs in the Unshared Visualizations condition, 12 pairs in the Shared View-Only Visualizations condition, and 12 pairs in the Shared Full-Access Visualization condition.

2.3.1 PRELIMINARY ANALYSES

To insure the task was equally difficult across conditions we administered the NASA TLX workload scale (Hart & Staveland, 1988) and CRT Scales (Frederick, 2005) on the posttest. Mean scores did not differ by condition. To insure that correctly identifying the serial killer reflected comparable insight across conditions, the posttest survey tested participant's recognition memory via multiple-choice questions for the eight pieces of evidence that might have led them to correctly identify the serial killer. Again, there were no differences across conditions.

Table 2.1 shows the correlations of measures on the pairs. These correlations of measure allow us to examine, across all conditions, whether visualization-related communication is associated with collaborative success. The table shows that, overall, when pairs identified the serial killer, they had also communicated more about the serial killer and talked more about the visualization. Active use of the visualization tool was *not* directly associated with communication, which could be due to partners' opening their visualization window once, then moving on to IM talk and document viewing.

2.3.2 INDIVIDUAL AND COLLABORATIVE PERFORMANCE

We first examined performance on the more simple of the two problems—the Raffield homicide. Although we did not ask pairs to solve the Raffield case, about one-third of the pairs did so anyway. We believe some pairs did so because it was an easy way to "get something done" when the pair had trouble identifying the serial killer. This line of thinking is supported by the fact that the correlation between identifying the serial killer and solving the Raffield homicide was $r = -.20$. There were also no differences across the four conditions for solving the Raffield homicide, which suggests that a visualization tool does not influence performance on a simple problem.

Next, we examined the more complex problem of the serial killer. If visualization improves individual performance, then that improvement could translate into a greater likelihood of collaborative success. Because the dependent variable, here the serial killer, is itself a discrete

52

variable, the appropriate analysis is a logistic regression (Hosmer & Lemeshow, 1989). A logistic regression assesses whether visualization conditions predict a dichotomous outcome (i.e., identified the serial killer or not). For individual participants, the logistic regression analyses showed a highly significant influence of condition on whether or not individuals identified the serial killer (logistic regression Likelihood Ratio $\chi2 = 12.1$, p $< .01$, $df = 3, 93$) with the No Visualization condition showing the greatest difference (Likelihood Ratio $\chi2 = 5.75$, $p = .01$).

We predicted in Hypothesis one that using a visualization tool would increase collaborative performance over performance in the control condition. Collaborative success for us was when both members of the pair correctly identified the serial killer. We conducted analyses at the pair level; those results are shown in Figure 2.6.

Pairs in the No Visualization condition performed worse than the other conditions, as predicted. Only 7.7% ($SE = 12.7$) of pairs in the No Visualization condition identified the serial killer— whereas 50% ($SE = 14.5$) of pairs in the Unshared Visualization condition, and 33.3% ($SE = 13.2$) of pairs in the Shared View Only Visualization condition, and 58% ($SE = 13.2$) of pairs in the Shared Full Access Visualization condition—identified the serial killer (logistic regression Likelihood Ratio $\chi2 = 9$, $p < .05$, $df = 3, 46$). Student's t tests revealed significant differences between the two best conditions—Full Access Visualization and Unshared Visualizations when compared to the No Visualization controls.
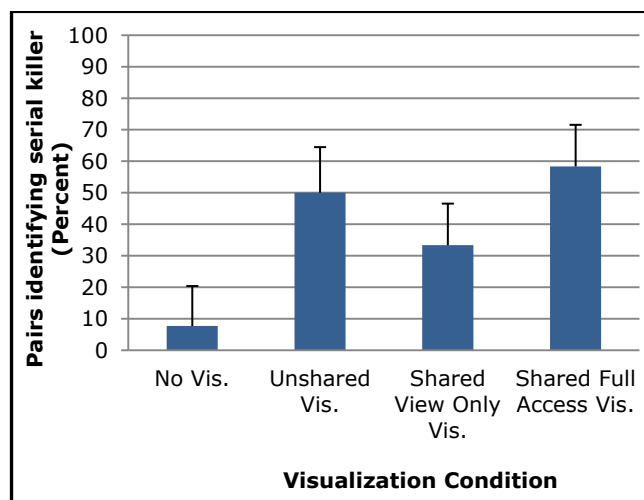
Figure 2.6 Percent of pairs solving the serial killer task by condition.

To test whether visualization helped the collaboration more than it did for individual members of each pair, we conducted a *nominal pair* comparison with the actual pairs. We randomly assigned the participants within each condition to someone else in the same condition—someone with whom they had never worked. Then we inspected the impact of visualization conditions on these nominal ("in name only") collaborators. The idea was to compare these nominal pairs with the actual pairs and then evaluate the extent that collaboration mattered when visualization was given to pairs. The results of this analysis are in Figure 2.7. They show that, controlling for condition, performance was worse by nominal pairs than by actual pairs (logistic regression Likelihood Ratio $\chi 2 = 3.04$, $p = .08$). Indeed, no nominal pair identified the serial killer in the No Visualization condition, and the top mean performance (in the Shared Full Access Visualization condition) was only 48% ($SE = 14.4$). These analyses suggest that, although visualizations aided individuals, collaborative performance was benefited from using the visualization tool.

The results support Hypothesis one. They show that visualization increases collaborative performance. However, the comparatively weaker performance of the pairs in the Shared View-Only Visualization condition suggests that *features* of the tool do matter. In what follows, we discuss tool use and communication in the three visualization conditions, as well as tests of hypotheses two and three.
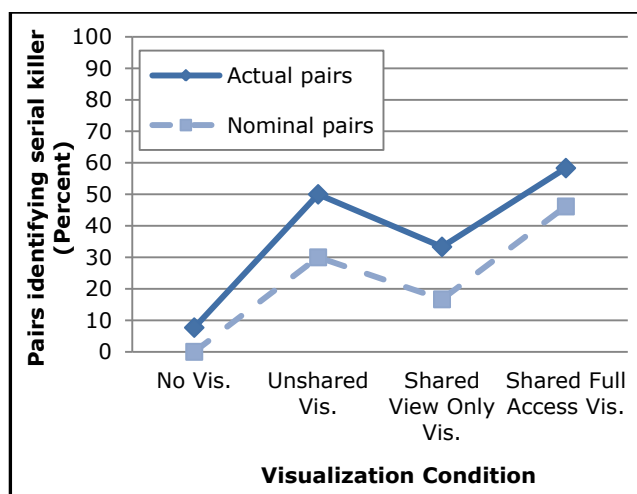
Figure 2.7 Percent of actual and nominal pairs solving the serial killer task, by condition.
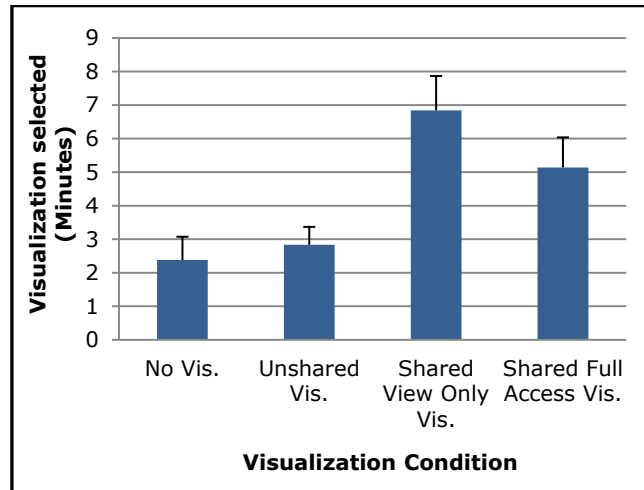
Figure 2.8 Mean number of minutes during which participants had the visualization selected, by condition.

2.3.3 VISUALIZATION TOOL USE AND COMMUNICATION

Hypothesis two predicted that access to a visualization tool would increase remote pair performance in complex problem solving when access to a visualization tool also increased information sharing and discussion by the pair. The first step was to examine whether or not access to the visualization tool changed pairs' behavior. And we did find that it changed pairs' behavior. In the No Visualization condition, the average participant spent 2.7 minutes with the spreadsheet selected. By contrast, in the Visualization conditions, the average participant spent 5.7 minutes with the network diagram opened (F [3, 43] = 4.1, p = .01). Pairs in the two Shared Visualization conditions used the visualization tool more than did pairs in the Unshared Visualization condition (F [2, 44] = 3.36, $p < 0.05$). As shown in Figure 2.8, Shared View-Only Visualization pairs used the visualization tool the most (M = 6.84 minutes, SE = 1.02), followed by Shared Full Access Visualization pairs (M = 5.14 minutes, SE = 0.89), followed by Unshared Visualization pairs (M = 2.83 minutes, SE = 0.54). A contrast revealed that this difference was significant when comparing both shared conditions against the unshared condition (F [1, 30] = 6.37, $p < 0.05$). These results indicate that sharing visualizations does encourage tool use.

Hypothesis three predicted that the Shared Full Access Visualization would best promote discussion and joint problem solving. Hence we tested whether the participants in the Visualization conditions, particularly in the Shared Full Access condition, communicated differently than those in the other conditions. We found no overall effect on the total amount of IM conversation in the pairs, but we did find a significant effect on talk about the network
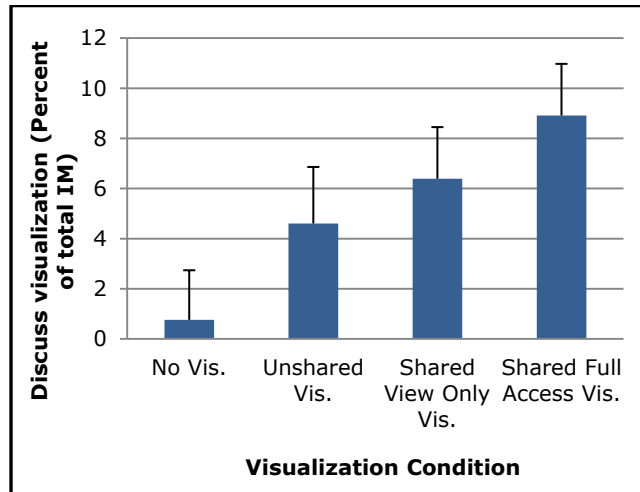
Figure 2.9 Mean percent of total IM lines during which pairs discussed the visualization, by condition.

diagram versus talk about the spreadsheet (F [3, 43] = 2.8, p < .05; see Figure 2.9). According to a Student's t test, pairs in the Shared Full Access Visualization condition talked significantly more about the network diagram (9% of IM lines) than did pairs in the other Visualization conditions (5% of IM lines), or pairs (talking about the spreadsheet) in the No Visualization condition (<1% of IM lines).

How was talking about the visualization relevant to identifying the serial killer? We looked at whether those who identified the serial killer talked differently with their partners in the three Visualization conditions. The correlational analyses showed interesting relationships across the three Visualization conditions. For example, the more that pairs talked about the network diagram, the more pairs discussed the serial killer (F [1, 30] = 11.5, p < .001), and the more pairs discussed the serial killer, the more likely they were to identify the serial killer (logistic regression Likelihood Ratio χ^2 = 3.6, p = .05). These analyses indicate that the visualization contributes to solving the complex serial killer case.

2.4 DISCUSSION

A simple visualization tool increased complex problem solving performance in remote pairs. In many cases, pairs identified the serial killer. But identifying the serial killer was also a very complex problem. Even after an hour of perusing documents and discussing cases, only 36% of pairs solved the case and caught the killer. The visualization tool did make a significant difference, improving not only individual performance, but also collaboration between pairs.

Access to the visualization tool encouraged discussion of a network diagram of case evidence more so than a simple spreadsheet that contained the same information. Discussion of the network diagram led pairs to discuss relevant evidence and hence contributed to their successful identification of the serial killer.

The features of the tool mattered, although differences among the tools tended to be overshadowed by the impact of having any visualization tool at all. Total manipulable access to the shared visualization (our Shared Full Access Visualization condition) encouraged pairs to use the tool and fostered more discussion and better performance—an average of 58% of the pairs solved the case. By contrast, when pairs had a tool that gave shared views but an inability to manipulate others' data (Shared View-Only Visualization condition), there was a dip in performance—an average of 33.3% of pairs solved the case. Although speculative, it is possible that when each member of the pair had his or her own visualization, and could only stare at the other person's diagram and manipulations, the two nonintegrated diagrams of data might have violated the "proximity compatibility" principle of display design (Wickens & Carswell, 1995), and confused pair members.

This study increases our understanding of how visualizations can aid collaboration. Our nominal pair analysis (see Figure 2.7) showed that real collaboration was valuable, but we do not know exactly how pairs came to aid one another; for example, we do not know whether pairs formed a common mental model of the problem (Fiore, Salas, Cuevas, & Bowers, 2003; Mohammed & Dumville, 2001), or whether they simply tried harder because the visualization was fun and motivating (Viégas et al., 2004). Additionally, since our top-performing Shared Full Access visualization also gave pairs a window into the entire integrated dataset, it remains unknown whether giving partners equal access to all available data is the key to collaborative success, or if that success hinged on the visualization pointing out important patterns or nodes in the data. Further research ought to study these more exact consequences of collaborative visualization tools in analysis. How are joint representations created, perceived and given meaning? These are queries worth being explored and understood, particularly when these joint representations are regarded as different from linguistic and gestural cognitive processing (Cheng, Lowe, & Scaife, 2001; Clancey, 1994; Zhang & Norman, 1994).

2.4.1 LIMITATIONS

This study cannot be generalized or applied to other genres of visualization tools, or to other task types (such as decision making), or to other remote, collaborative settings. For example, sharing information through IM may have also introduced barriers to the effective flow of information, or it may have made visualizations particularly effective in a way they would not otherwise be effective. Previous studies have shown that IM provides an effective channel of communication between partners (Scupelli et al., 2005) but an audio chat feature could help us understand the role different channels play in the use of visualization tools.

Also, participants were given predrawn social network diagrams. One could argue that if pairs took a more active role in creating the diagram, then they might also better understand their data (Suthers & Hundhausen, 2001). However, a recent trend in social network diagrams for analysis uses diagrams that are automatically generated from an existing dataset. Oftentimes the datasets have millions of different records. So a real challenge might be how to engage users in helping to create a dataset on this scale.

This study examined synchronous interactions. In distributed teams, colleagues often do not work simultaneously. Asynchronous, collaborative visualizations can encourage to knowledge discovery (Heer et al., 2007). Asynchronous communication and access to information visualization tools would be most similar to our Unshared Visualization condition. Pairs did quite well in this condition (50% solution rate). Thus our findings suggest that asynchronous teams would benefit from the use of such tools to solve complex problems.

2.4.2 SUMMARY

Information visualization, in the form of a network diagram, aided both individual and collaborative analysis. Real collaboration improved the performance of pairs over statistical pairings, particularly if pairs (a) had an integrated visualization that both could manipulate, and (b) when pairs discussed the visualizations they received. Doing so led to more relevant discussion of evidence and higher solution rates among pairs.

3 CAN EQUAL ACCESS TO INFORMATION INCREASE COLLABORATIVE SUCCESS?[2]

In a world of widespread access to information, large amounts of information can overwhelm collaborators, even when they have visualizations to help them. My first study found that visualizations can improve collaborative problem solving; however, it was unclear whether performance improved because the visualization pointed out important patterns or because the visualization gave partners visual access to all of the data. In a second study, I examined whether the visualization would still help a pair of collaborators if both collaborators had full access to all the evidence. I analyzed the success and discussion process of remote pairs of collaborators trying to identify a serial killer in multiple crime cases. In some instances, each collaborator had half of the evidence; in others, both collaborators had all the available evidence. These pairs of collaborators also used one of three tools: spreadsheet only (control condition), unshared visualization, or shared visualization. I found that visualizations improved analysis over the control condition, but the extent of this improvement depended on how much evidence each partner had. When each collaborator possessed all the evidence with visualizations, their discussion flagged and they showed evidence of more confirmation bias. They discussed fewer hypotheses and persisted on the wrong hypothesis.

3.1 INTRODUCTION

In the Vancouver serial killer example, the pattern of missing women emerged once all the caseloads from many detectives and several decades had been combined. The British Columbia Missing Women task force enabled this increased access to information. But what if the police department had a shared data repository in which detectives could easily search all available cases? Might a detective see a pattern sooner?

Within the intelligence community, where access to information has traditionally been severely limited, there have been recent efforts to reduce barriers to information access, in order to avoid

---

[2] The material presented in this chapter has been previously published as Balakrishnan, A.D., Fussell, S., Kiesler, S., & Kittur, A. (2010). Pitfalls of Information Access with Visualizations in Remote Collaborative Analysis. *Proc. CSCW 2010.* NY: ACM.

catastrophic lapses in analysis and delays in information sharing (Office of the Director of National Intelligence, 2008; Valledor, 2010). For example, Mike McConnell, the Director of National Intelligence, has emphasized a new culture of information sharing across the many agencies responsible for national security:

> The information sharing strategy is focused on developing a 'responsibility to provide' culture in which we unlock intelligence data from a fragmented information technology infrastructure spanning multiple intelligence agencies and make it readily discoverable and accessible. (Office of the Director of National Intelligence, 2008)

The hope is that with increased access to information, it will be easier to discover patterns and make sense of the data. New advances in cloud computing technology have made such sharing improvements more feasible (Adams, 2011). While improved information sharing is meant to streamline the work process for intelligence analysts, it also increases information overload for them.

Collaborative analysis combined with visualization tools might be an ideal solution to the information overload problem because it could provide both social and cognitive solutions. As I showed in Chapter 2 (Balakrishnan et al., 2008), visualizations such as Figure 2.1 have been shown to facilitate collaborative analysis (Edelson et al., 1996; Mark, et al., 2003a, 2003b). My first study found that visualizations can improve collaborative problem solving; however, it was unclear whether performance improved because the visualization pointed out important patterns or because the visualization gave partners visual access to all of the data. Having visual access to all the available information may have reduced the burden between collaborators to explicitly share facts with one another. In addition, visual access to all the evidence may have helped pairs to overcome coordination costs that arise from the time spent, and possibly wasted, in discussion (Shepperd, 1993).

Another advantage of visualizations is that they may help combat certain cognitive biases. Cognitive biases, particularly confirmation bias—the tendency to seek out information that confirms what one already thinks, and avoid information that disconfirms it—can cause analysts to persist on the wrong hypothesis (Nickerson, 1998). Having visual access to all the evidence

may have encouraged partners to seek out non-confirming evidence rather than focusing on information in common (Stasser & Titus, 1985).

This chapter examines remote pairs of analysts collaborating on the serial killer task described in Chapter 2. Success on this task depends on insight when combing through hundreds of pieces of evidence. I examine how the distribution of evidence (each partner has all the evidence or each has half of it) and the availability of visualization tools change how the pairs discuss the evidence and how successful they are in their problem solving.

3.1.1 DISTRIBUTION OF EVIDENCE

Collaborators may have different access to the myriad of raw data or evidence on a given problem for organizational, legal, political, and other reasons. Sometimes everyone has all the collected evidence; for instance, after the outbreak of swine flu, epidemiologists in all of Great Britain used a common tracking database of medical cases, called QSurveillance (QSurveillance, 2010). At other times, analysts have partial evidence. For instance, in the U.S., restrictions define which intelligence analysts can view which portions of intelligence data. One goal of this paper is to explore how the distribution of evidence influences collaborative analysis.

When each analyst has all of the data or evidence, the demand for timely exchange of raw facts is minimal, and discussion can focus on inferences and hypotheses drawn from the data. At the same time, having all the data raises the specter of information overload. To minimize such information overload, analysts may discuss limited hypotheses and attain a common mental model. Although many writers argue that groups need a shared mental model (e.g., Blockeel & Moyle, 2002; Kozlowski, Ilgen, & Klimoski, 2006), it can lead to confirmation bias. Thus, even in small groups with limited information to share, collaborators seldom attain knowledge gains and improved performance from full information (Mojzisch & Schulz-Hardt, 2005).

When analysts have only partial access to evidence, there is much more demand for information exchange; often the problem cannot be solved without it. For time-sensitive problems, valuable time will be spent simply making sure that everyone has the right information. To save time, analysts may decide to share lines of investigation or hypotheses, rather than raw data. For example, if a detective has noticed that many crimes take place near hospitals, he might share this observation with fellow detectives, rather than all his crime cases. If each analyst contributes

a unique perspective, the analysts may debate alternative hypotheses, thereby avoiding confirmation bias. Thus we propose that when analysts do *not* have all of the evidence themselves, they are likely to spend more time discussing hypotheses and relating them to the evidence than when they have all the evidence.

Hypothesis 1: Pairs of analysts will solve the problem more often, discuss the problem more, and generate more hypotheses and better-supported hypotheses, when each partner has partial evidence, than when each partner has all the evidence.

3.1.2 INFORMATION VISUALIZATION

Billman et al. (2005) report that distributed pairs using CACHE (Convertino et al., 2008), a system with visual data presentation for intelligence analysis, overcame *a priori* biases and did more effective data analysis. Mark et al. (2003a, 2003b) reported that remote pairs with visualizations communicated more than collocated pairs did. Their results and the results from my first study suggest that communication helps pairs take advantage of the visualization tool. From this work, I maintain:

Hypothesis 2: Pairs of analysts with a visualization tool will solve the problem more often, will discuss the problem more, and will generate more hypotheses about the data and better-supported hypotheses, than analysts without a visualization tool.

3.1.3 VISUALIZATIONS WITH ALL OR PARTIAL EVIDENCE

If visualization tools provide the benefits we have discussed above, the degree of benefit may depend on the way evidence is distributed across members of a collaborative team. Although visualizations may be expected to improve hypothesis generation, discussion, and problem solving regardless of how evidence is distributed among analysts, these benefits may be reduced when the analysts each have all the evidence, and therefore do not need to exchange information and discuss the problem as much.

Hypothesis 3: Visualizations will benefit collaborative analysis more when each partner has partial evidence than when each partner has all the evidence.

3.2 METHOD

I report the analyses of data from an experiment designed as a two-level factorial, with two information conditions (Half Evidence vs. All Evidence), and three visualization conditions

(None, Unshared Visualization, Shared Visualization). Participants worked in pairs randomly assigned to one of the three visualization conditions. I collected the data in the half-evidence conditions for my first study (Balakrishnan et al., 2008) and subsequently collected data for the all-evidence conditions to understand the significance of the distribution of information.

3.2.1 PARTICIPANTS

One hundred and eighty total participants participated in the experiment, described as a "Detective Mystery Study" (84 female, 96 male; 55% U.S. born; age range 18-64, median age approximately 22). Eighty-eight percent of the participants were undergraduate or graduate students. Participants were paid $15 for their participation. They were told the experiment would last 1.5 hours. There were no demographic differences between the participants across conditions.

3.2.2 PROCEDURE

Participants were seated separately, such that they could not see their partner or their partner's workstation. They role-played a pair of detectives of a police department, collaborating remotely to identify a possible serial killer (Scupelli et al., 2005). They had to work through many documents and reports to detect the serial killer. After working together on this task, they were each asked to complete two online reports on the results of their investigation.

Participants were trained to use either NetDraw (See Figure 2.1), the visualization tool adapted for this study, if they were in the visualization conditions, or the Excel spreadsheet, if they were in the control condition. Training took an average of 30 minutes.

After training, the pairs were left to work on the assignment for one hour. They were given an MSN Instant Messenger [IM] client and encouraged to use the client to talk with their partner. After an hour, or when the participants had completed their investigation and report, they each completed an online survey to elicit the evidence they used to identify the serial killer. (For a complete description of the task, see "Methods," Chapter 2.2.)

3.2.3 DISTRIBUTION OF EVIDENCE

The evidentiary documents and reports were available online and could be opened, searched, put in different or new folders, and manipulated freely. To insure that sufficient screen space was available to examine multiple documents at once, the participants each had access to two 17"

monitors placed side by side. Participants were also given paper versions of the instructions and worksheets.

In the case files, participants had witness and suspect interview reports, coroner's reports, crime statistics by police district zone, a map of the zone and adjacent zones, a bus route map, and a police department organizational chart. Participants could also use one worksheet for recording dates, weapons, and other relevant evidence for each case, another worksheet for recording different suspects, their connection to the victim, and alibis, and a third worksheet for recording when and where each crime took place, intended to support inter-case connections.

In the *Half Evidence* condition, each member of the pair had half of the caseload and evidence for the serial killer on their computer. In the *All Evidence* condition, each member of the pair had all of the cases and documents.

3.2.4 VISUALIZATION TOOL

Each pair was randomly assigned to one of three conditions, differing with respect to their use of a visualization tool. The visualization tool, NetDraw v.2, enabled participants to see social and information network relationships in the data because it linked names, places, events, and objects, thereby providing a visual analysis perspective to identify the serial killer. (For a more complete description of the visualization and the various conditions, see "Methods," Chapter 2.2.)

In the *No Visualization* condition, collaborative pairs did not have access to NetDraw. To ensure that they received the same information as others, they were given Microsoft Excel spreadsheets containing the same relationship information among the persons mentioned in the evidence documents.

In the *Unshared Visualization* condition, each partner in the pair had access to NetDraw and to an interactive and searchable social network diagram of their own evidence (either their own half or all of the evidence). They could not view their partner's visualization.

In the *Shared Visualization* condition, each member of the pair had access to NetDraw and to an interactive and searchable social network diagram of all the evidence. (In the previous study, this condition's full name was *Shared Full-Access Visualization*, to differentiate it from the

visualization condition that has been removed in this study.) This diagram could be manipulated and searched by both participants in the pair. Effectively, this condition meant that, in the Half Evidence condition, each partner could see a diagram of all the evidence even though they only had direct access to half of the supporting evidence on their own computer. In the All Evidence condition, each partner not only had all the evidence on their computer, but also saw a diagram of all the evidence.

Because pairs in the *Shared-View-Only Visualization* condition confused participants and resulted in the lowest rates of collaborative success in Study 1, this condition was removed for the second study.

3.2.5 MEASURES
As in Study 1, there were three main sources of data: participants' final reports, their posttest surveys, and IM logs of their discussions.

Identifying the Serial Killer

We determined whether participants correctly identified the serial killer from their written reports. We were interested mainly in the success of the *collaboration*, so both members of the collaborative pair had to have named the serial killer for the pair to be coded as having successful collaborative performance. However, the results were essentially the same at the individual level.

Discussion Process

We calculated how much the pair communicated by counting the total number of IM words they exchanged during a session. We also coded participants' discussion topics line by line. In total,

Topic	Definition	Example
Serial killer task	Pertains to solving the serial killer task or evidence pointing to the serial killer.	"I see a connection between 2 of my cold cases; they both involve a blunt object."
Clue Discussion	Discussion pertaining to one of the eight critical clues.	Detective A: "Hey, all of our victims ride the 500 bus." Detective B: "Ooh, good find!" or "That make[s] sense, they all lived near the 500 as well!"
Hypothesis Discussion	Discussion of a new hypothesis is introduced or confirmed.	"I think these four blunt instrument victims are connected." "I feel like it is a suspicious man on the bus."

Table 3.1 Conversational coding scheme.

there were more than 8,700 lines of IM (See Table 3.1 for the conversational coding scheme). An independent coder coded 7% of the data (Kappa = .71). All codes were at the individual level. Hypotheses were counted only the first time they were discussed, even if pairs revisited a certain hypothesis after considering other hypotheses in between. The reason for this coding decision was that prior research suggests that the consideration of unique hypotheses, not the total number of times a hypothesis is mentioned, contributes to problem-solving success.

Individual Characteristics

Prior research suggests that individuals' tendency toward cognitive reflection, as measured by a simple scale called the CRT, improves their ability to overcome confirmation bias (Frederick, 2005). We used CRT scale scores as a control variable in our analyses. We also administered the NASA TLX scale, a measure of task workload (Hart & Staveland, 1988).

3.3 RESULTS

I and my colleagues analyzed data from ninety pairs (180 participants), with fifteen pairs in each of the six conditions.

3.3.1 IDENTIFYING THE SERIAL KILLER

From the hypothesis that distributed evidence leads partners to discuss and debate problems more deeply, we predicted that pairs whose partners each had only half of the evidence would perform better than those pairs whose partners both had all of the evidence. We also predicted that visualizations would help pairs solve the problem. Because the dependent variable, identifying the serial killer, is a discrete variable, the appropriate analysis is a logistic regression (Hosmer & Lemeshow, 1989). This regression assesses whether the independent variables predict the dichotomous outcome, identifying the serial killer. We conducted analyses at the pair level. We found that performance depended on whether the pair had access to all of the evidence. Figure 3.1 shows the results of the analysis, which support Hypothesis 3, the interaction effect.

In the Half Evidence condition, only 13% (SE = 12.5) of pairs in the No Visualization condition identified the serial killer, while 46% (SE = 11.8) of pairs in the Unshared Visualization condition, and 60% (SE = 11.8) of pairs in the Shared Visualization condition identified the serial killer. Student's t tests show differences at the $p < .05$ level between No Visualization and the Shared Visualization condition. In the All Evidence conditions, however, all three conditions
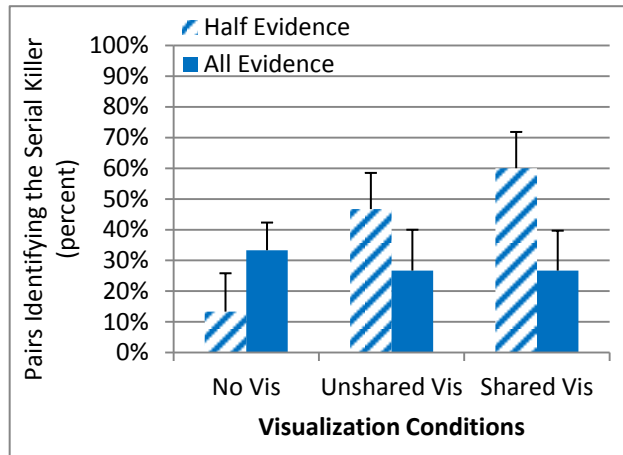
Figure 3.1. Percent of pairs solving the serial killer task by condition.

performed comparatively poorly: 33% (*SE* = 9.01) of pairs in the No Visualization condition, 27% (*SE* = 13.3) of pairs in the Unshared Visualization condition, and 27% (*SE* = 13.3) of pairs in the Shared Visualization condition identified the serial killer (logistic regression Likelihood Ratio χ2 = 9.3, *p* < .09, *df* = 5, 90; Cramer's Phi = 0.35). The two visualization conditions in the Half Evidence condition significantly outperformed both All Evidence visualization conditions (logistic regression Likelihood Ratio χ2 = 8.4, *p* < .05, *df* = 3, 120).

Because so many pairs failed to identify the serial killer, we rated each participant's reports based on his or her progress toward the solution on a four-point scale: 0 for *unsolved*, 1 for *suspected pattern*, 2 for *suspected perpetrator,* and 3 for *correct solution*. We conducted an ANOVA with the solution as the dependent variable, evidence condition and visualization condition were between groups factors, and CRT scores were a control. (Non-integer degrees of freedom may occur in these analyses, see [Littell, Milliken, Stroup, & Wolfinger, 1996]). We found a significant effect as to whether the pair solved the problem by evidence condition (*F* [1, 82.36] = 4.57, *p* < .05; Cohen's *d* = 0.46) and no effect by visualization condition. Individuals in the Half Evidence condition (*M* = 1.87, *SE* = .12) had significantly better solutions than those in the All Evidence condition (*M* = 1.34, *SE* = .13).

In summary, visualizations did increase problem solving success as predicted, but only when evidence was distributed. In the next section, I analyze participants' discussions in order to
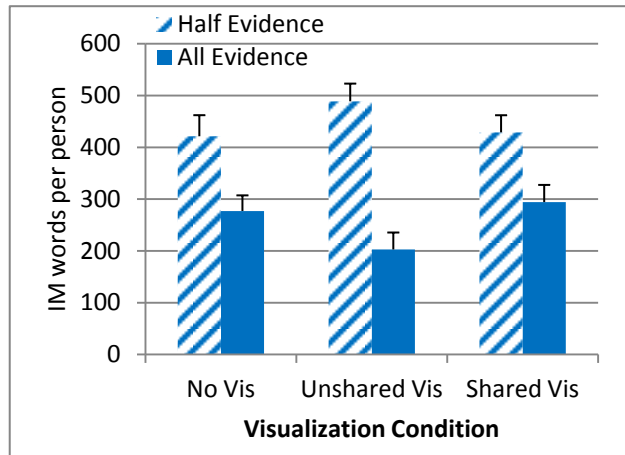
Figure 3.2 Average number of individual contributions of IM words by condition.

evaluate why shared and unshared evidence affected the success of their collaborative analyses so greatly.

3.3.2 DISCUSSION PROCESS

Total Talk

We predicted that pairs with half the evidence would discuss the problem more than pairs with all the evidence. We counted the total number of words each participant contributed to their IM discussion. We log transformed the data because they were skewed. In an ANOVA, the number of total IM words was the dependent variable, evidence condition and visualization condition were between groups factors, and CRT scores were a control. These results are seen in Figure 3.2. As predicted, individuals in the Half Evidence conditions ($M = 446$, $SE = 20.9$) exchanged significantly more words with their partners than individuals in the All Evidence conditions ($M = 256$, $SE = 15.2$; $F [1, 80.8] = 28.9$, $p < .01$).

We also predicted that the visualization tools would increase discussion among pairs. Overall, visualization condition did not affect the amount of discussion ($F [2, 80.8] = .06$, $p = .94$). However, the interaction effect between visualization and information conditions showed a trend in the predicted direction ($F [2, 80.9] = 1.90$, $p = .16$).

Overall, a greater number of IM words was significantly correlated with better solution rates ($r = .21$, $p < .01$). However, the importance of discussion varied by condition. In the two conditions where solutions were most likely, Half Evidence/Unshared Visualization and Half

68

Evidence/Shared Visualization, words were positively correlated with higher solution rates ($r = .34$, $p = .06$; $r = .54$, $p < .01$, respectively), whereas in the other conditions the correlations were lower.

Discussion of Serial Killer

Total IM words and total IM words that discussed the serial killer were highly correlated ($r = .76$, $p < .01$). For each participant, we summed the total number of words about the serial killer case and divided by the participant's total IM talk to control for individual variations in talk amount. In an ANOVA, the amount of serial killer discussion was the dependent variable, evidence condition and visualization condition were between groups factors, and CRT scores were a control. As predicted in Hypothesis 1, individuals in the Half Evidence conditions ($M = 256$, $SE = 15.2$) exchanged significantly more words with their partners about the serial killer case than individuals in the All Evidence conditions ($M = 153$, $SE = 15.0$; $F\,[1, 38.03] = 4.30$, $p < 0.05$). However, contrary to Hypothesis 2, those in the Unshared Visualization ($M = 198$, $SE = 18.4$) and Shared Visualization ($M = 203$, $SE = 21.5$) conditions did *not* talk more about the serial killer case than those without a visualization ($M = 218$, $SE = 19.1$; $F\,[2, 164] = .76$, *ns*).

Overall, discussion of the serial killer was significantly positively correlated with better solutions ($r = .47$, $p < .01$). This relationship was highest in the two conditions where there were the most solutions: the Half Evidence/Unshared Visualization and Half Evidence/Shared Visualization conditions ($r = .52$, $r = .57$, $p < .05$, respectively).

Discussion of Evidence

We predicted that having a visualization tool and half the evidence would increase sharing of pieces of evidence that were critical to solving the problem. We counted the number of critical pieces of evidence that partners shared with each other in their IM discussion, and compared that number to the number they recalled in the posttest survey. This analysis allowed us to compare the evidence discussed with the critical evidence recalled for each individual. In an ANOVA, the percent of evidence discussed was the dependent variable, evidence condition and visualization condition were between groups factors, and CRT scores were a control. We found a trend in the direction predicted in Hypothesis 1 (*ns*) only for evidence condition as a main effect. Participants

in the Half Evidence conditions discussed a higher percentage of the evidence with their partner ($M = 63.2\%$, $SE = 3.7\%$) than those in the All Evidence conditions ($M = 53.9\%$, $SE = 4\%$).

Hypothesis Generation

We predicted that generating a greater number of unique hypotheses would help pairs reach a solution. Consistent with this idea, the overall correlation between generating hypotheses and solutions was r = .34, p < .01. Furthermore, we hypothesized that having access to half the evidence would increase pairs' generation of unique hypotheses, and that access to a visualization tool would also increase hypothesis generation. In an ANOVA in which the total number of unique hypotheses contributed to IM discussion by each individual was the dependent variable, evidence condition and visualization condition were between groups factors, and controlling for CRT scores, we found a marginal main effect by evidence condition (F [1, 31.1] = 3.5, $p = .07$) but no visualization main effect (see Figure 3.3). We also found a significant interaction effect between information and visualization condition (F [1, 165] = 3.3, $p < .05$), suggesting that visualizations helped in the Half Evidence conditions but not in the All Evidence conditions. In the Half Evidence conditions, Student's t tests show that those in the Unshared Visualization ($M = 2.77$, $SE = .2$) and Shared Visualization ($M = 2.73$, $SE = .3$) conditions discussed a significantly greater number of unique hypotheses than participants in the No Visualization condition ($M = 2.33$, $SE = .22$). For some reason, those in the All Evidence-No Visualization condition generated the most hypotheses; however, there was no correlation
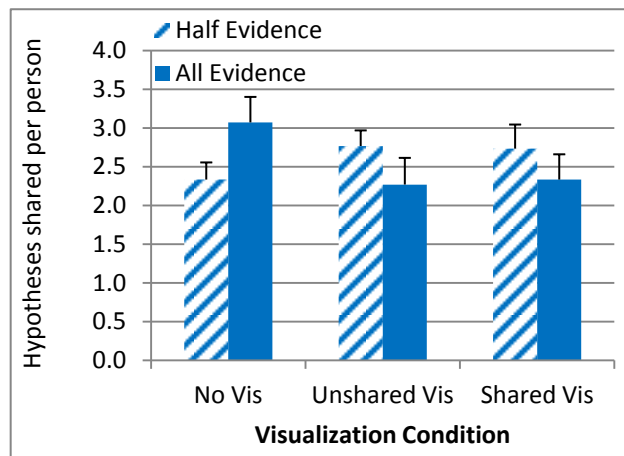
Figure 3.3 Average number of hypotheses shared per person across conditions.

between generating hypotheses and finding the correct solution in this condition, suggesting that their discussion was less fruitful.

3.4 DISCUSSION

This second study explored the impact of the distribution of information and a visualization tool on the process of collaborative problem solving. We found that using a visualization tool aids problem solving, but only when information is distributed between collaborative partners. This finding is contrary to the implicit assumption in much writing about data sharing that greater access to data will aid collaborations. In this study, when both partners had access to all of the evidence, they performed more poorly and showed more evidence of confirmation bias than the partners who each had access to only half of the evidence.

One possible explanation for these results is that participants in the All Evidence conditions suffered from information overload. They had twice the number of text documents and far more evidence to look at. However, in the Half Evidence conditions, participants required more evidence to solve the problem than they themselves possessed, so it would seem that they had just as great a workload as participants in the All Evidence conditions because they had to acquire additional information from their partner (with the added overhead of communicating, representing, and storing that information). Furthermore, as measured by the NASA TLX scale on the posttest survey, participants in the All Evidence conditions did *not* report feeling a higher workload than participants in the Half Evidence conditions.

Half Evidence conditions may also have implicitly provided a strategic structure to the collaborative process. For instance, each of the partners can first go through his or her evidence and share his/her perspective with the partner. By contrast, partners with all the evidence may arbitrarily sift through the evidence and fail to consider the partner's viewpoint. Prior research has shown that assuming what the other partner knows leads to lower rates of collaborative success (Nickerson, 1998). Significantly, participants in the All Evidence condition recalled as much evidence on the posttest as participants in the Half Evidence condition, even though they discussed a lower percentage of that evidence with their partner. This finding suggests that they lacked motivation to share information, perhaps because they assumed that their partner was aware of the same evidence.

Giving participants half the evidence may also have given them a sense of ownership and expertise about their own evidence. If people feel their contributions are important to collaborative success, they are less likely to show social loafing (Karau & Williams, 1993). With both partners actively sharing ideas, there is a greater diversity of ideas within the pair, which can be associated with better collaborative outcomes (Jehn, Northcraft, & Neale, 1999). Finally, a sense of differing expertise within groups helps reduce the tendency to focus on information that has already been shared, which mitigates confirmation bias (Stasser, Vaughan, & Stewart, 2000).

One might think those with partial evidence would generate a narrow perspective, based on their own data, that would anchor their point of view. However, these pairs knew that their partners had relevant evidence. This knowledge could be a critical reason why those in the Half Evidence/Shared Visualization condition outperformed participants in the other conditions. In that condition, partners had only half of the evidence, but they could see a diagram of all the evidence—including their partner's—on the screen. Having a visualization of all the evidence may have elicited conversation about the problem and the evidence. This conversation may have forced confrontation with disconfirming evidence for incorrect hypotheses, and thus combated confirmation bias.

In order to better explain the role of confirmation bias in this study, I performed a detailed tracing of the discussions in the Shared Visualization conditions, where the differences between the Half Evidence and All Evidence conditions were most stark. Figure 3.4 shows a diagram of all discussions in those conditions. Each dot represents one pair in either the All Evidence (whole circles) or Half Evidence (half circles) condition. Orange (or light grey in black and white) dots represent failures to solve the case, and blue dots (dark grey in black and white) represent identifying the serial killer. In the far right path, all pairs in the All Evidence conditions who started with the irrelevant Raffield case remained stuck there, whereas five out of the seven pairs in the Half Evidence conditions successfully moved on to solve the serial killer case. In the middle path, pairs in the Half Evidence condition who saw a serial killer pattern first also then understood a crucial connection between cases, whereas a majority of those in the All Evidence conditions who saw a pattern did not successfully identify a connection. In the far-left path, more All Evidence pairs noticed a connection between two cases than did those in the Half Evidence
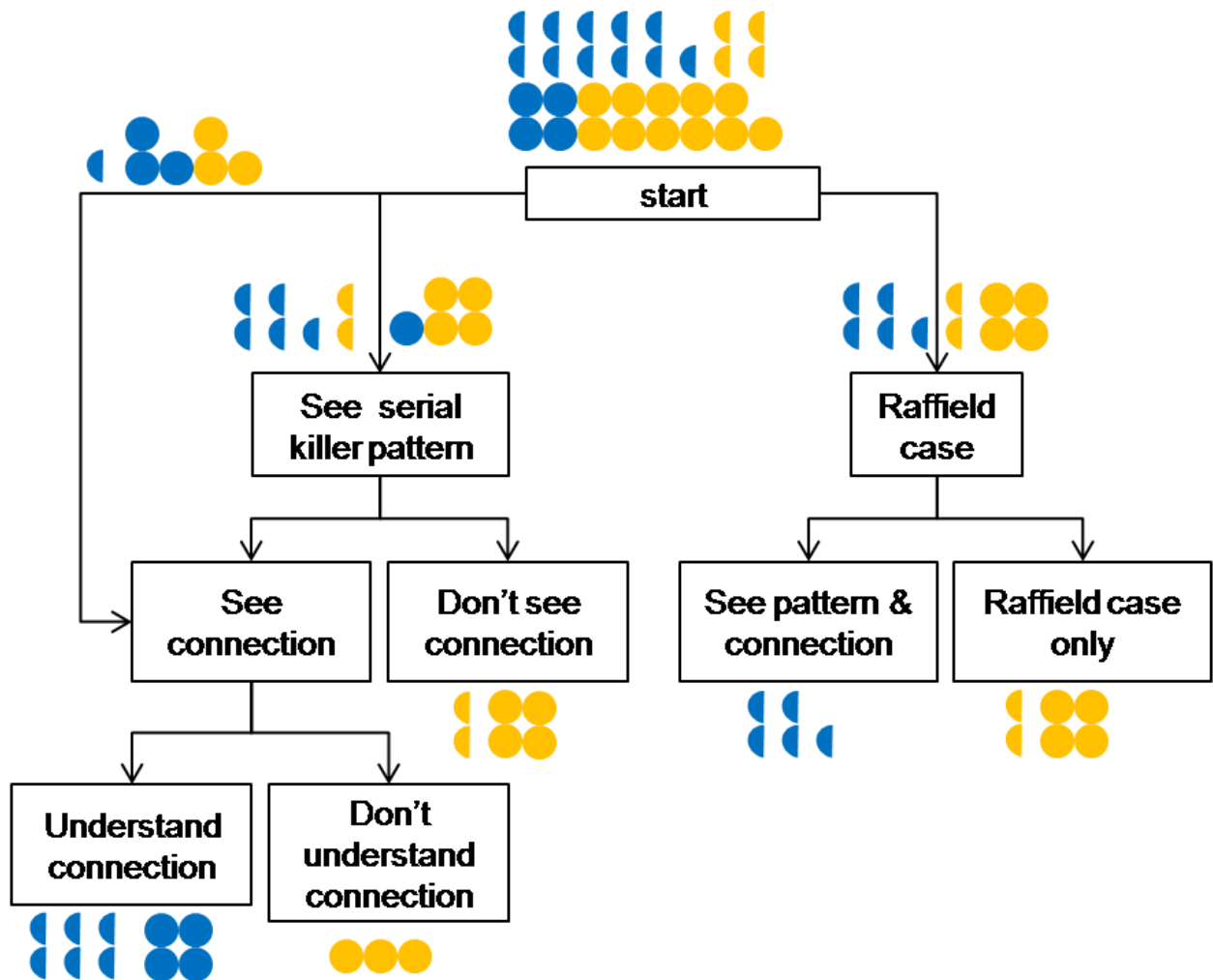
Figure 3.4 Process map of shared visualization pairs.

Full circles represent All Evidence condition pairs and half circles, Half Evidence condition pairs. The color blue designates pairs who solved the serial killer case while orange designates pairs who did not solve the case.

conditions, but seeing this connection translated to successfully understanding what the connection meant and identifying the serial killer only 50% of the time.

This analysis shows why we believe confirmation bias plagued those in the All Evidence conditions, and visualizations did not help them. If they started off on the wrong path, they were more likely to stay there, and even if they noticed an interesting clue, such as the connection between two cases or the serial killer pattern, they did not debate the data enough to come to a correct solution.

The detailed tracing also revealed a failure on the part of pairs in the All Evidence condition to partition the task. Only three pairs in the All Evidence condition proposed splitting up the evidence and examining parts of it in more depth—and in all three cases they quickly abandoned the idea. These pairs did not develop effective problem-solving strategies. Giving members of a pair only half the evidence automatically provides a structure for the problem-solving process: pairs know they must read their own information and then report back to their partner any interesting findings. Without any obvious task partitioning, pairs with all evidence appear to become lost in information overload.

3.4.1 LIMITATIONS

Although this study contributes to understanding how visualizations and the distribution of information can affect collaborative analysis, we have studied only one analytic task, limiting generalizability. Also, we studied people who had not worked together previously. Collaborators may build experience working with one another, improving their communication (however, this experience seems not to reduce confirmation bias [Heuer, 1999]). Also, our participants used IM, whereas real-world collaborations most likely rely on more than just one form of communication, including audio and video channels. Finally, in real-world environments, analysts are highly trained in the use of knowledge and visualization tools, whereas our participants may have suffered from inadequate experience, having been given only minimal training in its use.

3.4.2 SUMMARY

Visualizations improved remote collaborators' performance over the control condition, but this improvement depended on how much information each partner had. When each partner had all the evidence, discussion flagged, pairs discussed fewer hypotheses, and they persisted on the wrong hypotheses—in other words, they suffered from confirmation bias. These pairs seemed to be overwhelmed and did not systematically approach the task. My first two studies suggest that visualization tools might prompt collaborations to be more systematic, but this depends on collaborators using the visualization, finding relevant patterns, and ultimately using these findings to direct their analysis.

4 CAN PROCESS INTERVENTIONS STIMULATE VISUALIZATION USE AND IMPROVE COLLABORATIVE SUCCESS?

The first two studies of this dissertation reveal important findings, but several facts remain puzzling. Some pairs perform poorly even though they have been granted access to resources that should improve their opportunities for success, even with a partner, equal access to information, and visualization. This chapter details a third study that explores whether process interventions can lead pairs to collaborative success. Pairs with access to all the evidence and a shared visualization are given one of four sets of interventions: (1) no intervention (control); (2) coordination intervention; (3) sensemaking intervention; or (4) both coordination and sensemaking interventions. While pairs with any type of intervention outperform pairs without intervention, (3) performs best. Pairs given the sensemaking intervention uncover critical relationships within the visualization that lead to problem solving success and outperform other pairs.

4.1 INTRODUCTION

Studies one and two asked whether visualization tools could elevate the performance of remote pairs. The first study showed that visualization tools have the potential to lead collaborators to success, yet these rates of success were still surprisingly low. Further, the benefits of a visualization tool were negated when both partners had access to all the evidence.

But working with multiple partners should help the problem solving process in several ways. First, many people can cover more information than a single person. Second, many people may highlight more patterns and perspectives, which should aid the quality of the sensemaking (Pirolli & Card, 2005). Third, groups should be able to leverage the abilities of the most intelligent person in the group. In other words, even if one person finds the solution, the group should reap the benefit. And yet, as we have seen, pairs do not outperform individual problem solvers in pretests, and are oftentimes guilty of confirmation bias when given access to all the evidence. What is going on?

Prior research on collaborations *does* point to potential barriers that stand in the way of success. Indeed, having to collaborate may introduce more harm than good (Ringelmann, 1913; Moede, 1927; Levinger, Graves, & Peckham, 1974; Kravitz & Martin, 1986). Successful collaborations depend on the how successful collaborators coordinate efforts, share information, and leverage multiple perspectives (Devine, 1999; Mennecke & Valacich, 1998; Tudor, Trumble, & Diaz, 1996). Conversely, if these collaboration costs are left unmanaged they can overwhelm a team, which results in wasted time and opportunity (Yamane, 1996; Zawaki, 1994). And pairs with shared visualization access in the second study faced these kinds of collaboration costs (See Figure 4.1). Analysis reveals that while access to a visualization improves problem solving success, having a collaborative process that encourages coordination and analysis efforts is also critical to success.

In-depth analysis of pairs from study two—of those with access to the shared visualization—isolated the obstacles and points of failure that pairs encountered as they attempted to find the serial killer. Pairs here succumbed to two classes of failures: (1) *coordination failures*; and (2) *sensemaking failures*. Evidence of coordination failures came in several forms (See Table 4.1). Dividing evidence between partners, for example, imports a task structure to the collaborative process that undermines performance. By contrast, partners with all of the evidence never divided search processes, and started by arbitrarily sifting through the documents. Additionally,
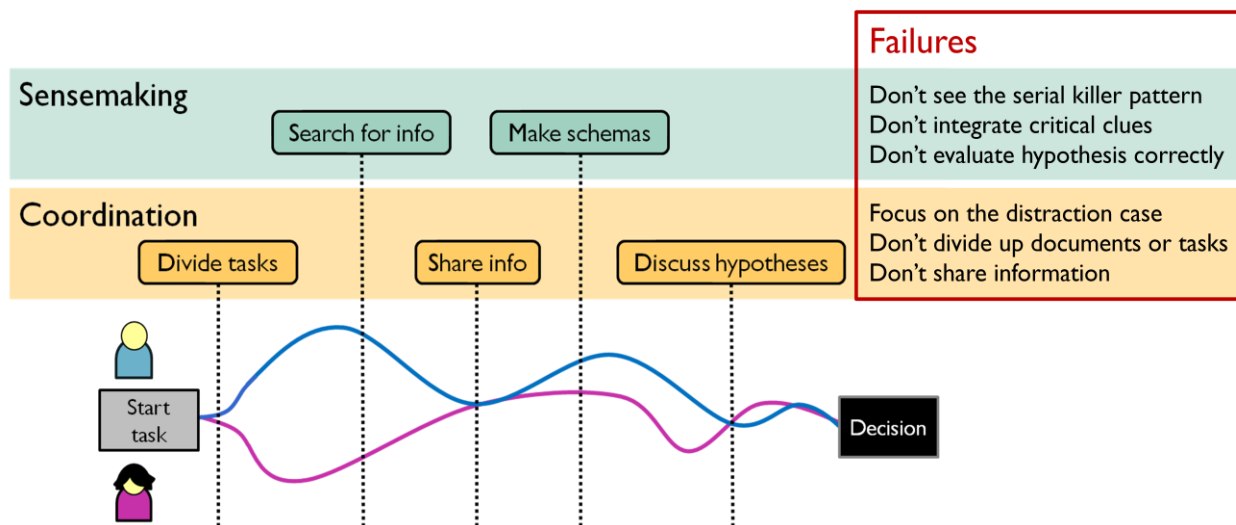
Figure 4.1. Failures or collaboration costs that pairs confront during their collaborative sensemaking or problem solving process.

pairs with all of the evidence ($M = 153$ words, $SE = 15.0$) talked significantly less and shared less information with their partner than their counterparts with only half the evidence ($M = 256$ words, $SE = 15.2$).

Additionally, pairs confronted *sensemaking failures* (See Table 4.2), oftentimes because they did not effectively search for information. In study two, 27% of pairs in the shared visualization condition with all of the evidence (versus 13% of pairs with half evidence) got off track and focused too narrowly on the unrelated homicide. This lack of focus represents an ineffectual search process.

As pairs continued their processes and their ideation became more sophisticated, they had to overcome analysis failures to be successful. However, pairs had difficulty integrating information into a cohesive narrative of facts. An astounding 30% of individuals with access to the shared visualization and all the information did not see any pattern they felt was indicative of a serial killer, as opposed to 16% of their half-evidence counterparts. These facts are astounding because the serial killer committed four of the seven murder cases. These pairs could not understand the significance of the information, so could not build upon that significance for future sensemaking or hypothesis generation. Even if pairs had all the evidence and a shared visualization, and even if they noticed important clues, such as the connection between two cases

	Coordination tasks		
	Divided up documents	Focused on distraction case	Sharing info: Average # IM words
Half evidence	Not applicable	13% (2/15 pairs)	$M = 256$ ($SE = 15.2$)
All evidence	0% (0/15 pairs)	27% (4/15 pairs)	$M = 153$ ($SE = 15.0$)

Table 4.1. Comparing the rates of coordination failures between the shared visualization, half evidence and shared visualization, all evidence conditions from Study 2.

	Sensemaking tasks		
	Missed serial killer pattern	Discussed facts but missed connection	Pursued wrong hypotheses
Half evidence	16% (5/30 individuals)	15% (2/13 pairs)	3% (1/30 individuals)
All evidence	30% (9/30 individuals)	64% (7/11 pairs)	27% (8/30 individuals)

Table 4.2. Comparing the rates of sensemaking failures between the shared visualization, half evidence and shared visualization, all evidence conditions from Study 2.

or the serial killer pattern or both—they were still unable to integrate all the independent facts into a cohesive whole. This was true for seven out of eleven pairs, in the all evidence condition, who saw at least a pattern for a serial killer, or a connection between two suspicious characters, or both. In contrast, only two out of thirteen pairs in the half evidence condition had difficulties integrating their facts. They discussed a pattern for a serial killer, or a connection between the serial killer and a witness, but they did not properly understand the connections necessary to incorporate facts into a solution.

Furthermore, pairs shared only two hypotheses with their partner, on average. This is a sign that their creativity and thoughtfulness during analysis was stunted, for some reason. Previously, it was suggested that pairs with all the information appeared to be plagued by confirmation bias. In other words, participants have suboptimal-evaluation process when they choose an incorrect hypothesis, especially over a correct alternative. When someone chooses the wrong suspect, it is obvious they have evaluated a hypothesis incorrectly. And only one individual in the shared visualization, half evidence condition, picked the wrong suspect, in contrast to eight individuals in the shared visualization, all evidence condition. These pariticpants did *not* return to their data, nor did they debate ideas rigorously enough to come to the correct solution.

One possible way to help pairs overcome or avoid these pitfalls is to intervene in their problem solving process. Facilitating the collaborative process through interventions could give pairs a more efficient process, or break them out of an inefficient process, allowing them time to reflect and process the information they have discovered so far, and encouraging them to use other available. Such facilitation may increase problem solving success and help pairs fulfill their potential to outperform individuals. In what follows, I describe a third study that explores whether interventions, in the form of discussion prompts, can lead pairs to a higher degree of collaborative success in complex problem solving.

4.2   FACILITATING COLLABORATIONS VIA INTERVENTIONS

An intervention is an action that affects another's affairs, especially a deliberate entry into a situation or dispute in order to influence events or prevent undesirable consequences (intervention, 2011). Interventions are meant to interrupt a certain course of action, to attain an alternate, preferred course. The conventional use of the word indicates the purposeful interruption of another's dysfunctional behavior (e.g., "to intervene," "to have an intervention"),

78

particularly when the dysfunction has detrimental effects for them or others. It has also been shown that interventions that encourage groups to target integral components of a task can help individuals who are going down the wrong path (Hackman, Brousseau, & Weiss, 1976).

Knowing that simple interventions can impact both individual and group problem solving success, it is a surprise that the visualization intervention of the first two studies did not lead to significantly better performance. Access to the visualization did *not* necessarily keep pairs on task or lend any more structure to the group problem-solving process. In study one, we saw that the visualization tool has the potential to help participants. Could that potential be improved through interventions that would make the visualization more powerful? And what types of interventions are appropriate for this type of collaborative problem-solving task?

4.2.1 OVERVIEW OF INTERVENTION LITERATURE IN SMALL GROUP RESEARCH

Prior research on team processes and theories as to why groups often fail to perform better than individuals is already available (Steiner, 1972). The hope here is that studies from both social psychology and organizational psychology could improve team effectiveness through interventions that target specific parts of the process. Researchers have organized different intervention types by various categorization schemas, usually based on whatever part of the problem-solving process their research focuses (Hackman et al., 1976; Miranda & Bostrom, 1999; Kozlowski & Ilgen, 2006). To better survey researcher's intervention types, I created a new schema to categorize group process interventions from data that has been extracted from
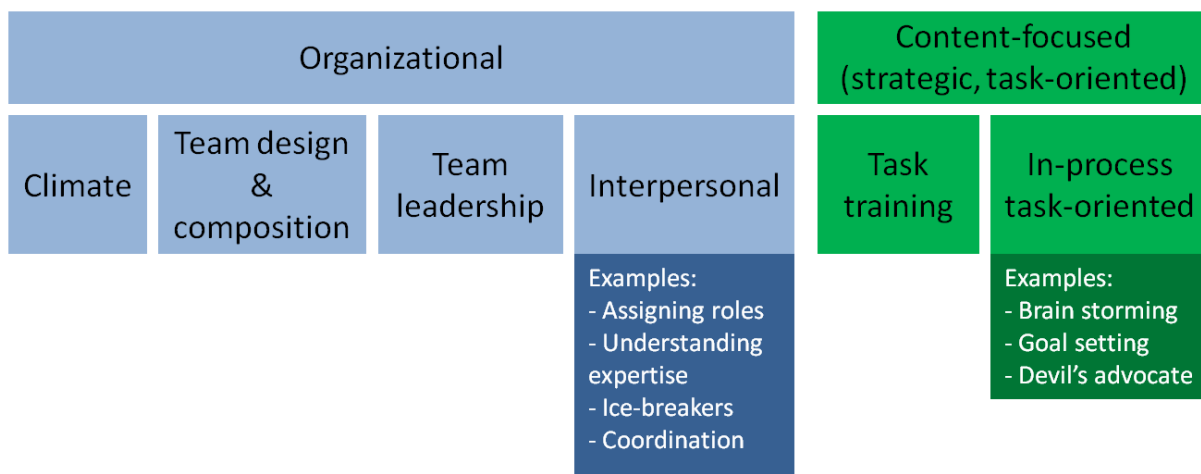
Figure 4.2. Schema of group intervention categories

variety of meta-reviews, theoretical constructs, and research studies (See Figure 4.2). In general, prior research has defined two main categories of group interventions: organizational and content-focused (Miranda & Bostrom, 1999; Woolley, 1998; Eden, 1990).

Organizational interventions

Organizational interventions center on the team itself, rather than on the team performing a task. Organizational interventions can be broken down into four types: (1) team climate; (2) team design & composition; (3) team leadership; and (4) interpersonal qualities (Kozlowski & Ilgen, 2006; Salas, DiazGranados, Klein, Burke, Stagl, Goodwin, & Halpin, 2008; Beer, 1980). Team climate refers to external attributes of the team, such as satisfaction with the work environment and openness to innovation (Schneider & Bowen, 1985; Anderson & West, 1998). Team design and composition interventions leverage research about how diversity and expertise can improve group effectiveness; it aims to improve team performance based on the actual members who compose the team (Woolley, Gerbasi, Chabris, Kosslyn, & Hackman, 2008; See Mannix & Neale, 2005 for a good meta-review on team-composition). Team leadership has been a hot topic within the business realm, and it often focuses on managerial leadership development as an intervention to promote highly productive teams (Collins & Holton, 2004). Interpersonal interventions focus on team building and relationships among groups; it aims to improve productivity and success (Salas, Rozell, Mullen, & Driskell, 1999; Bradley, White, & Mennecke, 2003; Klein, DeRouin, & Salas, 2006).

While all four types of interventions address critical components of group processes this work focuses on interpersonal interventions. These interventions can occur prior to the group task (in team building activities), and also during the group task (in improving information sharing). Early interpersonal interventions, such as the Delphi Technique and the Nominal Group Technique, aimed to *reduce* overall communication between team members (Linstone & Turoff, 1975; Delbecq, Van de Van, & Gustafson, 1975). The thought was that reduced communication efforts would allow a team to focus on the task at hand. However, other interventions focus on building team cohesion, trust, and also on ensuring proper coordination among functions of individuals within the group (e.g., open communication and division of labor). Ultimately, new ways to coordinate joint efforts among group members should reduce coordination costs within the collaboration. For example, just telling individuals to share relevant information, or to find

the most able group member, has been shown to improve group performance (Henry, 1995). Other research on interpersonal interventions, on the other hand, suggests that such prompts have mixed results. People function better as a team (i.e., team members perceive higher levels of team cohesion) but don't necessarily *perform* better (Lipshitz & Sherwood, 1978).

Studies one and two indicated that collaborators have difficulty with many facets of a multistage problem solving process, including difficulties with interpersonal tasks. Having good "interpersonal functionality," such as sharing information and dividing up labor during the task, was critical for being successful at this particular collaborative task—finding the serial killer before he kills again. Recall also that pairs with all of the evidence in study two never divided their documents with their partner; comparatively, they performed poorly. This leads me to believe that an interpersonal intervention that focuses on helping pairs coordinate their efforts will improve collaborative performance.

> *Hypothesis 1: Pairs who receive an interpersonal (or coordination) intervention will outperform those who receive no intervention.*

Content-focused interventions

The second main category of group interventions is content-focused; it is also sometimes referred to as "strategic" or "task-oriented" (Miranda & Bostrom, 1999; Woolley, 1998). These interventions, as opposed to organizational interventions, focus on the actual group task, rather than on team dynamics. Content-focused interventions can be further classified into two types: task training and in-process, task oriented. Training is a common tool used in high-pressure domains, such as in the military (Kozlowski & Ilgen, 2006; Cannon-Bowers & Salas, 1988). Classic theory on training says that teams learn best by doing, honing their skills through practice, repetition and/or simulations (Dyer, 1984). Training is very specific to the content and goals of the task. Many training sessions are designed to mimic potential real-life scenarios that the team will encounter, and training occurs prior to actually performing of the task.

In contrast, the second type of content-focused intervention is in-process, task-oriented. These interventions serve as aids to teams during the actual process of collaboration. Like training materials, task-oriented interventions are highly task-dependent. For example, a task that required idea generation may be improved with tactics such as brainstorming and presenting

groups with visual stimuli (Osborne, 1957; Wang, Cosley, & Fussell, 2010). For more complex tasks, such as evaluating multiple options, strategies such as "playing devil's advocate" would be more appropriate, a strategy still widely used and effective for helping individuals and groups achieve greater levels of success (Cosier, 1978). Such interventions can break up detrimental patterns, can help groups to adopt useful strategies, and can force problem solvers to think more critically (Hackman et al., 1976). As a consequence of these facts, I explore how in-process, task-oriented interventions might help pairs during their collaborative problem-solving process.

Studies one and two showed that collaborators had difficulty with task-oriented facets of a multistage problem-solving process; in particular, they had difficulty with sensemaking failures, such as the inability to find a pattern for a serial killer. Prior research has shown that in-process, task oriented interventions can help partners in the collaborative sensemaking process. Okhuysen and Eisenhardt (2002) have looked at how interventions impact knowledge integration within groups. Interruptions, such as encouraging individuals to question others, influence task pacing; this has been shown to enhance performance on ambiguous tasks. Okhuysen argues that such interruptions are vital to knowledge integration. By changing the focus of work from the larger primary task to a secondary subtask, to the intervention task, groups reflect upon the impact of the subtask on the primary task, leading to increased discussion and oftentimes to strategy changes. This research suggests that in-process, task-oriented interventions designed to help pairs focus and discuss subtasks of the sensemaking process may improve problem solving success. This line of thinking leads to a second hypothesis:

> *Hypothesis 2: Pairs who receive a task-oriented (or sensemaking) intervention will outperform those who receive no intervention.*

Combining organizational and content-focused interventions

Both interpersonal and task-oriented processes are vital to a team's effectiveness. While attending to one aspect may improve effectiveness, it is possible that interventions that address both aspects of team processes may be necessary to see performance gains (Zaccaro & McCoy, 1998; Reagon-Cirincione, 1994). However, too much process can become burdensome; it can increase the collaboration costs and take away resources that might be otherwise used for actual problem solving (Steiner, 1972). Our previous studies suggest that collaborators are already

overwhelmed by the problem-solving task itself, as well as by their access to information. If we provide participants with two types of intervention mechanisms, it may encourage discussion but also distract them from their task. Hence

> *Hypothesis 3: Pairs who receive both coordination and sensemaking interventions will outperform those who receive no intervention but not those who receive either the coordination intervention only or the sensemaking intervention only.*

4.3   METHOD

To test these three hypotheses a third laboratory experiment was constructed; again, it used the "detective mystery" paradigm of studies one and two (Balakrishnan et al. 2008, Balakrishnan et al., 2010). Study three, then, is a 2x2-level factorial with two interpersonal intervention conditions and two strategy intervention conditions (no intervention (control), coordination-only prompts, analysis-only prompts, and both prompts). Participants were randomly assigned to one of the four possible conditions.

4.3.1   PARTICIPANTS

One hundred twenty participants participated in the experiment, described as a "Detective Mystery Study" (73 female, 47 male; 64% U.S. born; age range 18-39, median age approximately 22.5). Participants were again recruited via online resources. As in previous studies, fluency in English was a requirement to participate in the study. In addition, participants were also required to be current undergraduate or graduate students. Participants were paid $15 for their participation and were told the experiment will last 1.5 hours. There were no demographic differences between the participants across conditions.

4.3.2   PROCEDURE

Participants were randomly paired and seated apart from their partner, such that they could not see their partner or their partner's workstation. Pairs role-played as detectives of a police department, collaborating remotely to identify a possible serial killer. As in studies one and two, pairs in study three also worked through a large collection of documents and reports to find a serial killer. Again, participants were each asked to complete two online reports on the results of their investigation.

Participants in the visualization conditions were trained to use NetDraw (See Figure 2.1), the visualization tool adapted for this study, or the Excel spreadsheet if they were in the control condition. Training took an average of 30 minutes.

After training, participants were left to work on the assignment for one hour. Just as in studies one and two, pairs were given an MSN Instant Messenger [IM] client and encouraged to use the client to talk with their partner. After an hour, or when the participants had completed their investigation and report, they each completed an online survey to elicit the evidence they used to identify the serial killer. (For a complete description of the task, see the methods section in chapter two).

4.3.3 DISTRIBUTION OF EVIDENCE

Each participant had access to all of the cases and documents relevant to catching the serial killer. As before, the evidentiary documents and reports were available online and could be opened, searched, put in different or new folders, and manipulated freely. To ensure that sufficient screen space was available to examine multiple documents at once, participants each had access to one 23" monitor. Also, participants were given paper versions of the instructions and worksheets.

4.3.4 VISUALIZATION ACCESS

Each participant had access to the *shared visualization*. Once again, the visualization tool is NetDraw v.2, a software application for drawing 2D social network diagrams. In the shared visualization condition, each member of the pair had access to NetDraw and an interactive and searchable social network diagram of all the evidence. This diagram could be manipulated and searched by both partners.

4.3.5 INTERVENTION OR DISCUSSION PROMPT MANIPULATION

Participants were randomly assigned to one of four intervention conditions: no intervention (control), coordination intervention, sensemaking intervention, and both interventions (See Table 4.3).

In the *no intervention* (control) condition, participants were given the same regular task instructions as other conditions, then told to begin the same detective mystery task of studies one

and two. These participants were not interrupted for the duration of their task, except for the usual time remaining notifications that interrupted all conditions.

In all three *intervention* conditions, participants were given the regular task instructions and instructed to begin the detective mystery task as in previous conditions. After participants had been working for 10 minutes, they were handed additional paper instructions that contained discussion prompts to help guide participants through their problem-solving process. Participants in the three intervention conditions were *not* given additional time to solve the case. Prior research on the timing of an intervention, within a collaborative process, has shown that such interventions are most successful after problem solvers have had time to develop an understanding of the problem, and then encouraged to reappraise their approach to the problem (Woolley, 1998). Pretesting of the intervention timing revealed that after 10 minutes most participants had begun working on the task and were prepared for discussion prompts.

The *coordination intervention* condition consisted of a discussion prompt that encouraged pairs to discuss particular strategies to achieve their goal; it specifically asked the pair to consider dividing up the documents and tasks. This specificity of the prompt was based on the fact that pairs with all the evidence, and the shared visualization from study two, lacked a coordination strategy for division of labor.

The *sensemaking intervention* condition consisted of discussion prompts that encouraged pairs to

Instructions	These instructions suggest a strategy by which you and your partner may go about your task more efficiently. Think of different ways to achieve your goal of finding a suspected serial killer and which approaches are best, given your allotted time. Please discuss your options with your partner.
Coordination intervention	Divide up the documents and work between you and your partner.
Sensemaking intervention	Find cases that you think are relevant to your task. Use the shared interactive diagram to find patterns or similarities between cases. Use the shared interactive diagram to find unusual patterns or links to focus on people of interest.

Table 4.3. Full text of discussion prompts given to participants by intervention condition.
Participants in the both interventions condition received both the coordination and sensemaking intervention discussion prompts.

discuss particular strategies to achieve their goal; specifically, they asked the pair to consider finding relevant cases, and suggested they use the visualization tool to find patterns or interesting linkages. Essentially, these prompts were strategies specific to the nature of the task. In addition, they served to keep pairs on track and focused.

In the *both interventions* condition, participants were given the coordination *and* sensemaking interventions.

4.3.6 MEASURES

As in study one and two, there were three main sources of data: participants' final reports, their posttest surveys, and the IM logs of their discussions. In addition, I examined participants' activity logs to better understand their exact actions and task process

Identifying the Serial Killer

Participants' correct identification of the serial killer was found in their written reports. Both individual and collaborative performances were analyzed, where collaborative performance depends on both members of a team correctly identifying the serial killer.

Discussion Process

How much a pair communicated was determined by counting the total number of IM words they exchanged during a session. Participants' discussion topics were also coded, line by line. In total, more than 6,300 lines of IM were analyzed (see Table 3.1). Data was divided between three independent coders. (Kappa = .74). As in study one and two, hypotheses were only counted as such the first time they were discussed, even if pairs revisited it after considering other hypotheses before returning.

Task Process

Video logs of participant's actions were analyzed to better understand how the manipulation of discussion prompts affected the task process. Each participant's activity logs were recorded using Camtasia Studio screen recording & video editing software (http://www.techsmith.com/camtasia/?gclid=CLaS-avxqKkCFUOo4AodYkZzKQ). Activity logs were also reviewed to examine how each participant used the visualization tool, what views participants found, and on what documents participants focused their attention on.

Finally, to ensure that correct identification of the serial killer reflected comparable insight across conditions, the posttest survey participant's recognition memory was also tested (with multiple choice questions) for the eight pieces of evidence that led to the serial killer.

Individual Characteristics

We used CRT scale scores as a control variable in our analyses to measure cognitive reflection. We also administered the NASA TLX scale, a measure of task workload (Hart & Staveland, 1988).

4.4 RESULTS

This section details the condition manipulations on collaborative and individual performance, communication, and visualization use. Results are based on data from 60 pairs (120 participants), 15 pairs in each of the four conditions.

4.4.1 PERFORMANCE MEASURES

I considered: (1) the effects of condition on our participant's collaborative performance in the serial killer task; (2) an individual member's performance in the serial killer task; and (3) in individual performance on the unrelated homicide case.

Pair level analysis

I explored the effect of our intervention on pairs who solved the serial killer task and pairs who did not (See Figure 4.3). For a pair to solve the case, both members of the pair had to correctly identify the serial killer. I predicted that pairs, given any type of intervention, would solve the serial killer task more often than those without intervention. To test this hypothesis, I conducted a logistic regression at the pair level to assess whether intervention condition predicted the dichotomous outcome. The condition did significantly predict solve rates (Likelihood Ratio $\chi^2 = 8.14$, $p < .05$, $df = 3, 60$; Cramer's Phi = 0.34). In fact, no intervention pairs performed the worst, with only 40% solving the case ($SE = 11.9$). Overall, and in comparison to pairs with no intervention, pairs who received any type of intervention solved the case significantly more often. Pairs who received the coordination intervention solved the case 73% of the time ($SE = 11.9$), pairs who received the sensemaking intervention solved the case 87% of the time ($SE = 11.9$), and pairs who received both interventions solved the case 60% of the time ($SE = 11.9$).
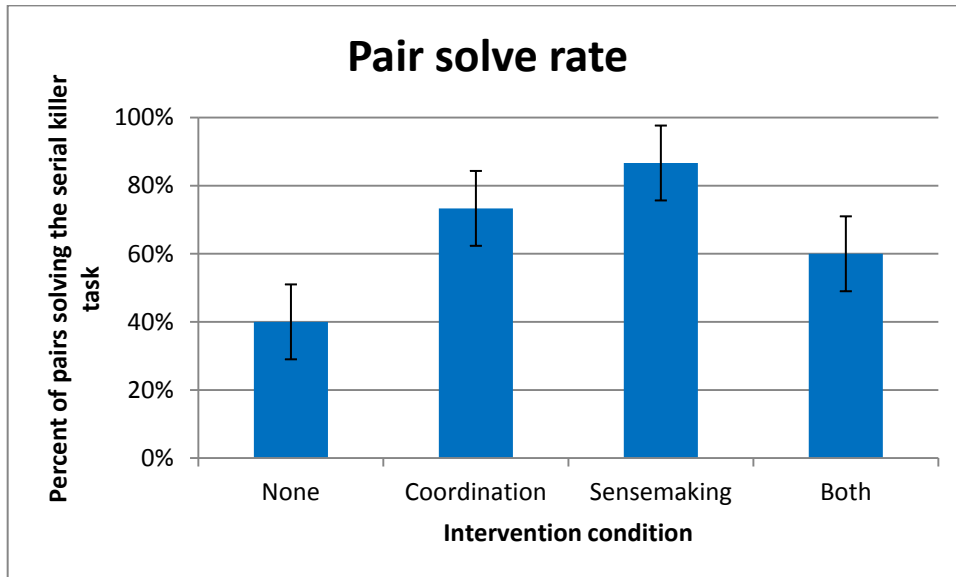
Figure 4.3. Pair solve rate (n = 60).

Remember that I had also predicted that not all intervention conditions would be equal. I argued that coordination intervention pairs and sensemaking intervention pairs would solve the case more often than both intervention pairs. Post hoc comparisons of all four conditions, using Fisher's exact test for pairwise comparisons, show that pairs in the sensemaking intervention condition were significantly more likely to solve the case than no intervention pairs ($p < .01$); meanwhile, pairs given the coordination intervention strongly tended to solve the case more often than pairs with no intervention ($p = .07$).

Individual level analysis

Next, I considered if the intervention condition resulted in individuals identifying the serial killer (See Figure 4.4). There was a main effect for intervention condition (Likelihood Ratio $\chi 2 = 9.57$, $p < .05$, $df = 3, 120$; Cramer's Phi = 0.28). And similar to pair performance, sensemaking intervention and coordination intervention in individuals significantly outperformed individuals given no intervention (no intervention $M = 53\%$, $SE = 8.18$; coordination intervention $M = 77\%$, $SE = 8.18$; sensemaking intervention $M = 87\%$, $SE = 8.18$; both interventions $M = 63\%$, $SE = 8.18$).
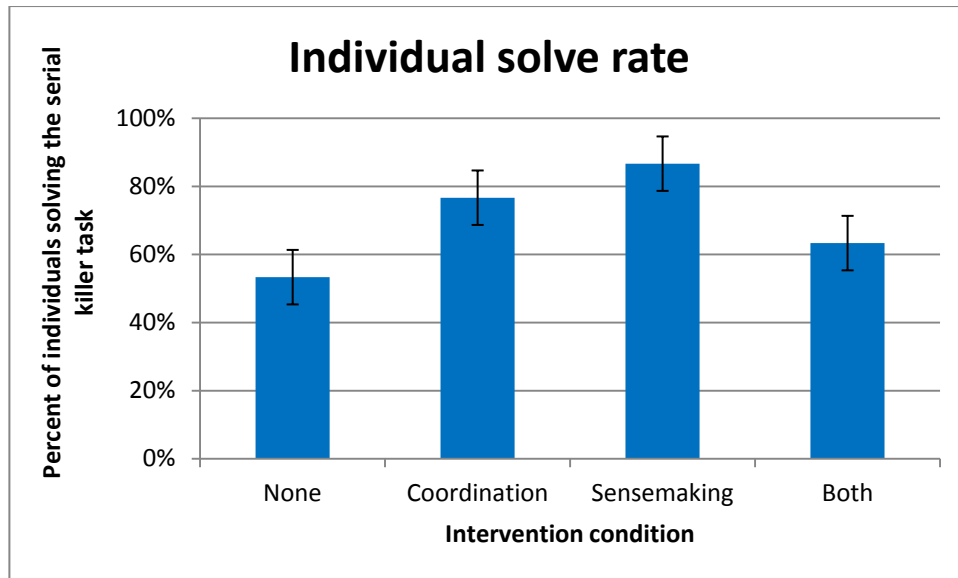
Figure 4.4. Individual solve rate (n = 120).

Decision consensus

While these results for individuals are similar to pair performance, closer examination reveals that no intervention individuals performed better than no intervention pairs (53% individuals versus 40% pairs). This discrepancy is due to the fact that pair performance requires both members of the pair to correctly identify the serial killer in their final reports. But I wanted to know whether pairs with no intervention were more likely to have a disagreement with their partner. As it turns out, no intervention pairs were more likely to disagree more often than sensemaking intervention pairs (Likelihood Ratio $\chi2 = 6.16$, $p < .05$, $df = 1, 30$; no intervention $n = 4/15$ pairs disagreed; coordination intervention $n = 1/15$ pairs; sensemaking intervention $n = 0/15$ pairs; both interventions $n = 1/15$ pairs). This suggests that the discussion prompts used in the intervention conditions helped pairs reach consensus, leaving no intervention pairs at a disadvantage.

Solution trends

Because so many pairs failed to identify the serial killer, I coded each participant's solution for his or her progress towards a solution based on two different facets: pattern recognition and decision correctness. The pattern recognition scale focused on whether or not participants identified a pattern for a serial killer between four out of the seven cases, a critical insight required to solve the case. This was a dichotomous variable. I tested the effects of intervention

condition on pattern recognition (See Figure 4.5). It *did* predict whether or not individuals saw a pattern (Likelihood Ratio $\chi 2 = 11.47$, $p < .01$, $df = 3, 120$; Cramer's Phi $= 0.31$). Eighty-six percent (26 out of 30, $SE = 5.54$) of individuals with no intervention saw the pattern necessary to identify the serial killer, as did 100% ($SE = 5.54$) of coordination intervention individuals, and 93% (28 out of 30, $SE = 5.54$) of sensemaking intervention individuals, and 77% (23 out of 30, $SE = 5.54$) of both intervention individuals. Fisher's exact test for pairwise comparisons revealed that coordination intervention in individuals resulted in seeing a pattern more often than those with no intervention or both interventions. Individuals who received the sensemaking intervention also saw a pattern more often than those with both interventions ($p < .05$).

Although finding a pattern among the cases was a major step towards the correct solution, participants also had to correctly identify and decide upon their primary suspect. Each individual's solution was rated on a scale from zero to two on how correct their decision was. A zero represented that individual chose an incorrect suspect. A one represented solutions wherein the individual was only partially correct; e.g., they identified the correct suspect, as one of several possible suspects, but failed to choose the correct possible suspect. A score of two was reserved for those solutions that declared outright the correct suspect as the serial killer. I
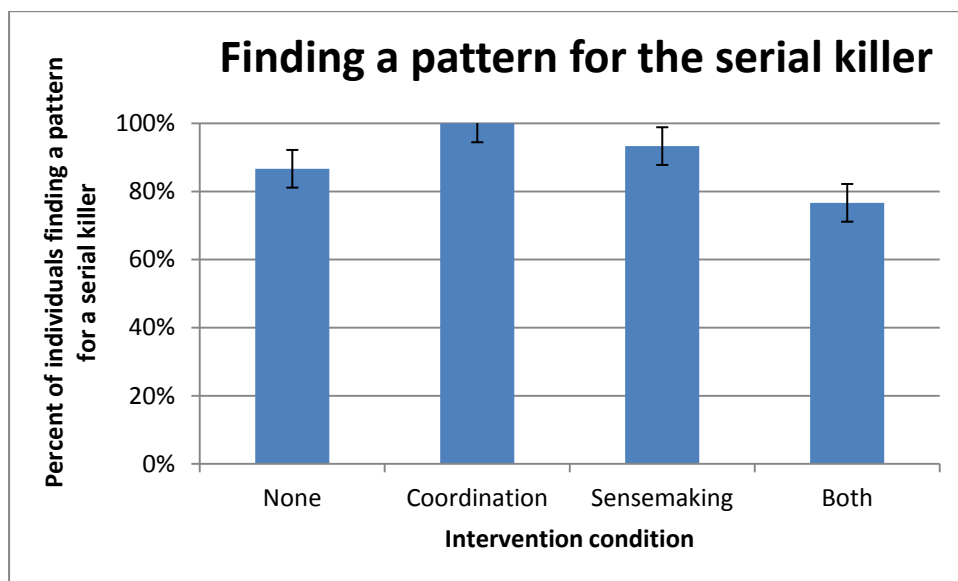
Figure 4.5. The percent of individuals who found a pattern for a serial killer by condition.

To find a pattern for a serial killer, participants had to correctly identify four out of the seven cases as being related or connected in their final report.

considered the effect of condition on decision correctness (See Figure 4.6). I conducted a logistic regression; the level of decision correctness was the dependent variable; the intervention condition was a between group factor. I then found that a significant main effect of condition on correctness levels (Likelihood Ratio $\chi2 = 18.49$, $p < .01$, $df = 6, 120$; Cramer's Phi $= 0.28$). Pairwise comparisons of conditions showed significant differences between sensemaking intervention and no intervention individuals (Likelihood Ratio $\chi2 = 12.08$, $p < .01$, $df = 2, 60$). Additionally, there were marginally significant differences between sensemaking and coordination interventions (Likelihood Ratio $\chi2 = 5.87$, $p = .05$, $df = 2, 60$), as well as in both interventions (Likelihood Ratio $\chi2 = 5.14$, $p = .07$, $df = 2, 60$). The coordination intervention condition was also significantly different than both interventions (Likelihood Ratio $\chi2 = 6.29$, $p < .05$, $df = 2, 60$).

Sensemaking intervention individuals had the highest solve rates; they also had high rates of finding a pattern for the serial killer, and they most often decided on the correct suspect. These
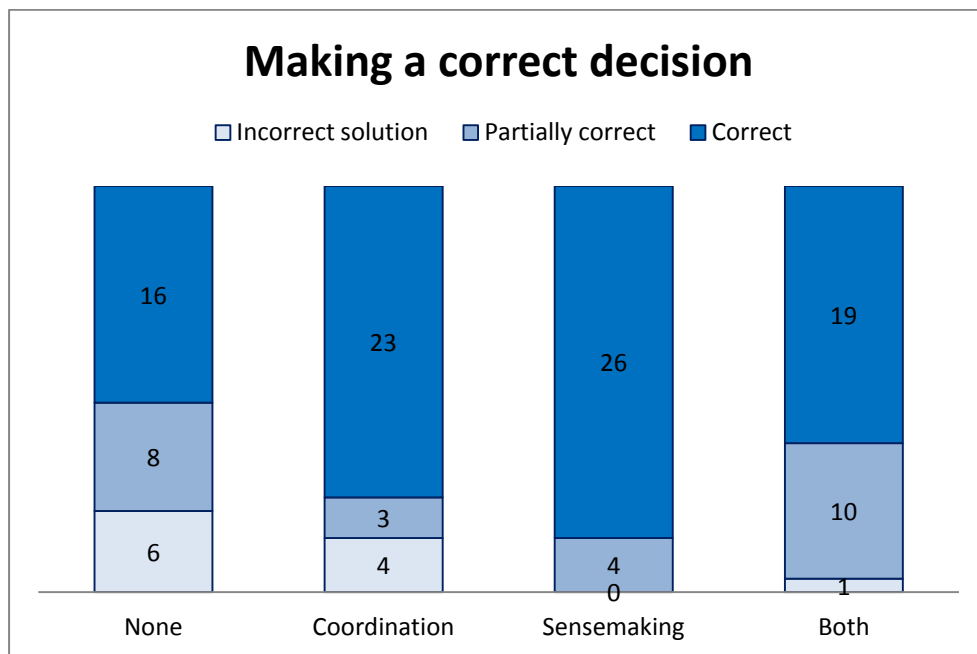
Figure 4.6. Number of individuals per level of decision correctness.

An incorrect solution means that the participant identified an incorrect suspect as the serial killer. Partially correct solutions had either a pattern for a serial killer without any primary suspect or had the correct suspect among several other suspects. A correct solution required that the participant explicitly identified the correct suspect as the serial killer.

results suggest that the sensemaking intervention elevated the effectiveness of participant's search and analysis process, increased their potential to find the hidden pattern, and helped them to the correct decision based on evidence rather than conjecture.

Distraction case solve rate

I explored whether or not participants had solved the unrelated homicide case, or "Raffield" case, which functioned as a distraction from the serial killings. Identifying the serial killer was negatively correlated with solving the Raffield homicide, $r = -.40$. In previous studies, I had found that participants who do not solve the serial killer case often fail because they waste their efforts on the unrelated, distraction case. In study two, participants with access to all the evidence and a shared visualization got off track and focused their attention on the Raffield case more often than their half-evidence counterparts. Hence, I wanted to know if the discussion prompt interventions helped individuals maintain their focus and avoid solving the Raffield case.

I considered whether intervention condition affected whether or not individuals solved the Raffield homicide case (See Figure 4.7). Condition did significantly predict Raffield solution rates (Likelihood Ratio $\chi 2 = 14.35$, $p < .01$, $df = 3, 120$; Cramer's Phi $= 0.35$). Individuals in the no intervention and both interventions conditions solved the Raffield case most often, 30% and 21% ($SE = 6.63$) of the time respectively. Those in the coordination and sensemaking
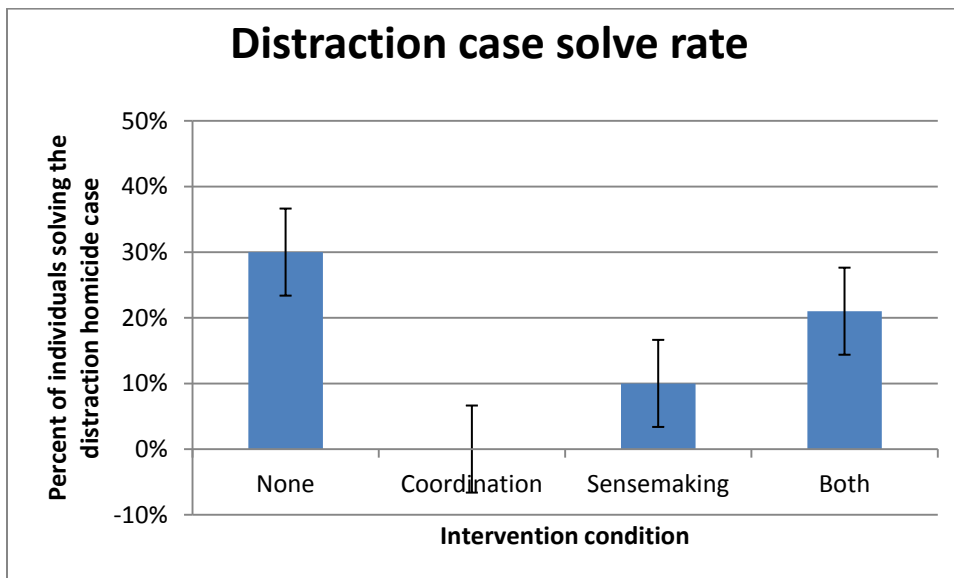
Figure 4.7. Percent of individuals per condition that solved the distraction Raffield homicide case.

intervention conditions solved the Raffield homicide less, 0% and 10% ($SE = 6.63$) respectively. Fisher's exact tests of pairwise comparisons showed that individuals in the no intervention condition solved the Raffield case significantly more often than individuals in the coordination intervention and sensemaking intervention conditions ($p < .05$). Individuals given both interventions solved the case significantly more often than those in coordination intervention condition ($p < .05$).

In summary, discussion prompt interventions did improve collaborative problem solving. The coordination intervention helped pairs find a pattern while sensemaking prompts helped pairs find a pattern and make the right decision. However, pairs given both interventions did not reap the same benefits; these pairs lost focus and wasted time on the distraction case.

4.4.2 COMMUNICATION

Next, I explored IM conversations between pairs to uncover possible factors of success (See Figure 4.8 and Figure 4.9). In this section, I review how much pairs talked, what topics they discussed, what evidence they shared, and the types of hypotheses they generated.

Total Talk and Topic Discussion

In the previous two studies, the more pairs talked over IM, the better they performed. I believed that pairs given any sort of discussion prompt interventions would be encouraged to talk more
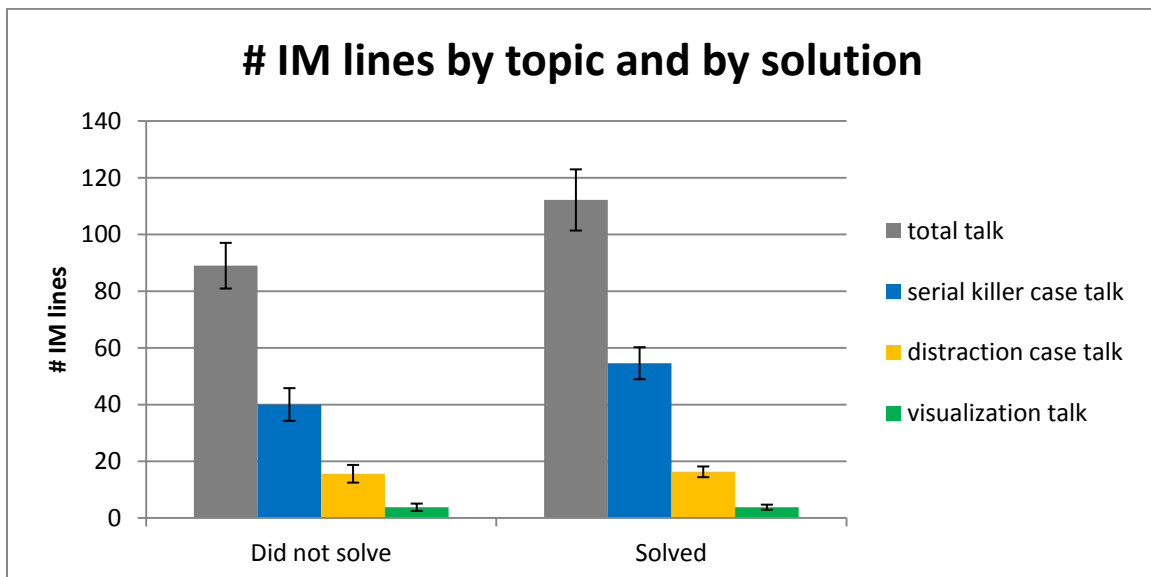
Figure 4.8. Average number of instant messaging lines by topic across solutions.

with their partner. However, total talk was *not* correlated with solving the serial killer case, and there was no significant effect of condition on the total amount of IM chat between partners (F [3, 56] = .81, $p = ns$). (See Figure 4.8 and Figure 4.9).

Next, I looked at topics of conversation between pairs. Previous studies showed a strong positive relationship between talking about the serial killer case and solving the case. However unlike previous studies, discussion of the serial killer case was not correlated with problem solving success (See Figure 4.8). I then considered if there were differences between conditions and how much time they spent chatting about the serial killer case. I was curious to see if the discussion prompt interventions encouraged pairs to focus their discussion on matters that related to their main task. I added the total number of IM lines pairs spent talking about the serial killer case and then divided the sum by the pair's total IM talk to control for variations of talk amount. What I saw was a trend that indicated that condition did impact the amount of time pairs spent
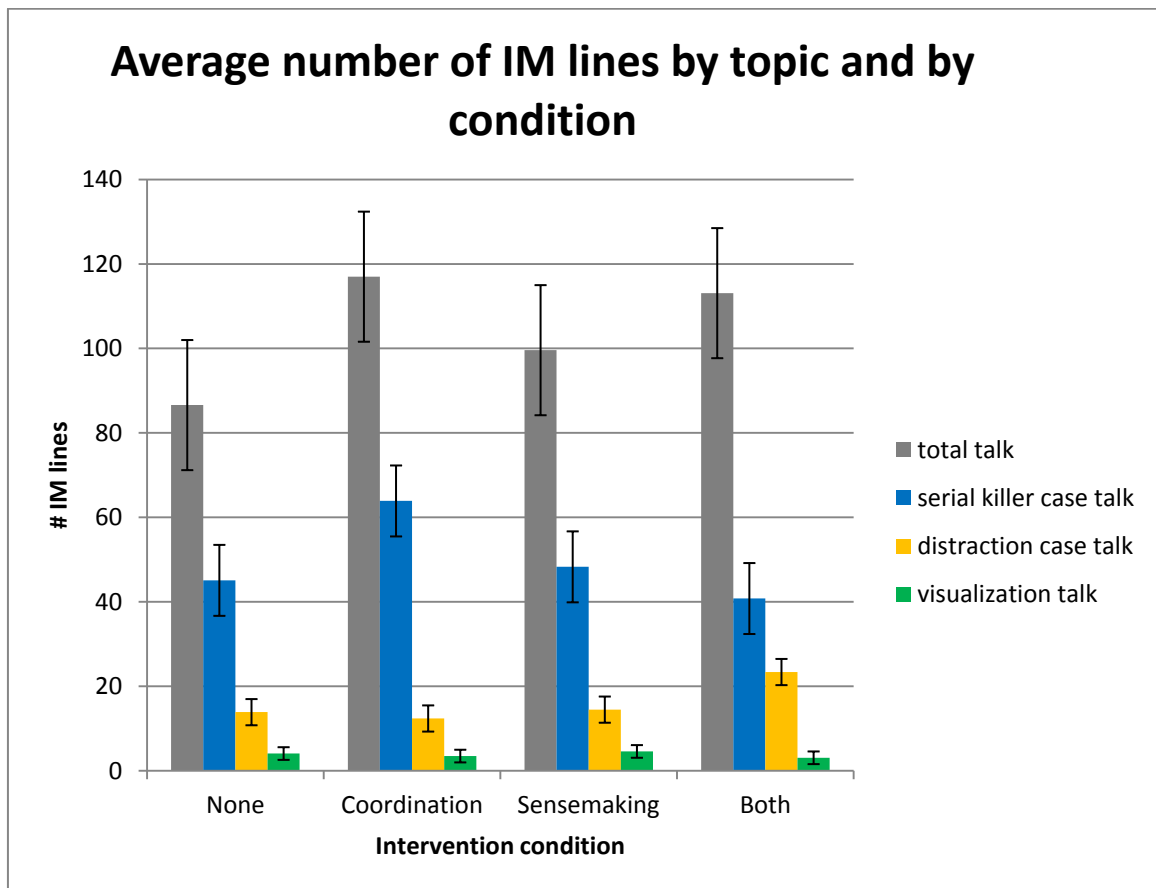
Figure 4.9. Average number of instant messaging lines by topic of conversation across conditions.

discussing the serial killer task (F [3, 56] = 1.99, p = .1). Post hoc comparisons between conditions using a student's t-test showed that coordination intervention pairs spent significantly more time discussing the serial killer task than their both interventions counterparts (M = 54%, SE = 4.5; M = 39%, SE = 4.5, respectively; p < .05).

The second topic of conversation I coded for was the unrelated Raffield homicide case. My motivations for exploring this factor were similar to those of exploring the distraction case solve rates. In previous studies, the percent of total discussion spent on the Raffield case indicated that the pair had gone off track. I added the total number of IM lines spent talking about the Raffield case and divided that sum by the pair's total IM talk to control for variations of talk amount. I looked at whether condition predicted the percent of total discussion time spent on the distraction Raffield homicide case. Here, the condition had a significant effect (F [3, 56] = 3.82, p < .05). A student's t-test showed that both interventions pairs (M = 24%, SE = 3.1) talked significantly more about the distraction case than those give the coordination intervention only (coordination intervention M = 9%, SE = 3.1, sensemaking intervention M = 18%, SE = 3.1, both interventions M = 16%, SE = 3.1).

A logistic regression was performed to better understand the effect of condition on the percent of total talk devoted to the Raffield case, with an interaction factor for solving the serial killer task. There was a marginally significant main effect for condition (Likelihood Ratio χ2 = 7.07, p =.06, df = 3, 60) and a significant interaction effect (Likelihood Ratio χ2 = 8.06, p < .05, df = 3, 60). This analysis revealed that the amount of "Raffield talk" did not impact the likelihood of solving the case for pairs in the sensemaking or coordination intervention conditions. However, for those in the no intervention condition, the more pairs discussed the Raffield case, the more likely they succeeded in the serial killer task. Conversely, for pairs in the both interventions condition, the more they talked about the Raffield case, the less likely they were to succeed in the serial killer task. This analysis echoes solution rates data; i.e., pairs with both interventions were led astray and lost focus, thereby lowering their chances of correctly identifying the serial killer.

Third, IM conversations were coded for the number of lines that referenced the visualization tool. I had reasoned that pairs given the sensemaking intervention would have more talk focused on the visualization because the discussion prompts explicitly encourage the use of the

visualization tool. But I found no significant effect of condition on the total percentage of talk referring to the visualization (F [3, 56] = .25, $p = ns$).

Discussion of Evidence

I hypothesized that discussion prompt interventions would increase sharing between pairs; specifically, I expected pairs to share more pieces of evidence critical to solving the problem. As in study two, I counted the number of critical pieces of evidence partners shared with each other in their IM discussion, then compared that number to the number they recalled in the posttest survey. This analysis gave us the percent of evidence discussed between pairs compared to the total critical evidence recalled. In an ANOVA, the percent of evidence discussed was the dependent variable, the evidence condition and visualization condition were between group factors, and CRT scores functioned as a control. And I found that there were no significant effects on the percent of evidence discussed.

Hypothesis Generation

I had also hypothesized that pairs given discussion prompt interventions, in particular the sensemaking intervention, would increase a pair's hypotheses generation and sharing over IM. But there were no significant differences across condition in the total number of unique hypotheses discussed by pairs.

4.4.3 VISUALIZATION USE

Next, I explored how much time individuals spent using the visualization, as well as an individual's use of particular visualization features. In study one, I found that time spent using the visualization was highly correlated with solving the serial killer task. In this study, however, the percent of time spent using the visualization was not correlated with solving the case (r = -.01, ns). Additionally, there were no significant differences in time spent on the visualization by condition. (But note that five videos were not captured properly so could not be used in analyses.)

I wanted to dig deeper into participants' visualization tool use and explore specific actions or views they encountered. Did the interventions change the manner in which individuals interacted with the visualization? I had hypothesized that discussion prompt interventions, specifically the sensemaking intervention that directs users to use the visualization during analysis, would

encourage individuals to engage with the visualization tool and find critical insights that could help them during the course of their investigation. I came up with four attributes of visualization use that could potentially impact problem solving success, which would signal more sophisticated, engaging use with the visualization tool. (Please view Figure 4.10 through Figure 4.14 for visual descriptions of these attributes). The first aspect is the search capability within the network diagram (See Figure 4.11). Participants could search for people with common characteristics, such as people who shared a place of employment or a mode of transportation (e.g., the bus). Here, I was mainly interested in whether or not participants used this functionality at all. Figure 4.15 and Figure 4.16 show the percent of participants who used the attribute search functionality at least once during the course of their session. Overall, most all individuals *did* search the visualization at least once, and there were no differences across condition. This is important because it shows that participants were actively engaging with the visualization, not passively consuming its pre-formatted, initial state.

Logistic regression helped me understand the effect of condition, attribute search, and the interaction factor for solving the serial killer task (See Figure 4.17). There was a significant main effect for condition (Likelihood Ratio $\chi2 = 9.62$, $p < .05$, $df = 3, 120$), and a marginally significant interaction effect (Likelihood Ratio $\chi2 = 3.21$, $p = .07$, $df = 1, 120$). This analysis revealed that for those in the no intervention condition, the use of attribute search functionality did not help them solve the case. Conversely, for pairs in the both interventions condition, the use of attribute search functionality did improve their chances of success. This analysis suggests that the discussion prompt interventions make the visualization tool more powerful, which ultimately leads to higher rates of complex problem-solving success.

I was also interested in participants finding a specific pattern within the visualization (See Figure 4.12). If participants searched for victims with a common characteristic, such as the time of day they were killed, the type of weapon used to kill them, or the part of the body that was injured, a pattern should quickly emerge between serial killer's four victims. Seeing this pattern on the visualization could point participants to the appropriate cases, helping their analysis process. I looked at whether or not participants had come across this visual pattern during their visualization use at least once. To find this pattern, participants needed to use the search functionality and, as expected, these two attributes were highly correlated ($r = .27$, $p < .01$).
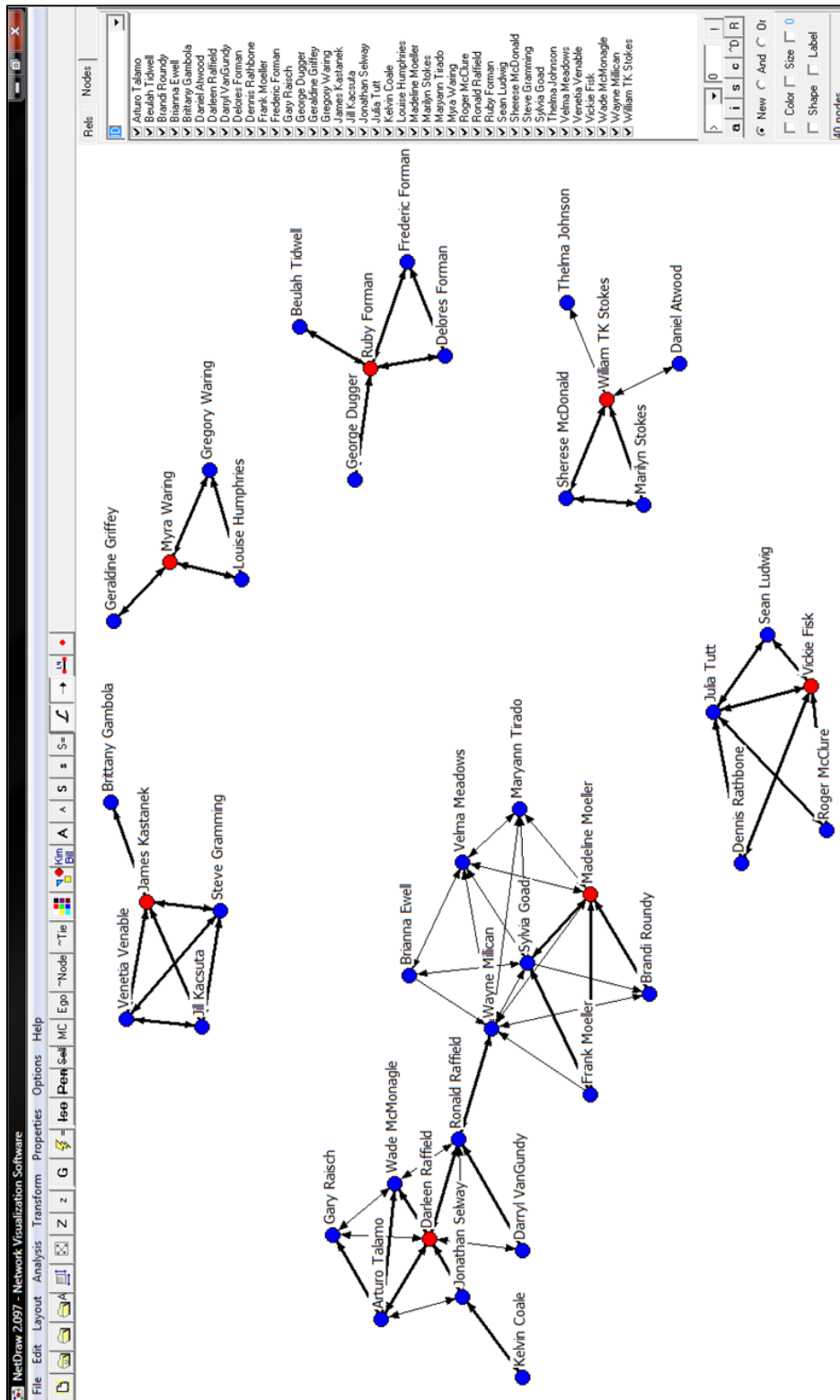
Figure 4.10. Image of NetDraw, the visualization tool used in all three laboratory studies. Here, it is shown in its initial state.
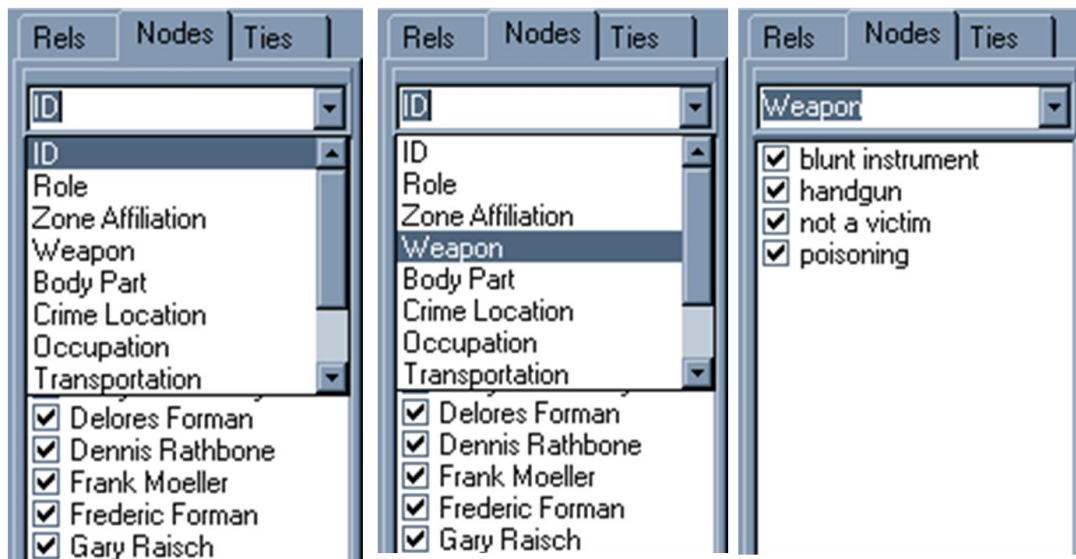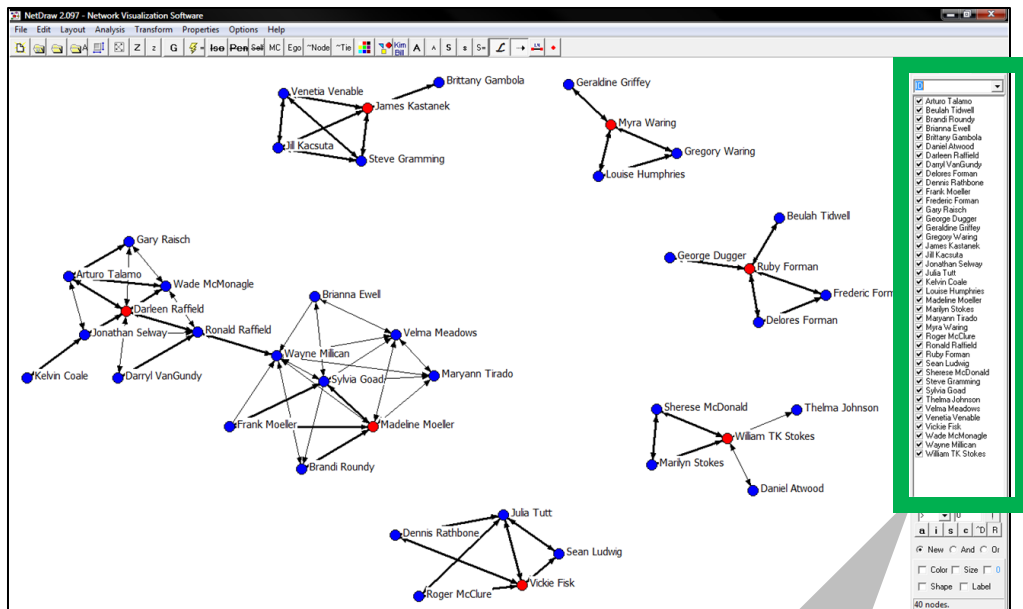
Figure 4.11. Search functionality within NetDraw.

The green box on the right hand side of the Netdraw window in the top image outlines the search box within the visualization application. The bottom three callout images represent a possible search within the application. Participants could use a drop down menu to search for characters in their case files with common characteristics, such as role (i.e., whether they were a victim of a crime or a suspect), zone affiliation (i.e., which police district zone they were affiliated with), etc. In the middle bottom image, the participant picks the attribute "weapon," which calls up the four attribute options that can be seen in the bottom right image.
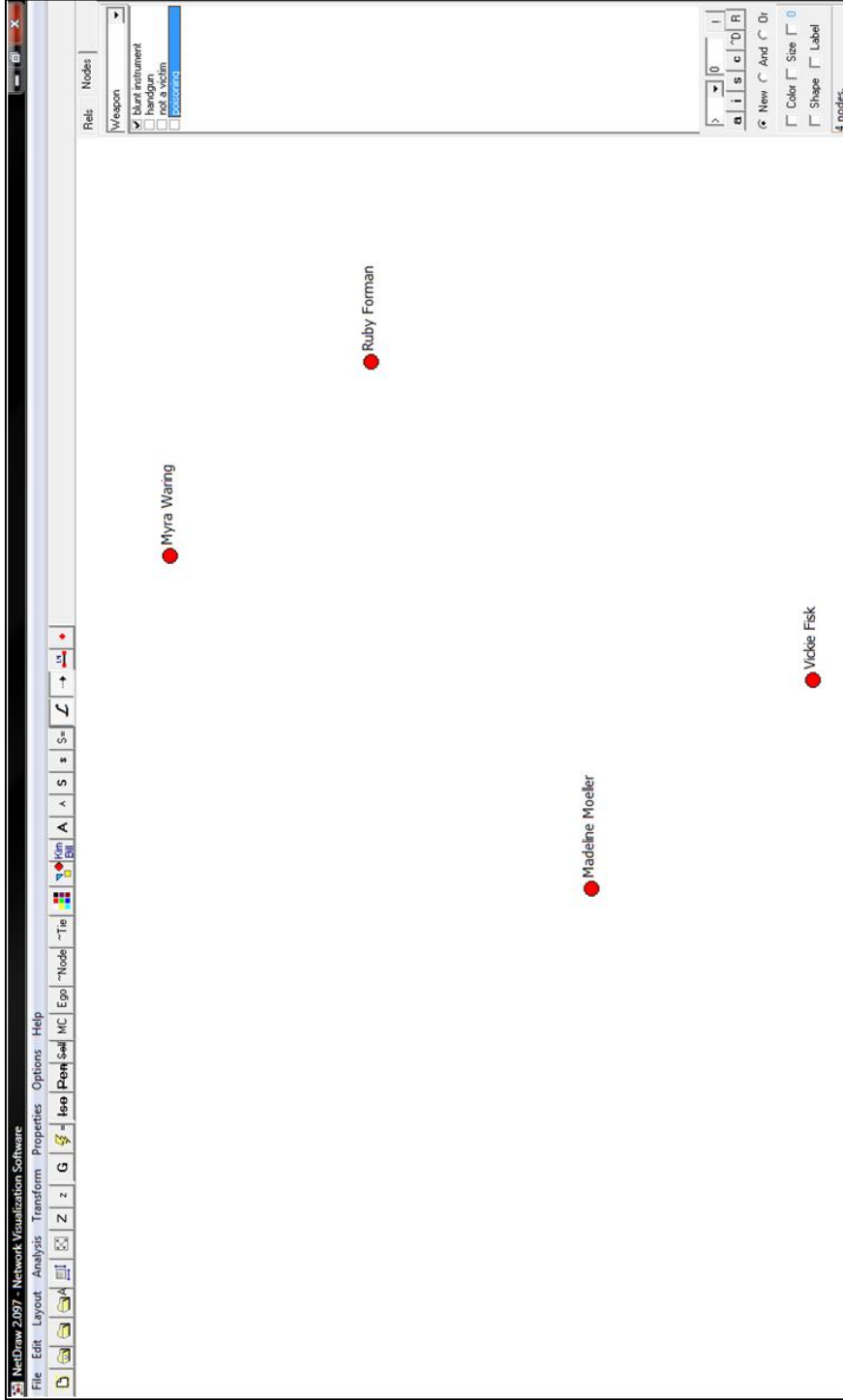
Figure 4.12. Pattern for serial killer in the visualization.

If participants searched for victims who were murdered by "blunt instrument" the network visualization tool would present the image above. These four nodes represent the four victims who compose the pattern for a serial killer.

Figure 4.13. Unusual link between the serial killer and another fictional character.

In the network diagram depicting the fictional characters from all the case files, there were two cases that were linked through the serial killer, Wayne Millican, and the husband of another murder victim, Ronald Raffield. Participants could notice this unusual link which could potentially lead them to a critical clue.
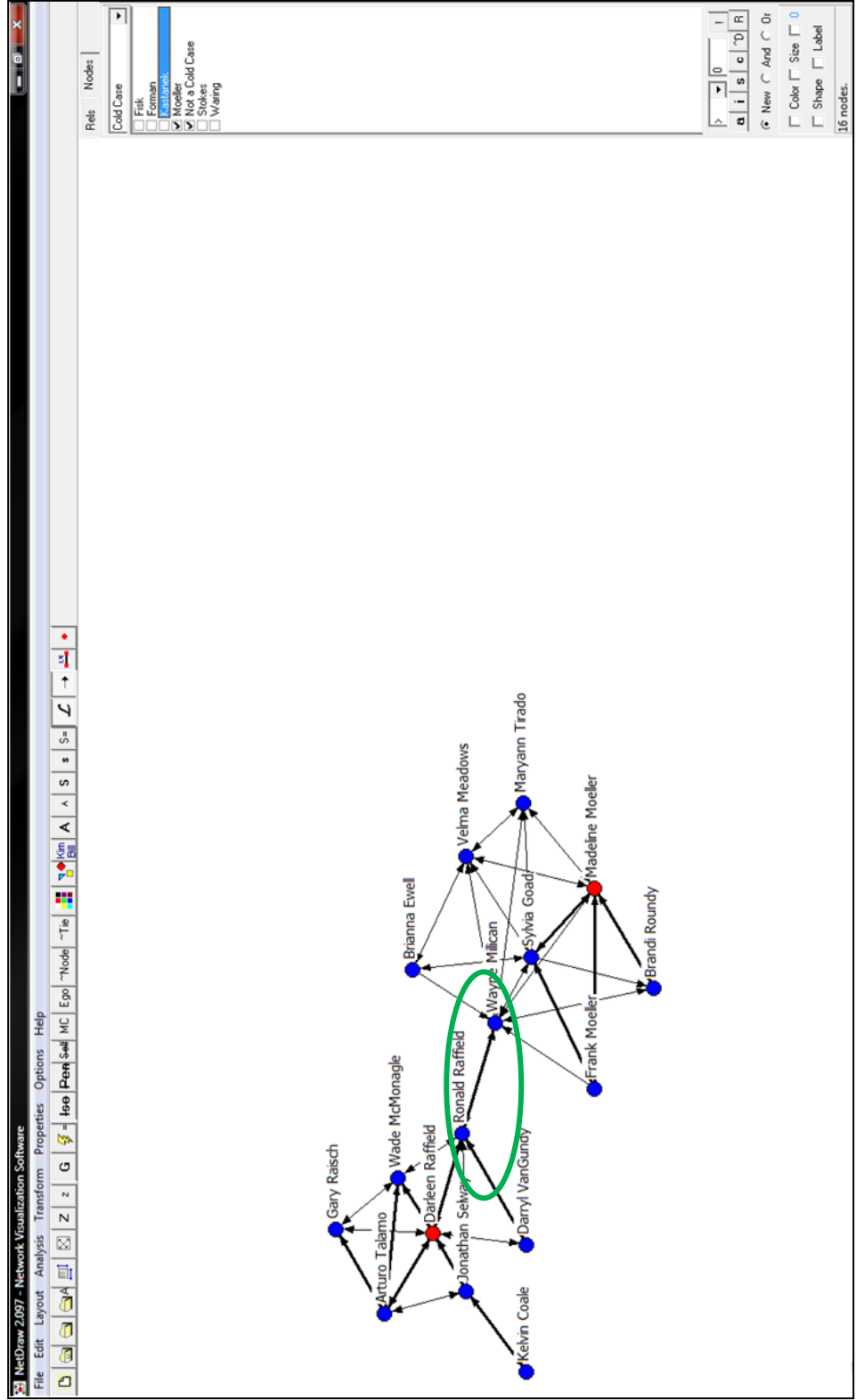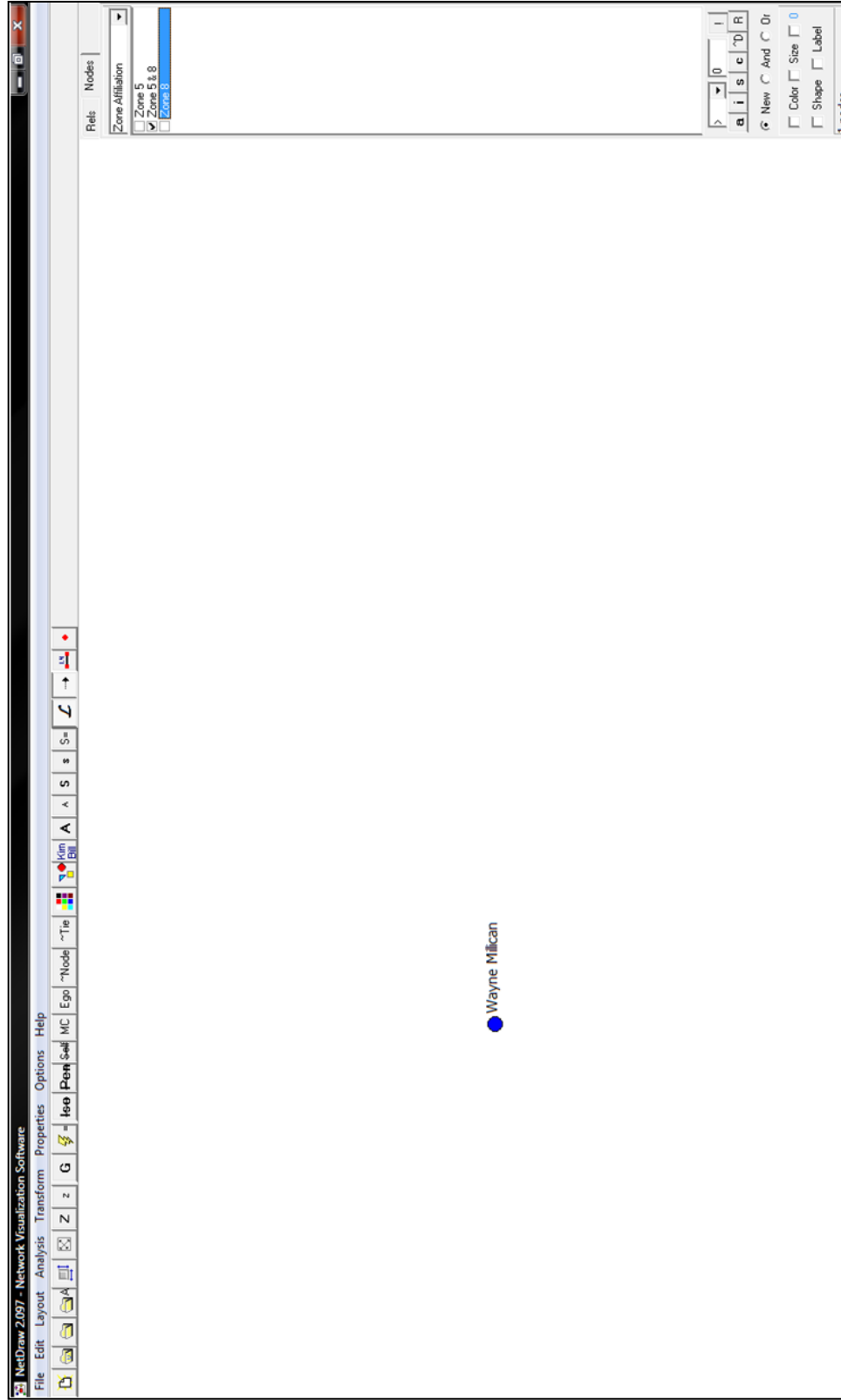
Figure 4.14. Zone affiliation, highlighting the serial killer's unique multiple geographic zone affiliation.
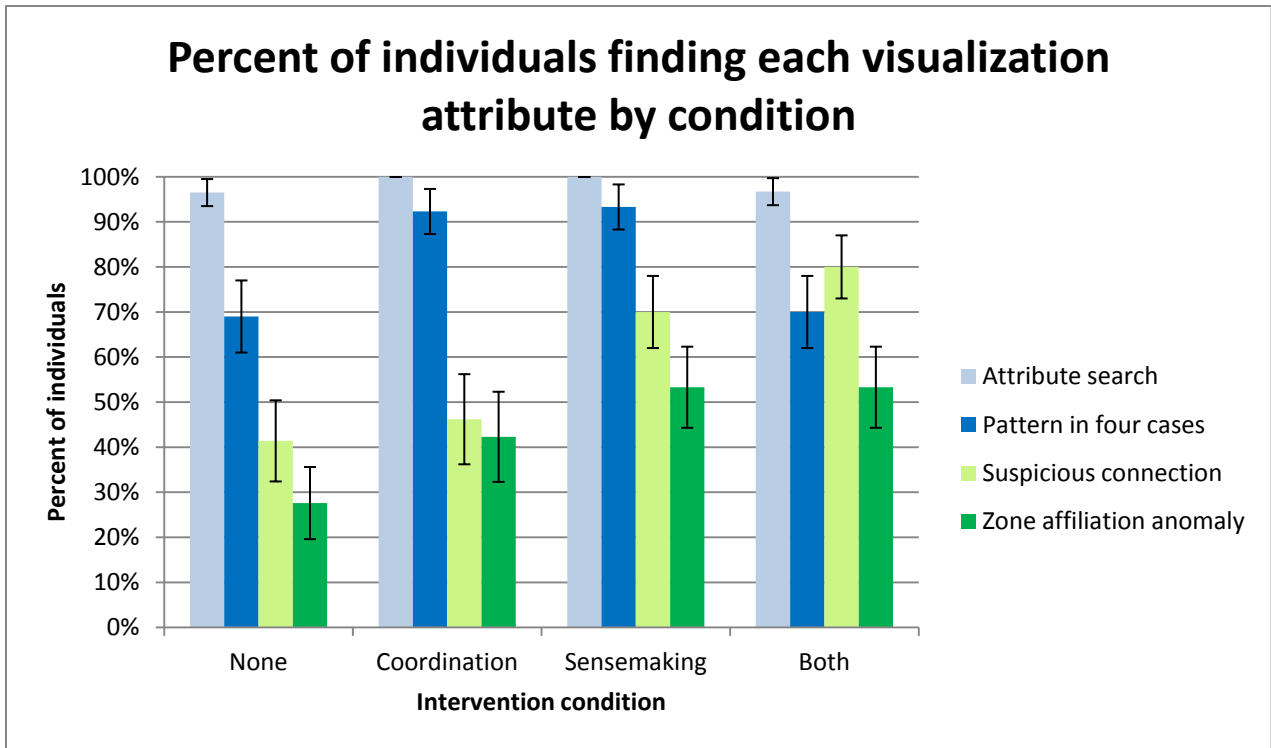
Figure 4.15. Percent of individuals finding each visualization attribute by condition.
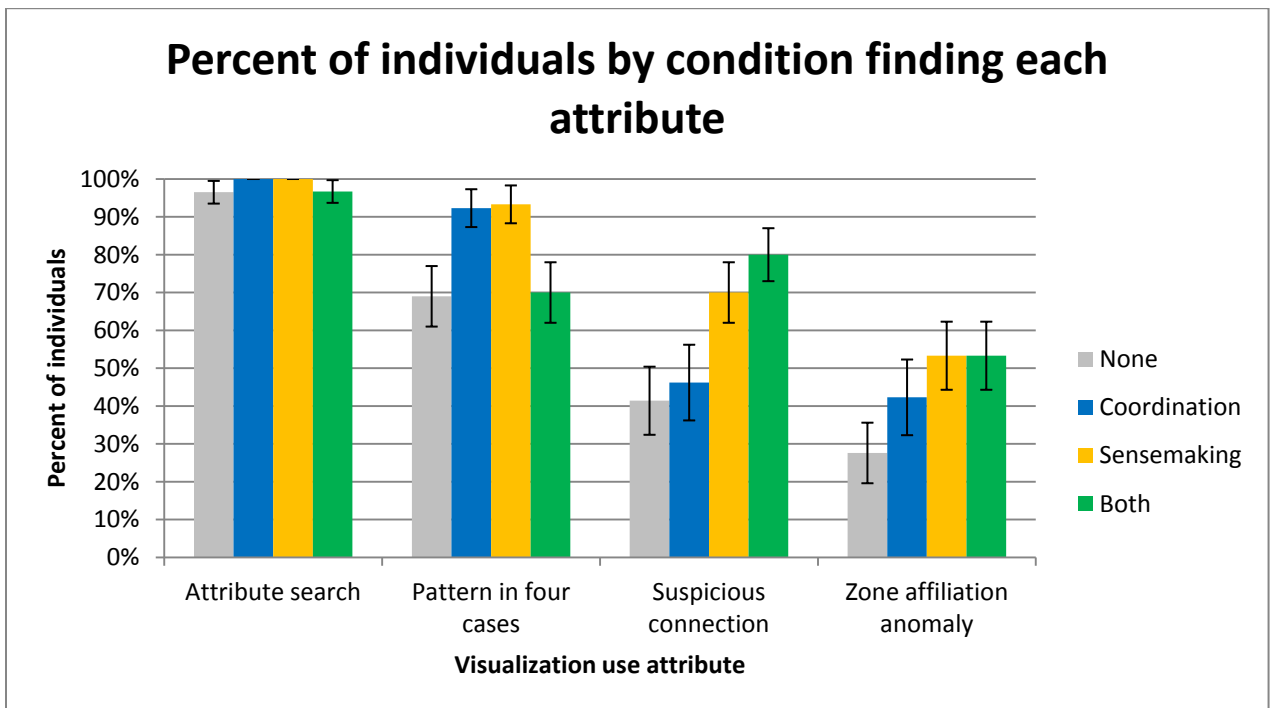
Figure 4.16. Percent of individuals by condition finding each visualization attribute.
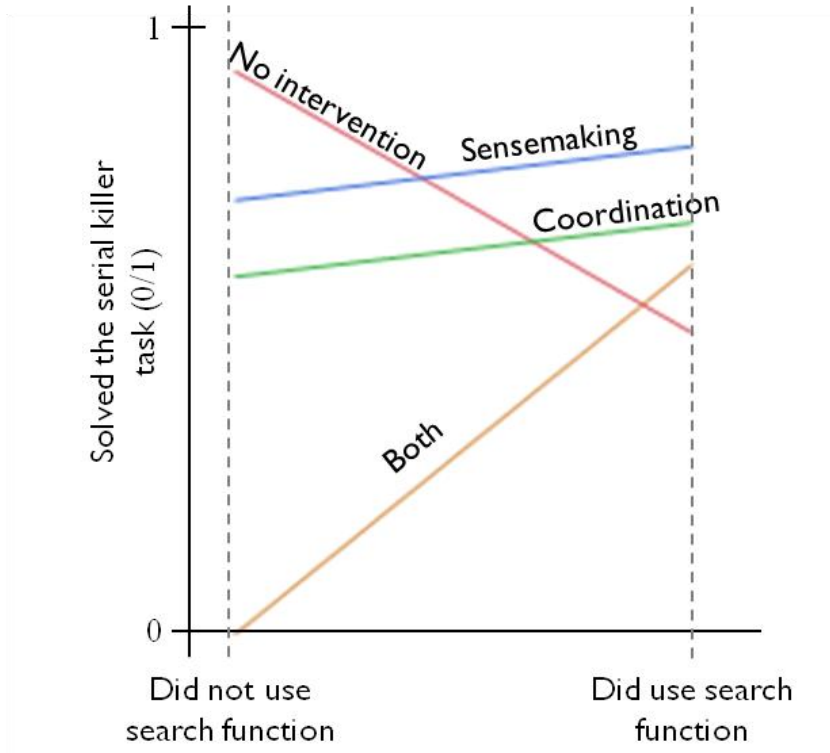
Figure 4.17. Shows the interaction profile of condition and attribute search function for solving the serial killer task.
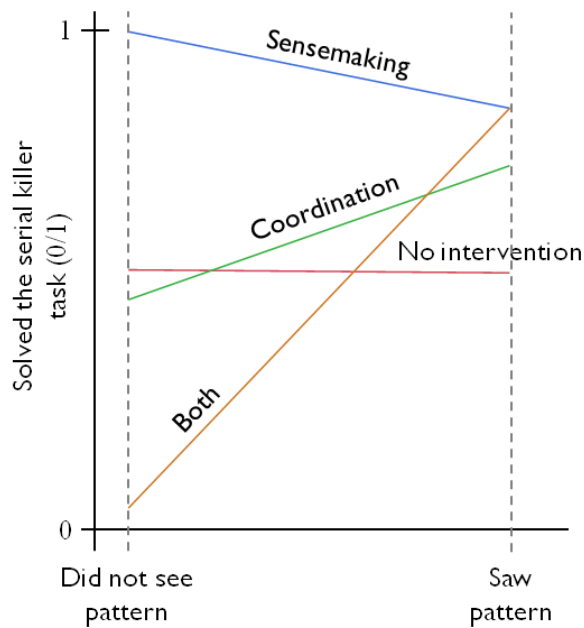
Figure 4.18. Shows the interaction profile of condition and seeing a pattern in the four cases for solving the serial killer task.
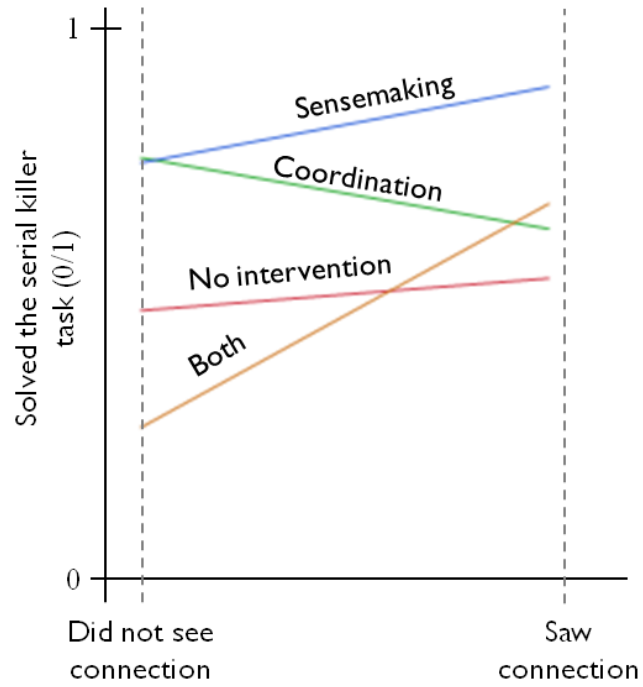
Figure 4.19. Shows the interaction profile of condition and seeing a connection between the serial killer and another character for solving the serial killer task.
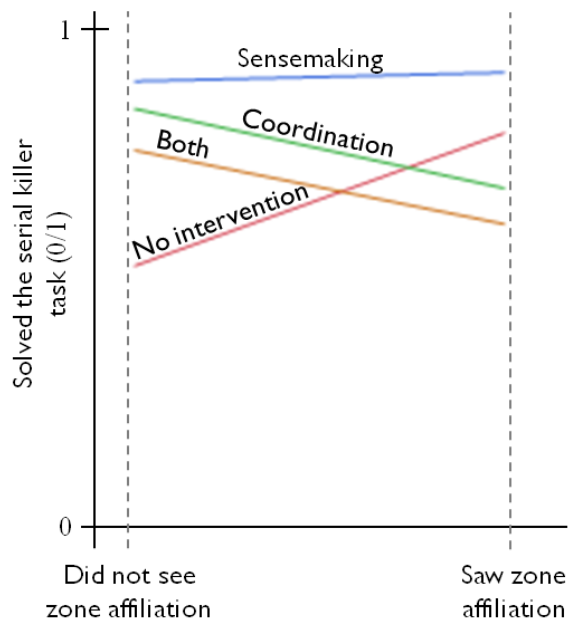
Figure 4.20. Shows the interaction profile of condition and seeing the zone affiliation anomaly for solving the serial killer task.

More importantly, seeing this pattern in the visualization was highly correlated with problem solving success ($r = .30$, $p < .01$). Next, I considered whether intervention condition influenced their likelihood of finding a pattern in the visualization. There was a significant main effect of condition (Likelihood Ratio $\chi2 = 10.89$, $p < .05$, $df = 3, 115$, Cramer's Phi $= .31$; Figure 4.15 and Figure 4.16). Fisher's pairwise comparisons revealed that individuals in coordination intervention and sensemaking intervention conditions were more likely to find this pattern in the visualization than individuals in either the no intervention or both interventions conditions ($p < .05$; coordination intervention $M = 92\%$, $SE = .05$; sensemaking intervention $M = 93\%$, $SE = .05$; no intervention $M = 69\%$, $SE = .08$; both interventions $M = 70\%$, $SE = .08$).

Logistic regression helped me understand the effect of condition in seeing the pattern, and the interaction factor for solving the serial killer task (See Figure 4.18). There was a significant interaction effect (Likelihood Ratio $\chi2 = 10.66$, $p < .05$, $df = 3, 120$). This analysis suggests that seeing the pattern and solving the case depends on what condition the individual is in. Additionally, this interaction effect moderates the main effect by condition, suggesting that this factor is very important to understanding differences between conditions in complex problem-solving tasks. Indeed, this interaction effect shows that seeing the pattern was critical to the success of pairs in the both intervention conditions, but that seeing the pattern did not significantly impact the probability of success for other conditions.

Participant's use of the first two visualization attributes (using the search capability and finding a pattern for a serial killer) relate to finding interesting patterns and similarities in the dataset. Attributes three and four focus on unusual and important links or connections. The third attribute was whether or not participants saw a suspicious connection between the serial killer and the husband of a murder victim, which linked two separate cases together (See Figure 4.13). For each participant, I coded whether or not they had focused on this unique relationship. This linkage was important because it pointed to critical information within a document that connected the serial killer to several other important clues; e.g., he rode the same bus to work as many of his victims, he had access to construction tools used in the murders, etc. I considered how condition helped focus in on this key relationship between these two characters. Condition did significantly predict whether or not participants would find this connection (Likelihood Ratio $\chi2 = 12.89$, $p < .01$, $df = 3, 115$, Cramer's Phi $= .33$; See Figure 4.15 and Figure 4.16).

Moreover, individuals in the sensemaking intervention and both interventions were more likely to find this pattern in the visualization than individuals in either the no intervention or coordination intervention conditions (Fisher's pairwise comparisons, $p < .05$; sensemaking intervention $M = 70\%$, $SE = .08$; both interventions $M = 80\%$, $SE = .07$; no intervention $M = 41\%$, $SE = .09$; coordination intervention $M = 46\%$, $SE = .1$). Logistic regression helped me understand the effect of condition on seeing the suspicious connection, and the interaction factor for solving the serial killer task (See Figure 4.19). But there was no significant interaction effect.

The fourth and final attribute of visualization use was the zone affiliation anomaly (See Figure 4.14). Each character was connected to a geographic zone based on the location to a homicide case. Zone affiliation was a characteristic that could be searched; it is also related to the third attribute, that is, related to the fact that serial killer and the husband of a murder victim were themselves linked. Only one character in all of the cases could be linked to two different geographic zones, and that was the serial killer. Participant's activity logs were searched and coded for whether or not they had found this specific zone affiliation anomaly. I considered the effect of condition on finding the zone affiliation anomaly. There was a trend towards a condition predicting whether other not participants found and searched for zone affiliation between characters (Likelihood Ratio $\chi 2 = 5.45$, $p = .1$, $df = 3, 115$, Cramer's Phi = .22). (See Figure 4.15 and Figure 4.16). Post hoc comparisons between conditions showed that participants in the sensemaking intervention condition and the both interventions condition were most likely to see this anomaly, and so were significantly more likely to see this attribute than those given no intervention (Fisher's pairwise comparisons, $p < .05$; sensemaking intervention $M = 53\%$, $SE = .09$; both interventions $M = 53\%$, $SE = .09$; no intervention $M = 28\%$, $SE = .08$; coordination intervention $M = 42\%$, $SE = .1$). The logistic regression used to understand the effect of condition on seeing the zone affiliation anomaly, and an interaction factor for solving the serial killer task, found no significant interaction effect. (See Figure 4.20).

In sum, individuals in all conditions used the visualization on average for the same amount of time. However, access to discussion prompt interventions encouraged more sophisticated use of the visualization tool. In particular, the coordination intervention and sensemaking intervention helped individuals find important patterns in the data, whereas the sensemaking intervention and both interventions helped individuals focus on an unusual linkage and an anomalous fact.

	Coordination tasks		
	Divided up documents	Focused on distraction case (percent of talk spent on distraction case)	Sharing information Average # IM lines
No intervention	20% (3/15 pairs)	16% of total IM	$M = 86$ lines ($SE = 15.4$)
Coordination intervention	80% (12/15 pairs)	10% of total IM	$M = 117$ ($SE = 15.4$)
Sensemaking intervention	20% (3/15 pairs)	18% of total IM	$M = 100$ ($SE = 15.4$)
Both Interventions	33% (5/15 pairs)	24% of total IM	$M = 113$ ($SE = 15.4$)

Table 4.4. Comparing the rates of coordination tasks between all four intervention conditions.
Note that the coordination intervention condition and the both interventions condition received the coordination discussion prompts.

	Sensemaking tasks		
	Missed serial killer pattern	Discussed facts but missed connection	Pursued wrong hypotheses
No intervention	13% (4/30 individuals)	50% (3/6 pairs)	20% (6/30 individuals)
Coordination intervention	0% (0/30 individuals)	20% (2/10 pairs)	13% (4/30 individuals)
Sensemaking intervention	7% (2/30 individuals)	0% (0/9 pairs)	0% (0/30 individuals)
Both Interventions	23% (7/30 individuals)	25% (2/8 pairs)	3% (1/30 individuals)

Table 4.5. Comparing the rates of sensemaking tasks between all four intervention conditions.
Note that the sensemaking intervention condition and the both interventions condition received the sensemaking discussion prompts.

4.4.4 PROCESS ANALYSIS

Finally, I wanted to know whether the interventions helped pairs avoid the coordination and sensemaking failures of study two. Likewise, I performed a detailed tracing of each pair's discussion across all four conditions and constructed similar process task tables. (See Table 4.4 and Table 4.5).

Coordination tasks

In the second study, I had found that none of the pairs who were given access to all of the data and a shared visualization divided the documents between partners. I was curious as to whether or not participants given the coordination prompt divided up documents with their partner. I examined every pair discussion for evidence that the pair had decided to divide up the documents between the two of them. (See Table 4.6 for an example of a conversation that shows a pair

Detective A	Should we perhaps divide up the work by looking at different zones?
Detective B	sure, I'll do zone 5
Detective A	ok
Detective B	there are two cases in zone 5 file that are very similar: middle aged caucasian [sic] white being hit with blunt object in their homes, both had recently been divorced or separated from husbands
Detective A	Yeah, two of [my] cases involved wounds with blunt objects in the head

Table 4.6. IM conversation depicting a division of labor.

This excerpt is from Pair 20 who was in the coordination intervention condition. The pair decided to divide up the documents by zone, reviewed their cases, and shared their findings.

dividing up their documents). I found that the coordination prompt did encourage a division of labor when it was given to participants alone. But the same prompt did not encourage pairs to divide up their documents if they had been given both intervention discussion prompts. Twelve out of fifteen pairs in the coordination-only intervention condition decided to divide up documents, whereas only 5 out of fifteen pairs in the both intervention condition did. Even though pairs given both interventions were encouraged to divide up the documents, they did not, and so they did not benefit by leveraging their partner.

The IM conversation excerpt in Table 4.6 provides insight into why dividing up the documents helped pairs. Pair 20, in the coordination intervention condition, decided to divide up the documents by zone, reviewed their cases, and then immediately reported their findings. This conversation helps illustrate how a division of labor could help pairs be more efficient in their task, rather than have each member having read all seven cases. Additionally, when sharing their findings with one another, this pair was able to more easily see the cases relevant to the serial killer task.

Again, there were no significant differences in how much information was shared across the four intervention conditions, but there was a difference in what pairs in each condition discussed. In particular, pairs given both discussion prompts spent the most time discussing the unrelated Raffield case, which is additional evidence that both intervention condition pairs did not effectively coordinate their efforts.

Sensemaking tasks

In the second study, I had also found evidence to show that pairs with all of the data were more likely to experience sensemaking failures than their half data counterparts, such as missing a pattern for the serial killer or pursuing a wrong hypothesis. I was interested in exploring whether or not a sensemaking intervention could help pairs avoid these types of failures, and I found a similar effect in the coordination tasks. Once again, pairs only given access to the sensemaking prompts appear to have been helped by the proposed strategies, whereas pairs given access to both sets of discussion prompts did not effectively incorporate the proposed strategies into their problem solving process.

In comparison to no intervention individuals, fewer individuals who had either the coordination intervention or the sensemaking intervention missed a pattern for a serial killer (no intervention = 4/30 individuals, coordination intervention = 0/30, sensemaking intervention = 2/30). In contrast, individuals given both interventions missed a pattern for a serial killer the most often (7/30 individuals).

I looked next for direct discussion of two critical clues that had to be connected together if participants were going to correctly identify the serial killer. I found in study two that pairs with all the data might have found and discussed these two critical clues, but that it didn't necessarily follow that they registered the connection—some still failed to solve the case. The pairs in the sensemaking intervention condition who discussed these two clues understood the connection between them and went on to successfully solve the case; not so for those in the other conditions.

Finally, participants could also pursue a wrong hypothesis by identifying an incorrect suspect in their final reports. Individuals with the sensemaking intervention most often decided upon the correct suspect as the serial killer. None of these individuals identified an incorrect suspect. These individuals also had a marked improvement over those without an intervention (6/30 individuals chose an incorrect suspect) and coordination intervention individuals (4/30). Individuals from both interventions identified the incorrect suspect only one out of thirty times. However, it is important to remember that, while they may not have focused on an incorrect suspect, these individuals were performing poorly on the other sensemaking tasks, missing the pattern for a serial killer along with failing to connect critical clues into a cohesive narrative.

The conversation excerpt from Pair 15, a sensemaking intervention pair, helps illuminate why a sensemaking intervention could help the problem solving process (See Table 4.7). One of the sensemaking interventions prompts encouraged pairs to use the visualization to find patterns in the data. In this conversation, the pair is using the visualization tool to explore different weapons associated with victims. The partners remain focused on the more prevalent or likely pattern rather than exploring the two unrelated gunshot cases.

4.5   DISCUSSION

In this third study, complex problem solving performance was increased with a visualization and process interventions. Discussion prompt interventions helped facilitate the collaborative process, reducing the burden on the pair to structure their teamwork and problem solving strategies. While all types of discussion prompt interventions lead to better performance, the types of interventions varied in effectiveness. Pairs given the sensemaking intervention performed best. In depth analysis of their problem solving process revealed that they were able to successfully complete multiple components of the task, which led to problem solving success. In particular, individuals given the sensemaking intervention found two critical elements of the case within the visualization: the pattern for a serial killer; and an unusual relationship leading to the correct suspect. Additionally, sensemaking intervention pairs had more correct solutions, they correctly identifying the pattern for a serial killer and made the correct decision on their primary suspect. Taken together, the data suggests that the sensemaking intervention helped keep participants focused on their task, helped them use the visualization tool to aid their search and analysis process, and fostered sophisticated visualization use, resulting in better decision making.

Detective A	i just looked at this diagram, only 2 people shot.
Detective B	hmmm
Detective A	more killed by blunt instrument
Detective B	ok that's more likely then lets eliminate all but blunt trauma

Table 4.7. IM conversation depicting focus on a serial killer pattern.

This excerpt is from Pair 15 who was in the sensemaking intervention condition; it shows how the pair decides to focus their attention on the more prevalent or likely pattern for a serial killer among the seven cases.

The coordination intervention improved problem-solving success, but it was not as effective as the sensemaking intervention. The coordination intervention encouraged pairs to divide up their documents in a way that streamlined their efforts; it also encouraged them to stay focused on finding the serial killer, and discouraged pairs from spending time on the distraction case. However, it did not result in more effective analysis process, as evidenced by having lower rates of correct decisions and by seeing fewer important clues within the visualization.

What can we take from this analysis of visualization use? The coordination intervention and sensemaking intervention appear to help participants find patterns in the data, while the sensemaking intervention also helped participants to spot unusual links. What becomes clear is why pairs in the sensemaking intervention condition were most successful; these pairs saw both the pattern and the unusual link in the visualization.

Being given both the coordination *and* sensemaking interventions could have given participants the best of all worlds. But no best world was seen. Pairs who were given both interventions did not divide up the documents with their partner, often missed the pattern for a serial killer, and could not connect critical clues into a cohesive narrative that pointed to a successful end. These pairs also veered off the correct path and wasted time discussing the distraction case. This suggests that while giving problem solvers some structure to their collaborative process, too much structure may be overwhelming and reduce overall task effort. Giving pairs too many options for strategies may have resulted in the pairs not choosing any option to follow thoroughly.

The findings from study three suggest that discussion prompt interventions can serve as a buffer between process failures and success. These interventions were able to help pairs overcome a variety of obstacles that confronted them. The coordination interventions helped pairs stay on task and share information with one another. The sensemaking intervention successfully improved the search and analysis process by helping pairs find the pattern for a serial killer and correctly integrating the information for a path to success.

4.5.1 LIMITATIONS

This study has similar limitations as the first two studies. Again, I focus on one analytic task, which limits generalizability. Also, partners had not previously worked together, which may

have impacted their communication. Participants also relied on IM to talk with one another. Other forms of communication, such as audio and video, may reduce barriers to open communication.

Teams who work together repeatedly often evolve and mature, developing stable patterns of work through repetition (Morgan, Salas, & Glickman, 1993; Glickman, Zimmer, Montero, Guerette, Campbell, Morgan, & Salas, 1987). The participants used in my laboratory studies did not have the opportunity to develop a stable pattern of work. They were trained on similar tasks, but this repetition is not equivalent to being a member of a long-standing team.

Note that I studied the effects of only two types of discussion prompt intervention. I did not explore other types of interventions, such as team building interventions, which could have also improved collaborative processes. I chose the two types of interventions I did because of the specific nature of the task, and because of the particular problems participants faced in previous studies.
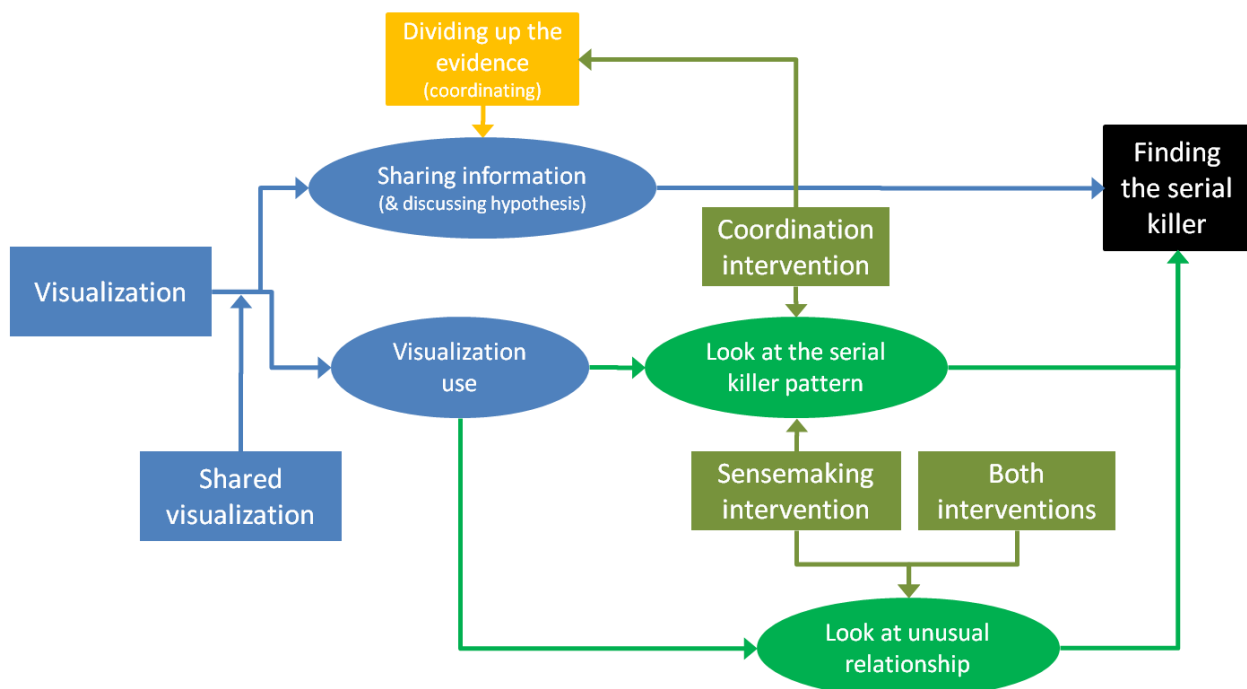
Figure 4.21. Model of main findings from all three studies.
Blue boxes represent findings from Study 1, yellow boxes represent findings from Study 2, and green boxes represent findings from Study 3.

While participants were trained on how to use the visualization tool and in investigative analysis via two training examples, all participants were still novice users who lacked the in-depth training of professional analysts. These discussion prompt interventions could be seen as additional training or instructions, or as evidence that the pre-task training was inadequate. However, the discussion prompts were created to be general rather than specific. They did not point out the clues in the case that would lead directly to the solution. Additionally, the results of study three highlight important facts regarding participant abilities to use the visualization tool and to perform investigative analysis. On average, participants in all four conditions used the visualization tool the same amount of time and actively engaged in the tool's search functionality. Also, those who did not solve the serial killer crime often solved the Raffield homicide, which required the same types of investigative analysis skills. This suggests that the discussion prompts provided different kinds support to different participants, not additional task training.

4.5.2 PUTTING IT ALL TOGETHER

To place study three's findings in context, I created a mapping of the main findings from all three studies and how they lead a path to success (See Figure 4.21). In study one (shown by the blue boxes), I found that a visualization—and in particular, a *shared* visualization—can help collaborators solve problems by increasing discussion between partners and using the visualization. In study two (shown by the yellow boxes), I found that giving collaborators half the evidence lead to a natural division of labor between the two partners and also encouraged discussion, which both lead to higher levels of success. Study three (represented by the green boxes), reveals important insights on visualization use. To be successful, collaborators had to use the visualization in order to recognize an unusual relationship and to see the serial killer's pattern. The sensemaking intervention helped partners find both these aspects, while the coordination intervention helped partners recognize a pattern and encouraged division of labor. Being given both interventions helped individuals find an unusual relationship within the visualization. Taken together, the findings from these three studies suggest that a visualization can help collaborative problem-solving, but this depends on the visualization being useful to the task, as well as on the visualization being able to inspire, or even being part of, a healthy collaborative process.

4.5.3 DESIGN IMPLICATIONS

In this last study, discussion prompt interventions were used to lend structure to the collaborative process that was found to be missing from study two. Discussion prompts are useful tactics for many types of problem-solving tasks, but ideally the visualization tool would be able to inspire such structures independently. The findings from study three suggest that having a visualization tool for sharing information and important insights is vital to success. While the visualization tool used in this study had the capacity to lead users to such important insights, only those users who were also prompted to *connect* these insights with important analysis and discussion realized the full benefits of the visualization tool. The visualization should encourage collaborators to find patterns or similarities within the visualization, search for unusual links or relationships, and think deeply about what these insights mean to their current task.

4.5.4 SUMMARY

Taken together, the findings from these three studies suggest that a network visualization can help collaborative problem-solving, but this depends on whether the visualization is useful to the task, and whether it is able to inspire or be part of a healthy collaborative process. In this last study, discussion prompt interventions were used as a way to lend structure to the collaborative process that was revealed to be missing from study two. These interventions were able to help pairs overcome the variety of obstacles that confronted them. Sensemaking prompts improved performance the most because they successfully improved the search and analysis process by helping pairs to find the pattern of a serial killer and to correctly integrate all information necessary to succeed.

5 DISCUSSION & CONTRIBUTIONS

The goal of this dissertation has been to better understand how network visualization tools affect collaborative problem solving. Collaborators are oftentimes separated, working together on problems remotely; they're also often faced with a volume of information that can greatly complicate collaboration. Visualizing this information can help collaborators sort through large quantities of data, and because of this visual analytics is a growing field. But visualizations help only when they promote effective problem-solving behaviors such as division of labor and open communication, and few behavioral studies have been performed to develop a deeper understanding of the impact of visualization tools on collaborators working to solve complex problems. The benefits of these tools are not always realized despite the fact that visual analytic tools have the technical capability to help problem solvers handle large volumes of data. Collaborations can also introduce additional barriers to success. To better clarify how visual analytic tools can improve collaborative problem solving, I conducted three studies, describing the various conditions and processes necessary to ensure the effectiveness visualization tools. In these three laboratory studies (chapters two, three, and four), I sought to understand the links between a visualization tool, and the coordination, communication, and sensemaking processes of remote collaborators.

I first asked whether or not visualizations should be shared between partners. The findings from study one revealed that sharing visualizations is not straightforward and the design of shared visualizations should consider the additional cognitive effort required when using one or multiple views of a visualization. Additionally, study one showed how visualizations could impact the communication between pairs.

Next, I tested the robustness of visualizations under conditions of information overload. I asked: "Does the visualization improve collaborative performance when pairs have equal access to data?" Whereas the visualization tool improved performance for collaborators with half the data, the same visualization tool did *not* improve performance when each collaborator had access to all the data. In other words, remote collaborators seemed to be overwhelmed. These results required

a detailed analysis of each pair's collaborative process to better understand the impact of information overload on visualization use in collaborative problem solving.

The results of study two led me to wonder if discussion prompt interventions would positively impact collaborations. Thus study three asked whether process interventions could help remote pairs overcome information overload by fostering visualization use, thereby improving team performance. I based process interventions on small group research and on my intimate knowledge of the laboratory task process. I compared the effects of different combinations of interventions on performance and effectiveness of remote pairs, measuring visualization use and performance in a more sophisticated way, by which I mean that the research moved well beyond measuring the total time spent using the tool, or thinking of success as something all or nothing way. "Success" in complex tasks is more than a binary, and to highlight this complexity I also created a schema to connect main findings from studies one, two, and three.

In the sections that follow, I discuss the contributions of this research and outline future research directions.

5.1 CONTRIBUTIONS
Fundamentally, the research in this dissertation contributes to three main areas: human-computer interaction (HCI), behavioral research, and visualization design.

5.1.1 HCI RESEARCH
This dissertation contributes the idea that collaboration design plays a critical role in the effectiveness of a visualization tool. The design of a particular collaboration refers to both external and internal components of that team. Findings from study one seemed to indicate that if a visualization tool would improve collaborative performance, it must also encourage positive collaborative behaviors, such as information sharing. When a remote pair has a jointly manipulable, shared-visualization, and shares the information they have, they will also have a higher rate of complex problem-solving success. Surprisingly, not all shared visualizations were equal. A shared view of a collaborator's visualization tool did not have the same beneficial effects as a fully shared visualization. The results of study one showed that the particular design of a visualization tool could significantly alter the effectiveness on team processes and performance.

The findings of study two were perhaps more intriguing. Study two revealed how external factors, such as information access, influence visualization use, and team effectiveness, affected collaborative performance. More specifically, information overload lead to an absence of the coordination mechanisms necessary for incorporating a visualization tool successfully into the collaborative, analytic process. Individuals were overwhelmed with information. They lost sight of important assets available to them, including access to a partner and the visualization tool itself.

Study three focused on rebuilding coordination and sensemaking procedures in response to information overload. Study three shows that if coordination and sensemaking procedures are in place, visualization use can improve performance. Scaffolded sensemaking leads remote pairs to discover critical visual insights, and increases the likelihood of their integrating these insights into a cohesive narrative. Why was the sensemaking intervention more successful than the coordination intervention? The nature of the problem-solving task itself required sensemaking for success; coordination, on the other hand, *could* make a team more efficient, but it did not lead necessarily to higher levels of understanding and the construction of meaning between the pair.

Somewhat counter intuitively, the results of study three show that the use of coordination and sensemaking prompts does not achieve additional benefits over using one of the two by themselves. Study three suggests that providing too much structure within collaboration reduces the potential benefits of structure. Collaborations incur costs, and these costs must be balanced with the needs of the task. Further research is necessary to better understand the generalizability of these findings, but this thesis offers insight into the complex relationship between collaborative visual analytic tools and the critical mechanisms supporting collaboration between a remote pair.

5.1.2 BEHAVIORAL RESEARCH

This dissertation provides contributions to three areas in behavioral research: information behavior, information sharing between dyads (pairs), and process intervention.

Information Seeking and Sensemaking

This research focuses squarely on collaborative problem solving and the process of sensemaking. The research of Pirolli & Card (2005) and also Bates (1989) influenced how I organized my

analysis. These works identified critical components of the information search and sensemaking process. Both of these models focus on the information behavior at an individual level. Throughout the dissertation, I have explored the information search and sensemaking process as a fundamentally collaborative endeavor. Findings from studies one, two, and three suggest how collaboration processes and visualization use influence information seeking and sensemaking behaviors.

The berrypicking approach to information seeking emphasizes the dynamic nature of information search. Information seekers, it was already understood, may change the direction of their search based on new information (See Figure 5.1 from Bates, 1989). But my research emphasizes that this information comes in a variety of forms, including visual forms. Hence the original berrypicking schematic has been enriched by visualization icons that illustrate the impact of visualization tools on the search process. Users of visual information may also change their query based on visual cues. Analysis from study three pointed out critical visual elements that motivated successful search and sensemaking processes. Additionally, Bate's berrypicking delineates a four-level hierarchy for search activities: move, tactic, stratagem, and strategy.
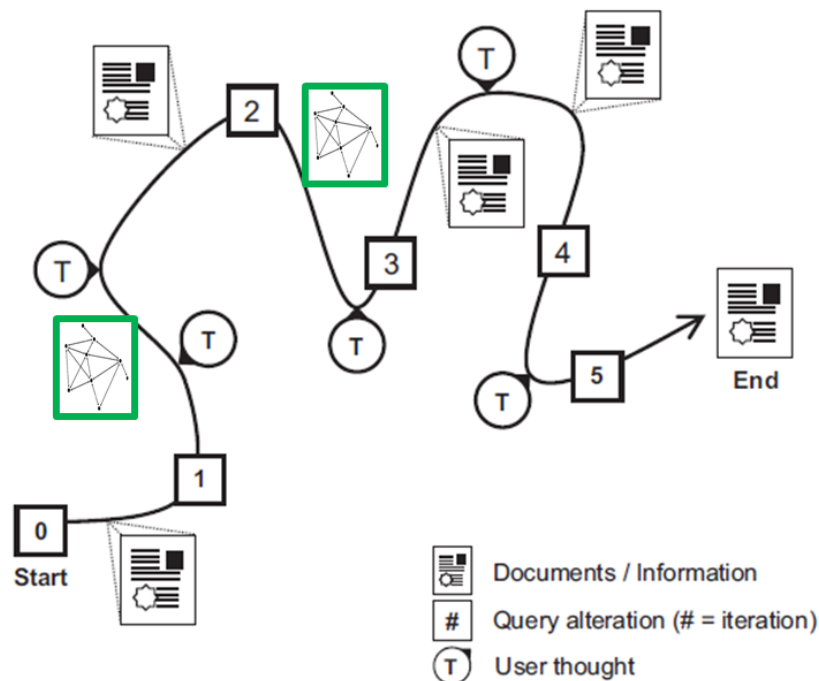
Figure 5.1. Berrypicking enhanced by access to a visualization tool.

Moves are single actions and tactics involve multiple moves. Strategems combine both actions and tactics while strategies involve all three sublevels. Findings from my dissertation also suggest that a visualization tool affects all four of these levels. A visualization of available information available can point toward the next document to search, the next several documents to search, patterns to watch out for, and offer multiple scenarios to fruitfully investigate.

My research also implies additional sensemaking processes in collaboration. Figure 1.5 illustrates the sensemaking loop (Pirolli & Card 2005). At the structural level, collaborative tasks include coordinating and reaching consensus. Coordination as a process is relevant across the entire duration of a task. Study two highlights the importance of coordination within a team, specifically in their division of labor.

Information sharing is another major component of collaboration. Information sharing also occurs at multiple stages of the sensemaking process, and it can be divided into distinct types: evidence sharing; discussing schema; and discussing hypothesis. Findings from study one and two show that information sharing across all three types is critical to problem-solving success.

Additionally, all three studies stressed that consensus in decision making is non-trivial. Reaching consensus involves comparison among alternatives and potentially justifying reasoning—it is more than just elucidating potential hypotheses. While agreement on final solutions was measured, this work did not focus on the process of reaching consensus. Further research is needed to better understand what effort is required to obtain this consensus (Mohammed, 2001).

Finally, this dissertation points to a critical failure in collaborative problem solving: when problem solvers lose focus on the problem itself. In all three studies, many pairs wasted their time and attention on a "red herring," trying to solve an unrelated homicide case. These problem solvers gone astray were not interrupted, nor did they give up on their task of finding the serial killer. Indeed, they found important clues and came to a solution with respect to the distraction case. However, they had lost sight of their ultimate goal on their way to trying to achieve it. Information seeking and sensemaking in this particular type of task is difficult for individuals because it is not always clear what the correct solution is, or even the format that the solution should take. What processes or tools can prevent problem solvers from losing focus? Further

research is needed to better understand how pairs can become better information seekers, so that they can more readily distinguish irrelevant information from relevant.

Information sharing

This dissertation also has implications for information sharing patterns among dyads, or pairs. Stasser and Titus found that when teams are given disparate information, they tend to focus on the information that individuals have in common; this can lead to confirmation bias (1985). The findings from study one suggest a possible solution to the tendency for sharing confirmatory information. The shared visualization offered a view of partners without complete access to information. Knowing your partner has access to unique information may be enough to stimulate sharing of unshared or disparate information.

Findings from study two point to possible benefits of dividing up information, despite the possibility of negative outcomes such as focusing on common information. Dividing up information between pairs facilitated the coordination processes, which were critical to helping teams be effective. Findings from study two also revealed that the potential for confirmation bias exists under information overload. Further research on information access is needed to better understand how to avoid information overload, and on how to avoid this tendency toward confirmation bias.

Small group interventions

This dissertation also provides theoretical contributions to intervention literature within small group research. I developed a schema to organize existing intervention literature. This schema structures possible target points that can influence small groups, and also points to opportunities for future research. The findings from study three support the idea that content-based (in-process, task-based) interventions have the ability to improve team performance over interpersonal interventions (Lipshitz & Sherwood, 1978). Interpersonal interventions are not completely ineffectual; they improve team effectiveness, such as coordination, but this does not necessarily result in large improvements in performance.

5.1.3 DESIGN IMPLICATIONS

This research has design implications as well. This dissertation explicitly lays out the many difficulties remote collaborators confront when conducting a complex task, particularly when trying to share a visual analytics tool. Additionally, my findings highlighted areas where problem

solvers needed help, providing opportunities for new tool and feature design. In all three studies, a visualization tool had the ability to improve collaborative problem solving but this depended on the pair adopting successful problem solving strategies. In study three, I found that discussion prompt interventions were a useful tactic for improving collaborative problem solving. Ideally, the visualization tool would be able to inspire process structures such as those outlined, as well as encourage positive behaviors at the individual level.

There exists a rich history of research on information visualization design and analytics (Card, Mackinlay, & Shneiderman, 1999). Much of this research has focused on the development of software and network architectures and also synchronization protocols that are necessary for supporting collaboration (Heer & Agrawala, 2006; Anupam, Bajaj, Schikore, & Schikore, 1994). However, my research has implications on behavioral questions whose answers support collaborative data analysis activities and associated tasks (Viegas & Wattenberg, 2006; Wattenberg & Kriss, 2006; Heer & Agrawala, 2008). One such example of this sort of research is Heer and Agrawala's set of design considerations for collaborative visual analytics (2008). While Heer and Agrawala's focus is on asynchronous collaborations, their considerations on focused coordination and communication are highly relevant to this dissertation (sections two and three of their paper). Likewise, my dissertation calls upon developers to focus on features that support coordination and sensemaking tasks.

Designing for Coordination

Many developers have focused on improving coordination between collaborators, such as adding features that improve awareness of one's partner's activities. My studies suggest that a shared visualization *does* provide a medium through which users can mark items that need to be discussed (Balakrishnan et al., 2008; Hill, Hollan, Wroblewski, & McCandless, 1992). A shared visualization also fosters a shared mental model that both partners can base their mutual understanding of the information on, simplifying the information sharing process (Balakrishnan et al., 2008; Brennan et al., 2006; Gergle, Kraut, & Fussell, 2004).

Additionally, study three showed how encouraging pairs to divide up their information and tasks created a division of labor and helped them improve their overall performance. Division of labor improves a team's efficiency by eliminating duplication of work. This is critical in situations when there is limited time, resources, or both. Remote pairs could organize themselves in a way

that access to data was automatically dispersed among their team (i.e., in study one partners were given access to half of the evidentiary documents). However, real-world analysts do not want an artificial division of information, preferring seamless access to information, as well as access to their partner's work (Badalamente & Greitzer, 2005). One main way, then, to implicitly facilitate division of labor is to create an awareness of partner efforts. This allows for pairs to plan around their partner's activities and allocate their efforts more efficiently (Carroll, Rosson, Convertino, & Ganoe, 2005; Dourish & Bellotti, 1992). In addition, designs for increased partner activity awareness may reduce the burden of sharing information. There are many ways to facilitate awareness of partner efforts, including visualizations and social network or group sites (Gutwin, 2002; Balakrishnan et al., 2010). Social network sites aimed at this sort of enterprise, such Chatter (https://www.chatter.com/) and Yammer (https://www.yammer.com/), seek to provide such awareness of collaborator activities in the common framework of a social network paradigm. Transparent, open, and engaging applications may encourage a natural coordination process among collaborators. Coordination improvements *do* result in higher efficiency within a pair, which is valuable for overall team cohesion and affect. However, previous research on small groups and findings from study three suggest that improving coordination within a team does not necessarily translate to the best overall task performance.

Designing for Sensemaking

Users need to connect insights with analysis to understand the importance of an insight. The visualization tool used in all three studies has the capability to lead users to these important insights. In a study of intelligence analysis tool development, Badalamente and Greitzer (2005) found that analysts want their tools to include a strategy template that helps them work. Taken together, this research suggests that tool design should incorporate strategy structures to help users through the collaborative sensemaking process, better linking insight and analysis.

Heer and Agrawala (2008) note that future opportunities for research agendas should consider automated techniques. And while I agree that a simple algorithm could automatically present patterns or unusual links to users, and concede that more sophisticated tools could be developed, my research purposefully decided on a different direction. I maintain, here and throughout, that in certain situations algorithms and tools will be unable to detect novel and unforeseen patterns in information where human experience and intuition might.

Sensemaking involves both information foraging and information schematization (Pirolli & Card, 2005; Heer & Agrawala, 2008). One major roadblock making information search difficult is the large volume of data available to users (Patterson, Woods, Tinapple, Roth, Finley, & Kuperman, 2001; Patterson, Woods, Tinapple, & Roth, 2001; Tinapple & Woods, 2003). Results from study two also point to this impasse (Balakrishnan et al., 2010). Visualizations do allow users to cut through large amounts of data in order to find critical pieces of information. However, studies two and three found that, even with access to the visualization with search functionalities, individuals often did not see important clues or facts within the visualization tool. During information schematization, problem solvers are trying to understand how *all* their clues fit together by integrating the various pieces into a coherent narrative. Participants in all three of the laboratory studies were not successful in connecting *all* the disparate "dots" available to them. Some failed to integrate the information necessary to find a pattern for a serial killer, while others were unable to connect the correct perpetrator to the correct crimes. Visualizations can make integration more effective by allowing users to visually connect various dots, and help users develop theories more rapidly (Wright et al., 2006; Stasko, Gorg, Liu, & Singhal, 2007). My findings suggest that the design of visual analytic tools should facilitate the search for facts and patterns within the dataset to help users best understand the data and how it hangs together.

An example of a design element that may support these tasks is "tagging." Tagging systems, such as Del.icio,us and Flickr, enable users to annotate specific items, such as photos, web pages, or news articles, with a user-defined keyword or set of keywords (Golder & Huberman, 2006). Tagging features help individual users search for similar information or objects, and also to re-find material previously tagged (Mathes, 2004; Golder & Huberman, 2006; Hammond, Hannay, Lund, & Scott, 2005). Tagging systems also have potential to benefit collaborations (Marlow, Naaman, Boyd, & Davis, 2006). By tagging an item with a specific keyword, the user implicitly shares and recommends an item to his collaborators based on some fact (Golder & Huberman, 2006; Willett, Heer, Hellerstein, & Agrawala, 2011). Searching for common tags or visually tagging an item as important could make patterns or unusual links more salient. Visual cues within information spaces help users better navigate to important information (Willett, Heer, & Agrawala, 2007). During pretests and think-aloud protocols, designed to help users brainstorm visualization features, users often suggested increased interactivity within the visualization tool. Thus I developed a prototype visualization tool that included a tagging feature as a proof of

concept, exploring whether or not users would find tagging useful. And tagging was found to be a possible feature, one that could support finding facts and patterns. (See Appendix A for a full description of the tagging tool).

A second example of how a tool can enable sensemaking is the "Apolo tool" (Chau et al., 2011). Rather than present users with all the available data at once, Apolo guides its users through the data incrementally and interactively; it allows them to explore a large network of data and to make sense of it gradually. If a user finds a piece of information that he deems important, Apolo also helps users find relevant information by specifying that important piece as an example. In short, Apolo uses machine learning to deduce what other pieces of information may be of interest to that user. Such tools can help users discover information gradually rather than overwhelming them all at once with the entire dataset. In all three studies, participants were given access to a visualization with their entire dataset already fully in view. Structuring the data exploration and search process into manageable pieces with a tool like Apolo may have helped participants to see patterns emerge, rather than having users search directly for meaningful patterns.

5.2 LIMITATIONS AND FUTURE DIRECTIONS

This dissertation has, of course, limitations and opportunities for future research. There is a greater need to conduct in-situ studies in the field—to move beyond the laboratory—to work with real analysts in their natural work environment. A critical limitation of this work is the external validity of the laboratory studies. The serial killer mystery paradigm used in all three studies was constructed to simulate tasks performed by intelligence analysts. While exploring these research questions with an artificial task, and from within the controlled environment of laboratory, studies did result in valuable findings. But it remains unclear how such research questions would play out in a real work environment (i.e. with a real serial killer). In the field, analysts are likely to have higher volumes of information, access to a greater variety of datasets and tools, be trained for specific analytical tasks, have preexisting relationships with their collaborators, and know how to handle their responsibilities within their work environment. A good question here would be: How do organizational factors, such as external incentives and competition, affect collaborations? In certain work environments, especially in high-risk work such as in the intelligence and finance industries, there may be constraints in place to reduce false positives, changing the way analysts conduct their sensemaking process. How do cultural

constraints such as these impact problem solving strategies? How do analysts manage to integrate information from a variety of sources? Do analysts focus on a particular line of questioning or several simultaneously while reviewing information?

Additionally, analysts go through extensive training on the way to becoming analysts. What are the most important components of these training programs, where what's important is carried over into their day-to-day experiences as an analyst? Even after training, it can take years for analysts to become experts in their field. How do analysts become experts? What is it that an expert analyst can do that other analysts cannot? A better understanding of how experts go about their sensemaking and coordination processes may uncover additional components critical to performance success.

Beyond issues of external validity, there are a number of interesting avenues of research left to pursue, which build upon this dissertation's work. Again, the focus of this dissertation is on synchronous remote collaborations between pairs. As distributed teams, in pairs or otherwise, are becoming more ubiquitous, many teams are geographically distributed over different time zones and rely more on asynchronous collaboration (Tang, Zhao, Cao, & Inkpen, 2011). While this does not alter the nature of the problem solving task (a problem still must be solved), it does affect communication and coordination processes within teams. Unlike synchronous collaborations, where work is done in parallel, work in asynchronous collaborations is done sequentially, where partners hand off their work to their collaborator. In some sense, this reduces the need for coordination on specific in-process tasks, but it also requires more effort to maintain a common mental model of the problem and its possible solution(s). Recent work done by colleague Ruogu Kang explores the use of annotations and interim work progress (Kang & Kiesler, under review). A design, similar to the tag-enabled visualization prototype, may serve as a visual equivalent to sharing notes with partners. Further research is needed to understand how this sort of sharing impacts performance, and on how teams can leverage sequential, hand-offs collaboration to their benefit.

In all three of this dissertation's studies, collaborators relied on instant messaging to communicate with their partner. In the real world, collaborators have a host of available means for communication, such as email, team rooms, wikis, audio chat, even screen-sharing or video chat. In a recent field study with Tara Matthews and Tom Moran from IBM Research –

Almaden, I found that business analysts rely heavily on Microsoft Powerpoint presentation decks to share information with their partners and clients (Balakrishnan, Matthews, & Moran, 2010). How do users choose what mode of communication to use and why? Does access to multiple forms of communication improve collaborative problem solving, either in its coordination processes, its sensemaking processes, or both?

In study three, I explored two different types of interventions. However, there are many other types of interpersonal interventions, as well as many alternative in-process task-oriented interventions. The "devil's advocate technique" could be useful when pairs discuss different hypotheses while brainstorming because it might help pairs think more creatively about the facts they have collected, leading them to seek after more salient patterns or linkages. Moreover, this dissertation did not explore interventions that were related to climate, team design, leadership or training. If, for example, a pair is given a team leader, is their process made more efficient? Are they more motivated to succeed?

In all three studies, I used the same visual analytic tool, NetDraw. I chose NetDraw because it was easy to learn and well-suited for this particular complex task. However, more sophisticated tools do exist, and they may simplify certain aspects of the task process. For example, many tools could easily identify patterns and outliers within the dataset. How would an increased capability to identify patterns and outliers affect performance? And how would the findings of this dissertation apply to other visualization tools, or future tools? A key contribution of this work is its elucidation of the work processes intrinsic to a collaborative problem-solving task, and how to measure the effectiveness of a visualization tool on work processes. Future work is needed to validate these visualization usage measures across different tools, different tasks.

Perhaps the most exciting research prospect lies in the development and subsequent testing of new visualization tools and features. As Heer and Agrawala (2008) point out, there are a number of design considerations that present opportunities for collaborative visual analytics; in particular, being able to share not just active real-time use, but also specific states of analysis. It remains unclear how (or if) new sharing mechanisms will improve collaborative problem solving. Thompson and colleagues' warn that "visualization tools have merely added to the plethora of existing resources that engineers need to search and sift through each day" (Thompson et al., 2006), and their warning ought not be dismissed. Increasingly more complex

tools do not necessarily mean an improved collaborative analysis. Sophisticated tools require more expertise and training in order to take full advantage of new features and techniques.

Last but certainly not least, a common theme that runs throughout this dissertation is that more information is not necessarily best. In study one, for example, when pairs in the shared view-only visualization were given access to two diagrams rather than one, their performance suffered. In study two, when pairs were given access to all of the data, their performance suffered. In study three, when pairs were given access to both types of intervention prompts, their performance again suffered. In an "age of information," the amount of data available to us seems to expand at a rate and by a volume that far exceeds our ability to conceive of its speed or size. The question for us then becomes how can we give problem solvers more information, in better ways, without allowing their performance to suffer.

5.3 CONCLUSION

This dissertation asks how visual analytic tools, specifically network visualization tools, impact collaborative problem solving in remote pairs. Using the "detective mystery" as an experimental paradigm, I conducted three laboratory studies to understand under what conditions visualizations improve coordination and sensemaking processes. This dissertation supports the idea that visualizations can improve collaborative performance, but it also stresses that this improvement depends on the tool's ability to support beneficial collaborative processes, such as information sharing, division of labor, and encouraged integration of visualization use with the sensemaking process. Future research ought to be done in order to better understand additional factors that affect collaborative problem solving. There are great opportunities for visualization tools to help analysts in real-world situations. These three studies have taken a strong step toward understanding what the key components of the collaborative problem-solving process are, such that visualization tools can best support those components. The design of new visual analytic tools can greatly benefit from this research because it explores the behavioral impact of visualization tools, uncovering in rich detail how remote partners can interact with each other through a collaborative, visual system.

6 REFERENCES

Adams, J. R. (2011). Leveraging the Cloud for the Intelligence Community. GovConExec
    Magazine, Spring 2011. Washington, D.C.: Executive Mosaic LLC. Retrieved from
    http://www.govconexec.com/2011/03/02/leveraging-the-cloud-for-the-intelligence-
    community/

analysis (2011). In *Merriam-Webster.com*. Retrieved from http://www.merriam-
    webster.com/dictionary/analysis

Anderson, N., & West, M. A. (1998). Measuring climate for work group innovation:
    Development and validation of the team climate inventory. *Journal of Organizational
    Behavior*, *19*, 235–258.

Andrews, K., & Heidegger, H. (1998). Information Slices: Visualization and Exploring Large
    Hierarchies using Cascading, Semi-Circular Discs. *Proceedings of Information
    Visualization 1998*. IEEE press.

Anupam, V., Bajaj, C., Schikore, D., & Schikore, M. (1994). Distributed and Collaborative
    Visualization. *Computer*, *27*(7), 37–43.

Argote, L., Gruenfeld, D., & Naquin, C. (2001). Group learning in organizations. M.E. Turner
    (Ed.), *Groups at work*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Badalamente, R. V., & Greitzer, F. L. (2005). Top ten needs for intelligence analysis tool
    development. *Proceedings of the 2005 International Conference on Intelligence Analysis*.

Balakrishnan, A. D., Fussell, S., & Kiesler, S. (2008). Do visualizations improve synchronous
    remote collaboration? *Proceedings of the ACM Conference on Human Factors in
    Computing Systems CHI 2008*, ACM Press.

Balakrishnan, A. D., Fussell, S., & Kiesler, S., Kittur, A. (2010). Pitfalls of information access
    with visualizations in remote collaborative analysis. *Proceedings of the ACM Conference
    on Computer-Supported Cooperative Work CSCW 2010*, ACM Press.

Balakrishnan, A. D.**,** Matthews, T., & Moran, T. P. (2010). Fitting an Activity-Centric System
    into an Ecology of Workplace Tools. *Proceedings of CHI 2010,* ACM Press.

Baron, R. A. (2006). Opportunity recognition as pattern recognition: How entrepreneurs "connect the dots" to identify new business opportunities. *Academy of Management Perspectives*, *20*(1), 104–119.

Bates, M. J. (1989). The design of browsing and berrypicking techniques for the online search interface. *Online Information Review*, *13*(5), 407–424.

Beaney, M. (2011). Analysis. *The Stanford Encyclopedia of Philosophy (Summer 2011)*. Edward N. Zalta, (Ed.), Retrieved from http://plato.stanford.edu/archives/sum2011/entries/analysis/

Beer, M. (1980). Organization change and development: A systems view. Glenview, IL: Scott, Foresman & Co.

Billman, D., Convertino, G., Pirolli, P., Massar, J. P. and Shrager, J. (2005). Collaborative intelligence analysis with CACHE: Bias reduction and information coverage. PARC UIR Tech Report. Palo Alto Research Center, CA. Retrieved from http://cscl.ist.psu.edu/public/users/ gconvert/mypapers/hcic2006_BillmanEtAl.pdf

Billman, D., Convertino, G., Shrager, J., Pirolli, P., & Massar, J. (2006). Collaborative intelligence analysis with CACHE and its effects on information gathering and cognitive bias. *Human Computer Interaction Consortium Workshop*.

Blockeel, H., & Moyle, S. (2002). Collaborative data mining needs centralised model evaluation. Lavrac, N., Motoda, H. and Fawcett, T., (Eds.), *Proceedings of the ICML-2002 Workshop on Data Mining Lessons Learned*, 21–28.

Boba, R. (2005). *Crime Analysis and Crime Mapping* .Thousand Oaks, California: Sage Publications, 6–7.

Bradley, J., White, B. J., & Mennecke, B. E. (2003). Teams and Tasks: A Temporal Framework for the Effects of Interpersonal Interventions on Team Performance. *Small Grup Research, 34*, 353.

Bradner, E., & Mark, G. (2002). Why distance matters: effects on cooperation, persuasion and deception. *Proceedings of CSCW 2002*, ACM Press, 235.

Brennan, S. E., Mueller, K., Zelinsky, G., Ramakrishnan, I.V., Warren, D. S., & Kaufman, A. (2006). Toward a Multi-Analyst, Collaborative Framework for Visual Analytics. *IEEE Symposium of Visual Analytics Science and Technology*.

Brewer, I., MacEachren, A. M., Abdo, H., Gundrum, J., & Otto, G. (2000). Collaborative geographic visualization: Enabling shared understanding of environmental processes. *Proceedings of InfoVis 2000: IEEE Symposium on Information Visualization*, 137–141.

Börner, K. (2001). iScape: A collaborative memory palace for digital library search results. *Proceedings of CHI 2001*, ACM Press, 1160–1164.

Cannon-Bowers, J. A., & Salas, E. (1998). Making decisions under stress: Implications for individual and team training. Washington, DC: American Psychological Association.

Cannon-Bowers, J. A., Salas, E. & Converse, S. (1983). Shared Mental Models in Expert Team Decision Making. In N. John Castellan, (Ed.), *Current Issues in Individual and Group Decision Making*. Hillsdale, NY: Lawrence Erlbaum Associates.

Card, S. K., Mackinlay, J. D., & Shneiderman, B. (1999). *Readings in information visualization: Using vision to think*. New York, NY: Morgan-Kaufmann.

Carroll, J., Rosson, M. B., Convertino, G.,& Ganoe, C. H. (2005). Awareness and teamwork in computer-supported collaborations. *Interacting with Computers*, *18*(1), 21–46.

Cartwright, D. & Zander, A. (1960). *Group Dynamics: Research and Theory*. New York, NY: Harper & Row.

Clark, N. K., & Stephenson, G. M. (1989). Group remembering. In P. B. Paulus (Ed.), *Psychology of group influence* (2nd ed.). Hillsdale, NJ: Erlbaum, 357–391.

Chau, D. H., Kittur, A., Hong, J. I., Faloutsos, C. (2011). Apolo: making sense of large network data by combining rich user interaction and machine learning. *Proceedings of CHI 2011*. ACM Press.

Cheng, P., Lowe, R., & Scaife, M. (2001). Cognitive Science Approaches To Understanding Diagrammatic Representations. *Artificial Intelligence Review*, *15*(1), 79–94.

Clancey, W. J. (1994). Situated cognition: How representations are created and given meaning. In Lewis, R. and Mendelsohn, P. (Eds.), *Lessons from Learning*. Amsterdam: North Holland, 357–391.

Clauser, J.K. & Weir, S.M. (1975). *Intelligence Research Methodology, An Introduction to Techniques and Procedures for Conducting Research in Defense Intelligence*. Washington, DC: Defense Intelligence School.

Collins, D. B., & Holton III, E. F. (2004). The effectiveness of managerial leadership development programs: A meta-analysis of studies from 1982 to 2001. *Human resource development quarterly*, *15*(2), 217–248.

Combs, B., & Slovic, P. (1979). Newspaper coverage of causes of death. *Journalism Quarterly*, *56*, 837–843, 849.

Convertino, G., Billman, D., Pirolli, P., Massar, J. P., & Shrager, J. (2008). The CACHE study: Group effects in computer-supported collaborative analysis. *Proceedings of CSCW 2008*, ACM Press, 357–391.

Cook, M. B., & Smallman, H. S. (2007). Visual evidence landscapes: Reducing bias in collaborative intelligence analysis. *Proceedings of Human Factors and Ergonomics*, *51*, 303–307.

Cosier, R. A. (1978). The effects of three potential aids for making strategic decisions on predictive accuracy. *Organizational Behavior and Human Performance*, *22*, 295–306.

Cummings, J. N., & Kiesler, S. (2005). Collaborative research across disciplinary and organizational boundaries. *Social Studies of Science*, *35*(5), 703.

Cummings, J. N., & Kiesler, S. (2007). Coordination costs and project outcomes in multi-university collaborations. *Research Policy*, *36*(10), 1620–1634.

Davis, J. (1992). *The Challenge of Opportunity Analysis*. Intelligence Monograph. Washington, DC: Center for the Study of Intelligence. CSI 92-003U.

Davis, J. H., Laughlin, P. R., & Komorita, S.S. (1976). The social psychology of small groups: Cooperative and mixed-motive interaction. *Annual Review of Psychology*, *27*, 501–541.

Dearth, D. H. (1995). National Intelligence: Profession and Process. In Douglas H. Dearth and R. Thomas Goodden (Eds.), *Strategic Intelligence: Theory and Application* (2nd ed.). Washington, DC: Joint Military Intelligence Training Center.

Delbecq, A. L., Van de Ven, A. H., & Gustafson, D. H. (1975). *Group techniques for program planning: A guide to nominal group and Delphi processes*. Middleton, WI: Green Briar Press.

Deutsch, M. (1960). The Effects of Cooperation and Competition Upon Group Process. In Dorwin Cartwright and Alvin Zander (Eds.). *Group Dynamics: Research and Theory*. New York: Harper & Row.

Devine, D. J. (1999). Effects of cognitive ability, task knowledge, information sharing, and conflict on group decision-making effectiveness. *Small Group Research*, *30*(5), 608–634.

DiMicco, J. M., Pandolfo, A., & Bender, W. (2004). Influencing group participation with a shared display. Proceedings of the *ACM Conference on Computer-Supported Cooperative Work CSCW 2004*, 614–623.

DiMicco, J. M., & Bender, W. (2007). Group reactions to visual feedback tools. *In Proceedings of the 2nd international conference on Persuasive technology* (pp. 132–143). Palo Alto, CA, USA: Springer-Verlag.

Dourish, P., Belotti, V. (1992). Awareness and coordination in shared workspaces. *Proceedings of CSCW 1992*, ACM Press 107–114.

Dyer, J. C. (1984). Team research and team training: State-of-the-art review. In F.A. Muckler (Ed.), *Human factors review*. Santa Monica, CA: Human Factors Society, 285–323.

Eccles, R., Kapler, T., Harper, R., & Wright, W. (2007). Stories in GeoTime. *Proceedings of IEEE Visual Analysis Science and Technology VAST'07,* IEEE Press, 19–26.

Edelson, D., Pea, R., Gomez, L. (1996). Constructive in the collaboratory. In B.G. Wilson (Ed.). *Constructivist learning environments: Case studies in instructional design*. Englewood Cliffs, NJ: Educational Technology Publications.

Eden, Colin (1990). The unfolding nature of group decision support - Two dimensions of skill. In C. Eden and J. Radford (Eds.). *Tackling strategic problems: The role of group decision support.* UK, London: Sage Press, 48–52.

Edmunds, A. & Morris, A. (2000). The problem of information overload in business organizations: a review of the literature. *International Journal of Information Managemen*t, *20*(1), 17–28.

Ellis, S. E., & Groth, D. P. (2004). A collaborative annotation system for data visualization. *Proceedings of the Working conference on Advanced Visual Interfaces*, 414.

Eppler, M. J. & Mengis, J. (2004). The concept of information overload: a review of literature from Organization Science, Accounting, Marketing, MIS, and Related Disciplines. *The Information Society*, *20*(5), 325–344.

Fandt, P., Richardson, W. & Conner, H. (1990). The Impact of Goal Setting on Team Simulation Experience. *Simulation and Gaming*, 21(4), 411–22.

Fiore, S., Salas, E., Cuevas, H., & Bowers, C. (2003). Distributed coordination space: toward a theory of distributed team process and performance. *Theoretical Issues in Ergonomics Science*, *4*(3), 340–364.

Foster, J. (2006). Collaborative information seeking and retrieval. *Annual Review of Information Science and Technology, 8*, 329–356.

Fraidin, S. N. (2004). When is one head better than two? Interdependent information in group decision making. *Organizational Behavior and Human Decision Processes*, *93*(2), 102–113.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, *19*(4), 25–42.

Galinsky, A. D., & Kray, L. J. (2004). From thinking about what might have been to sharing what we know: The effects of counterfactual mind-sets on information sharing in groups. *Journal of Experimental Social Psychology*, *40*(5), 606–618.

Gergle, D., Kraut, R. E., & Fussell, S. R. (2004). Language efficiency and visual technology: Minimizing collaborative effort with visual information. *Journal of Language & Social Psychology*, *23*(4), 491–517.

Glickman, A. S., Zimmer, S., Montero, R. C., Guerette, P. J., Campbell, W. J., Morgan, B. B., & Salas, E. (1987). *The evolution of teamwork skills: An empirical assessment with implications for training*. Orlando, FL: Naval Training Systems Center.

Golder, S. A., & Huberman, B. A. (2006). The Structure of Collaborative Tagging Systems. *Journal of Information Science, 32*(2).

Goodall, J. R., Lutters, W. G., & Komlodi, A. (2004). I know my network: collaboration and expertise in intrusion detection. *Proceedings of CSCW 2004*, ACM Press, 342–345.

Gottlieb, S., Arenberg, S., & Singh, R. (1994). *Crime Analysis, From First Report to Final Analysis*. Santa Barbara, California: Alpha Publishing.

Gotz, D., Zhou, M. X., & Aggarwal, V. (2006). Interactive visual synthesis of analytic knowledge. *Proceedings of 2006 IEEE Symposium on Visual Analytics Science and Technology,* 51–58.

Greif, I. & Sarin, S. (1987). Data sharing in group work. *ACM Transactions on InfoSys*, *5*, 187–211.

Gutwin, C. (2002). Traces: Visualizing the immediate past to support group interaction. *Graphics Interface*, 43–50.

Hackman, J.R. (1976). Group influences on individuals. In M.D. Dunnette (Ed.), *Handbook of Industrial and Organizational Psychology*. Chicago, IL: Rand-McNally.

Hackman, J. R., Brousseau, K. R., & Weiss, J. A. (1976). The interaction of task design and group performance strategies in determining group effectiveness. *Organizational Behavior and Human Performance*, *16*(2), 350–365.

Hammond, T., Hannay, T., Lund, B., & Scott, J. (2005). Social bookmarking tools (I). *D-Lib Magazine*, *11*(4), 1073–82.

Harkins, S. G., & Petty, R. E. (1982). Effects of task difficulty and task uniqueness on social loafing. *Journal of Personality and Social Psychology*, *43*(12), 14–29.

Hansen, P. & Jarvelin, K. (2005). Collaborative information retrieval in an information-intensive domain. *Information Processing and Management*, *41*, 1101–1119.

Harper, R. & Sellen, A. (1995). Collaborative Tools and the Practicalities of Professional Work at the International Monetary Fund. *Proceedings of CHI 1995*, ACM Press, 122–129.

Hart, S. G., & Staveland, L. E. (1988). Development of a multi-dimensional workload rating scale. In P. A. Hancock & N. Meshkati (Ed.), *Human mental workload*. Amsterdam: Elsevier, 139–183.

Hastie, R. & Dawes, R.M. (2001). Rational choice in an uncertain world: The psychology of judgment and decision making. Thousand Oaks, California: Sage Publications, Inc.

Heer, J., & Agrawala, M. (2006). Software Design Patterns for Information Visualization. *IEEE Transactions on Visualization and Computer Graphics. 12*(5). Sep/Oct 2006.

Heer, J., & Agrawala, M. (2008). Design considerations for collaborative visual analytics. *Information Visualization*, *7*(1), 49–62.

Heer, J., Viégas, F. B., & Wattenberg, M. (2007). Voyagers and voyeurs: Supporting asynchronous collaborative information visualization. *Proceedings of CHI 2007,* ACM Press, 1029–1038.

Heer, J., Viégas, F. B., & Wattenberg, M. (2009). Voyagers and voyeurs: Supporting asynchronous collaborative visualization. *Communicaitons of the ACM, 52*(1).

Henry, R. (1995). Improving Group Judgment Accuracy: Information Sharing and Determining the Best Member. *Organizational Behavior and Human Decision Processes*, *62*(2), 190–197.

Hendrix, T. D., Cross, I. I., James, H., Maghsoodloo, S., & McKinney, M. L. (2000). Do visualizations improve program comprehensibility? Experiments with control structure diagrams for Java. *ACM SIGCSE Bulletin*, *32*(1), 382–386.

Henningsen, D. D., Cruz, M. G., & Miller, M. L. (2000). Role of social loafing in predeliberation decision making. *Group Dynamics: Theory, Research, and Practice*, *4*(2), 168–175.

Heuer, Jr., R. J. (1999). *The psychology of intelligence*. Washington, D.C.: Center for the Study of Intelligence, Government Printing Office.

Hill, G. (1982). Group versus individual performance: Are N+ 1 heads better than one. *Psychological Bulletin*, *91*(3), 517–539.

Hill, W. C., Hollan, J. D., Wroblewski, D., McCandless, T. (1992). Edit wear and read wear. *Proceedings of ACM CHI 1992*, 3–9.

Hollingshead, A. B., McGrath, J. E., & O'Connor, K. M. (1993). Group task performance and communication technology. *Small Group Research*, *24*(3), 307–333.

Hosmer, D. W., & Lemeshow, S. (1989). *Applied logistic regression*. New York, NY: John Wiley & Sons.

intervention (2011). In *Merriam-Webster.com*. Retrieved from http://www.merriam-webster.com/dictionary/intervention

Isenberg, D. J. (1986). Group polarization: A critical review and meta-analysis. *Journal of Personality and Social Psychology*, *50*(1), 141–51.

Jahnke, H. N. (2003). *A History of Analysis*. Providence, RI: American Mathematical Society.

Janis, I.L. (1982). *Groupthink*: *Psychological studies of policy decisions and fiascos*. , Boston, MA: Houghton Mifflin.

Jarboe, S. (1996). Procedures for enhancing group decision making. In R.Y. Hirokawa & M.S. Poole (Eds.). *Communication and group decision making*. Thousand Oaks, CA: Sage Publications, 345–383.

Jehn, K. A., Northcraft, G. B., & Neale, M. A. (1999). Why differences make a difference: A field study of diversity, conflict, and performance in work groups. *Administrative Science Quarterly*, *44*, 741–763.

Jensen, E. (2007). Sensemaking in military planning: a methodological study of command teams. *Cognition, Technology, and Work*. Online First.

Johnson, D. & Johnson, R. (1985). The Internal Dynamics of Cooperative Learning Groups. In Richard Slavin, et al. (Eds.), *Learning to Cooperate, Cooperating to Learn*. New York: Plenum.

Johnson, D. & Johnson, R. (1981). Effects of Cooperative, Competitive, and Individualistic Goal Structure on Achievement: A Meta-Analysis. *Psychological Bulletin 89*(1), 47–62.

Johnston, R. (1997). *Decision Making and Performance Error in Teams: Research Results*. Arlington, VA: Defense Advanced Research Projects Agency.

Johnston, R. (2005). *Analytic culture in the U.S. intelligence community*: *An ethnographic study*. Washington DC: Center for the Study of Intelligence, Government Printing Office.

Kahaner, L. (1997). *Competitive intelligence: how to gather, analyze, and use information to move your business to the top*. Touchstone.

Kang, R. & Kiesler, S., (n.d.). Do Collaborators' Annotations Help or Hurt Asynchronous Analysis? Manuscript under review.

Karacapilidis, N., & Papadias, D. (2001). Computer supported argumentation and collaborative decision making: the HERMES system. *Information systems*, *26*(4), 259–277.

Karau, S. J., & Williams, K. (1993). Social loafing: A meta-analytic review and theoretical integration. *Journal of Personality & Social Psychology, 65*(4), 681-706.

Keel, P. E. (2006). Collaborative visual analytics: Inferring from the spatial organization and collaborative use of information. *IEEE Symposium on Visual Analytics Science and Technology*, 137–144.

Keel, P. E. (2007). EWall: A visual analytics environment for collaborative sense-making. *Information Visualization*, *6*(1), 48–63

Klahr, D., & Simon, H. A. (1999). Studies of scientific creativity: Complementary approaches and convergent findings. *Psychological Bulletin*, *125*, 524–543.

Kiesler, S., & Cummings, J. N. (2002). What do we know about proximity and distance in work groups? A legacy of research. *Distributed work*, *1*, 57–80.

Kight, L. (1996). *Elements of CI Success*. Briefing to SCIP Conference, Alexandria, VA, 28 March, 1996.

Klein, C., DeRouin, R.E., & Salas, E. (2006). Uncovering workplace interpersonal skills: A review, framework, and research agenda. In G. Hodgkinson & J.K. Ford (Eds.), *International review of industrial and organizational psychology* (vol. 21). Chichester, UK: Wiley.

Kohavi, R., Rothleder, N. J., & Simoudis, E. (2002). Emerging trends in business analytics. *Communications of the ACM*, *45*(8), 45–48.

Kozlowski, S. W. J., Ilgen, D. R., & Klimoski, R. (2006). Enhancing the effectiveness of work groups and teams. *Psychological Science in the Public Interest*, *7*(3), 77.

Kraut, R., Egido, C., & Galegher, J. (1988). Patterns of contact and communication in scientific research collaboration. *Proceedings of CSCW 1998,* ACM Press.

Kraut, R., Fussell, S. R., Brennan, S. E., & Siegel, J. (2002). Understanding effects of proximity on collaboration: Implications for technologies to support remote collaborative work. In P. Hinds & S. Kiesler (Eds.), *Distributed work.*. Cambridge, MA: MIT Press, 137–162.

Kraut, R., Fussell, S., & Siegel, J. (2003). Visual Information as a Conversational Resource in Collaborative Physical Tasks. *Human Computer Interaction*, *18*, 13–49.

Kravitz, D. A., & Martin, B. (1986). Ringelmann rediscovered: The original article. *Journal of Personality and Social Psychology*, *50*(5), 936–941.

Krizan, L. (1999). *Intelligence essentials for everyone*. Joint Military Intelligence College. NDIC PRESS. Retrieved from http://www.ndic.edu/press/8342.htm

Landgren, J. & Nulden, U. (2007). A study of emergency response work: patterns of mobile phone interaction. *Proceedings of CHI 2007,* ACM Press, 1323–1332.

Larkin, J., & Simon, H. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, *11*, 65–100.

Larson, J. R., & Christensen, C. (1993). Groups as problem-solving units: Toward a new meaning of social cognition. *British Journal of Social Psychology*, *32*(1), 5–30.

Larson, J. R., Christensen, C., Abbott, A. S., & Franz, T. M. (1996). Diagnosing groups: Charting the flow of information in medical decision-making teams. *Journal of Personality and Social Psychology*, *71*(2), 315–330.

Larson, J. R., Christensen, C., Franz, T. M., & Abbott, A. S. (1998). Diagnosing Groups: The Pooling, Management, and Impact of Shared and Unshared Case Information In Team-

Based Medical Decision Making. *Journal of Personality and Social Psychology*, *75*, 93–108.

Lascara, C., Wheless, G., Cox, D., Patterson, R., Levy, S., Johnson, A. E., & Leigh, J. (1999). TeleImmersive virtual environments for collaborative knowledge discovery. *Advanced Simulation Technologies*. San Diego, CA.

Laughlin, P. R., & Futoran, G. C. (1985). Collective induction: Social combination and sequential transition. *Journal of Personality and Social Psychology*, *48*, 608–613.

Laughlin, P. R., McGlynn, R. P. (1986). Collective induction: Mutual group and individual influence by exchange of hypotheses and evidence. *Journal of Experimental and Social Psychology*, *22*, 567–589.

Lavery, T., Franz, T., Winquist, J., & Larson, J. (1999). The Role of Information Exchange in Predicting Group Accuracy on a Multiple Judgment Task. *Basic and Applied Social Psychology*, *21*(4), 281–289.

Levine, J. M., & Moreland, R. L. (1990). Progress in Small Group Research. *Annual Review of Psychology*, *41*(1), 585–634.

Levinger, A. G., Graves, J., & Peckham, V. (1974). The Ringelmann effect: Studies of group size and group performance. *Journal of Experimental Social Psychology*, *10*(4), 371–384.

Linstone, H. A., & Turoff, M. (1976). *The Delphi Technique: Techniques and applications*. London, United Kingdom: Addison-Wesley.

Lipshitz, R. & Sherwood, J. J. (1978). The effectiveness of third-party consultation as a function of the consultant's prestige and style of intervention. *Journal of Applied Behavioral Science*, *14*, 493–509.

Littell, R., Milliken, G. A., Stroup, W. W., & Wolfinger, R. D. (1996). *SAS system for mixed models*. Cary, NC: SAS Institute.

MacEachin, D. J. (1994). *The Tradecraft of Analysis: Challenge and Change in the CIA*.

Malone, T., Crowston, K. (1994). The interdisciplinary study of coordination. *ACM Computing Surveys*, *26*, 87–119.

Malone, T.W. & Crowston, K.G. (1991). *Toward an interdisciplinary theory of coordination*. Massachusetts Institute of Technology, Center for Coordination Science, Cambridge, Mass.

Mannix, E., & Neale, M. (2005). What differences make a difference? The promise and reality of diverse teams in organizations. *Psychological Science in the Public Interest*, *6*, 31–55.

Mark, G., Carpenter, K., & Kobsa, A. (2003a). Are there benefits in seeing double? A study of collaborative information visualization. *Proceedings of CHI 2004*, ACM Press, 840–841.

Mark, G., Carpenter, K., & Kobsa, A. (2003b). A model of synchronous collaborative information visualization. *Proceedings of Info Vis 2003,* IEEE Press, 373-381.

Mark, G., Kobsa, A., & Gonzalez, V. (2002). Do four eyes see better than two? Collaborative versus individual discovery in data visualization systems. *Proceedings of InfoVis 2002,* IEEE Press, 249–255.

Marlow, C., Naaman, M., Boyd, D., & Davis, M. (2006). HT06, tagging paper, taxonomy, Flickr, academic article, to read. *Proceedings of the Seventeenth conference on hypertext and hypermedia*, 31–40.

Mathes, A. (2004). Folksonomies-cooperative classification and communication through shared metadata. *Computer Mediated Communication*, *47*.

McGrath, J. E. (1984). *Groups: Interaction and performance*. Englewood Cliffs, NJ: Prentice-Hall.

McGrath, J. E., & Berdahl, J. L. (1998). Groups, technology, and time: Use of computers for collaborative work. *Social psychological applications to social issues*, *4*, 205–228.

McGrath, J. E., & Hollingshead, A. B. (1994). *Groups interacting with technology: Ideas, evidence, issues, and an agenda*. Thousand Oaks, CA: Sage Publishers.

Mennecke, B. E., & Valacich, J. S. (1998). Information is what you make of it: The influence of group history and computer support on information sharing, decision quality, and member perceptions. *Journal of Management Information Systems*, *15*(2), 173–197.

Miranda, S. M., & Bostrom, R. P. (1999). Meeting facilitation: process versus content interventions. *Journal of Management Information Systems*, *15*(4), 89–114.

Mohammed, S. (2001). Toward an Understanding of Cognitive Consensus in a Group Decision-Making Context. *The Journal of Applied Behavioral Science*, *37*(4), 408–425.

Mohammed, S., & Dumville, B. (2001). Team mental models: Expanding theory and measurement through cross-disciplinary boundaries. *J. Org Behavior*, 22, 89–106.

Moede, W. (1927). Die Richtlinien der Leistungs-Psychologie. *Industrielle Psychotechnik, 4,* 193–207.

Mojzisch, A., & Schulz-Hardt, S. (2005). Information sampling in group decision making: Sampling biases and their consequences. In K. Fiedler & P. Juslin (Eds.), *Information sampling and adaptive cognition*. Cambridge: Cambridge University Press, 299–326.

Monk, A. (2003). Common ground in electronically mediated communication: Clark's theory of language use. In J.M. Carroll (Ed.), *HCI models, theories and frameworks: Towards a multidisiplinary science.* San Francisco, CA: Morgan Kaufmann, 265–289.

Morgan, B.B., Salas, E., & Glickman, A.S. (1993). An analysis of team evolution and maturation. *Journal of General Psychology*, *120*, 277–291.

Morris, M. R. (2008). A survey of collaborative web search practices. *Proceedings of CHI 2008*, ACM Press, 1657–1660.

Morris, M. R., & Horvitz, E. (2007). SearchTogether: an interface for collaborative web search. *Proceedings of the 20th annual ACM symposium on User interface software and technology*, 3–12.

Mullen, B. & Copper, C. (1994). The Relation Between Group Cohesiveness and Performance: An Integration. *Psychological Bulletin*, *115*, 210–27.

Newton, M. (2005). *Robert Pickton: The Vancouver Missing Women*. Retrieved from: http://www.trutv.com/library/crime/serial_killers/predators/robert_pickton/1.html

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, *2*, 175–220.

Novak, J., Wurst, M. (2005). Collaborative Knowledge Visualization for Cross-Community Learning. In S. O. Tergan & T. Keller (Eds.): *Knowledge and Information Visualization*, Berlin: Springer-Verlag Berlin Heidelberg, 95–116.

Office of the Director of National Intelligence. (2008). *Intelligence Community Information Sharing Strategy*. Retrieved from http://www.dni.gov/reports/IC_Information_Sharing_Strategy.pdf

Okhuysen, G. A., & Eisenhardt, K. M. (2002). Integrating knowledge in groups: How formal interventions enable flexibility. *Organization Science*, *13*(4), 370–386.

Olson, G. M., & Olson, J. S. (2000). Distance matters. *Human-computer interaction*, *15*(2), 139–178.

Orasanu, J. (1990). Shared Mental Models and Crew Performance. Paper presented at the 34th annual meeting of the *Human Factors Society*, Orlando, FL.

Osborne, A. (1957). Applied imagination. New York, NY: Scribner.

Palantir (n.d.) Retrieved from http://www.palantirtech.com/

Pang, A., & Wittenbrink, C. (1997). Collaborative 3D visualization with CSpray. *Computer Graphics and Applications*, *17*(2), 32–41.

Patterson, E. S., Woods, D. D., Tinapple, D., Roth, E. M., Finley, J. M., & Kuperman, G. G. (2001). Aiding the intelligence analyst in situations of data overload: From problem definition to design concept exploration. *Institute for Ergonomics/Cognitive Systems Engineering Laboratory Report, ERGO-CSEL.*

Patterson, E. S., Woods, D. D., Tinapple, D., & Roth, E. M. (2001). Using cognitive task analysis (CTA) to seed design concepts for intelligence analysts under data overload. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *45*, 439–443.

Paul, S. A. & Morris, M. R. (2009). CoSense: enhancing sensemaking for collaborative web search. *Proceedings of CHI 2009*, ACM Press.

Paul, S. A. & Reddy, M. (2010). Understanding together: Sensemaking in collaborating information seeking. *Proceedings of CSCW 2010*, ACM Press.

Pirolli, P., & Card, S. (1995). Information foraging in information access environments. *Proceedings of CHI 1995*, ACM Press, 51–58.

Pirolli, P., & Card, S. (1999). Information foraging. *Psychology Review*, 106(4), 643–675.

Pirolli, P., & Card, S. (2005). The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. *Proceedings of the International Conference on Intelligence Analysis*, 2–4.

Poltrock, S., Dumais, S., Fidel, R., Bruce, H., & Pejtersen, A.M. (2003). Information seeking and sharing in design teams. *Proceedings of GROUP 2003,* ACM Press, 239–247.

Poole, M.S. (1981). Decision development in small groups I: A comparison of two models. *Communication Monographs*, *48*,1–24.

Poole, M. S., & DeSanctis, G. (1989). Use of group decision support systems as an appropriation process. *Proceedings of the Twenty-Second Annual Hawaii International Conference on System Sciences: Emerging Technologies and Applications Track*, *4*.

Poole, M. S., & DeSanctis, G. (1990). Understanding the use of group decision support systems: The theory of adaptive structuration. *Organizations and communication technology*, *173*, 191.

Poole, M. S., Holmes, M., & DeSanctis, G. (1991). Conflict management in a computer-supported meeting environment. *Management Science, 37*(8), 926–953.

Priest, D. & Arkin, W. M. (2010, July 19). A hidden world, growing beyond control. *Washington Post*, Retrieved from http://projects.washingtonpost.com/top-secret-america/articles/a-hidden-world-growing-beyond-control/

QSurveillance (2010). Retrieved from www.qresearch.org/Public/QSurveillance.aspx

Qu, Y. & Hansen, D. L. (2008). Building shared understanding in collaborative sensemaking. *Proceedings of CHI 2008 Sensemaking Workshop*, ACM Press.

Reagan-Cirincione, P. (1994). Improving the accuracy of group judgment: A process intervention combining group facilitation, social judgment analysis, and information technology. *Organizational Behavior and Human Decision Processes*, *58*(2), 246-270.

Ringelmann, M. (1913). Research on animate sources of power: The work of man. *Annales de l'Instuit National Agronomique*, *12*, 1–40.

Russell, D. M., Stefik, M. J., Pirolli, P., Card, S. K. (1993). The cost structure of sensemaking. *Proceedings of InterCHI*. ACM Press.

Ryall, K., Forlines, C., Shen, C., & Morris, M. (2004). Exploring the effects of group size and table size on interactions with tabletop shared-display groupware. *Proceedings of CSCW 2004*, ACM Press, 284–293.

Safarik, M. E., Jarvis, J., & Nussbaum, K. (2000). Elderly female serial sexual homicide. *Homicide Studies*, *4*, 294–307.

Salas, E., DiazGranados, D., Klein, C., Burke, C. S., Stagl, K. C., Goodwin, G. F., & Halpin, S. M. (2008). Does team training improve team performance? A meta-analysis. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *50*(6), 903.

Salas, E., Rozell, D., Mullen, B., & Driskell, J.E. (1999). The Effect of Team Building on Performance: An Integration. *Small Group Research*, *30*, 309.

Sawant, N., Scharver, C., & Leigh, J. (2000). The tele-immersive data explorer: A distributed architecture for collaborative interactive visualization of large data-sets. *Proceedings of the Immersive Projection Technology Workshop,* Ames, Iowa.

Schneider, B., & Bowen, D.E. (1985). Employee and customer perceptions of service in banks: Replication and extension. *Journal of Applied Psychology*, *70*, 423–433.

Scholten, L., van Knippenberg, D., Nijstad, B. A., & De Dreu, C. K. (2007). Motivated information processing and group decision-making: Effects of process accountability on information processing and decision quality. *Journal of Experimental Social Psychology*, *43*(4), 539–552.

Scupelli, P., Kiesler, S., Fussell, S.R., & Chen, C. (2005). Project view IM: a tool for juggling multiple projects and teams. *Proceedings of the CHI 2005*. New York, NY: ACM Press, 1773–1776.

Seashore, S. (1954). *Group Cohesiveness in the Industrial Workgroup*. Ann Arbor, MI: University of Michigan Press.

Shaw, M.E. (1973). Scaling group tasks: A method for dimensional analysis. *JSAS Catalog of Selected Documents in Psychology*, *3*, 8.

Shepperd, J. A. (1993). Productivity loss in performance groups: A motivation analysis. *Psychological Bulletin*, *113*, 67–68.

Shneiderman, Ben. (1996). The eyes have it: A task by data type taxonomy for information visualizations. *Proceedings of Visual Languages 1996*. New York, NY: IEEE Press, 336–343.

Shrinivasan, Y. B., & van Wijk, J. J. (2008). Supporting the analytical reasoning process in information visualization. *Proceedings of CHI 2008*, ACM Press, 1237–1246.

Shrinivasan, Y. B., & van Wijk, J. J. (2009). Supporting exploration awareness in information visualization. *IEEE Computer Graphics and Applications Special Issue on Collaborative Visualization*, *29*(5).

Simonton, D. K. (2003). Scientific creativity as constrained stochastic behavior: The integration of product, person, and process perspectives. *Psychological Bulletin*, *129*, 475–494.

Slavin, R. (1989). Research on Cooperative Learning: Consensus and Controversy. *Educational Leadership 47*(4), 52–55.

Stasko, J., Catrambone, R., Guzdial, M., & McDonald, K. (2000). An evaluation of space-filling information visualizations for depicting hierarchical structures. *International Journal of Human Computer Studies*, *53*, 663–694.

Stasko, J., Gorg, C., Liu, Z., & Singhal, K. (2007). Jigsaw: Supporting investigative analysis through interactive visualization.

Stasko, J., Görg, C., & Liu, Z. (2008). Jigsaw: supporting investigative analysis through interactive visualization. *Information Visualization*, *7*, 118-132.

Stasser, G. (1992). Information salience and the discovery of hidden profiles by decision-making groups: A. *Organizational Behavior and Human Decision Processes*, *52*(1), 156–181.

Stasser, G., Taylor, L. A., & Hanna, C. (1989). Information sampling in structured and unstructured discussions of three-and six-person groups. *Journal of Personality and Social Psychology*, *57*(1), 67–78.

Stasser, G., & Titus, W. (1985). Pooling of unshared information in group decision making: Biased information sampling during discussion. *Journal of Personality and Social Psychology*, *48*(6), 1467–1478.

Stasser, G., & Titus, W. (1987). Effects of information load and percentage of shared information on the dissemination of unshared information during group discussion. *Journal of personality and social psychology*, *53*(1), 81–93.

Stasser, G., Vaughan, S. I., & Stewart, D. D. (2000). Pooling Unshared Information: The Benefits of Knowing How Access to Information Is Distributed among Group Members. *Organizational Behavior and Human Decision Processes*, *82*(1), 102–116.

Steiner, I. D. (1972). *Group process and productivity*. New York, NY: Academic Press.

Steiner, I. D., & Rajaratnam, N. (1961). A model for the comparison of individual and group performance scores. *Behavioral Science*, *6*(2), 142–147.

Stewart, D. D., Billings, R. S., & Stasser, G. (1998). Accountability and the discussion of unshared, critical information in decision-making groups. *Group Dynamics: Theory, Research, and Practice*, *2*(1), 18–23.

Suthers, D., & Hundhausen, C. (2001). Learning by Constructing Collaborative Representations: An Empirical Comparison of Three Alternatives. *Proceedings of Euro CSCL 2001*. Maastricht, Netherlands: Maastricht McLuhan Institute, 577–592.

Tang, J. C., Zhao, C., Cao, X., & Inkpen, K. (2011). Your time zone or mine?: a study of globally time zone-shifted collaboration. *Proceedings of CSCW 2011*, ACM Press, 235–244.

Thompson, R. S., Rantanen, E. M., & Yurcik, W. (2006). Network intrusion detection cognitive task analysis: textual and visual tool usage and recommendations. *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting*, Human Factors and Ergonomics Society, 669–673.

Tinapple, D., & Woods, D. (2003). Message overload from the inbox to intelligence analysis: how spam and blogs point to new tools. Proceedings of the *Human Factors and Ergonomics Society Annual Meeting, 47*, 419–423.

Tolcott, M. A., Marvin, F. F., and Bresoick, T. A. (1989). *The confirmation bias in military situation assessment*. Reston, VA: Decision Science Consortium.

Treverton, Gregory F. *Reshaping National Intelligence in an Age of Information.* Cambridge, MA: Cambridge University Press, 2001.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*, 207–232.

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, *90*, 293–315.

Tudor, T. R., Trumble, R. R., & Diaz, J. J. (1996). Work-teams: Why do they often fail? *S.A.M. Advanced Management Journal*, *61*(4), 31-40.

Valledor, J. C. (2010). *Connecting the Dots: Enduring Challenges in the Nation's Information Sharing Environment*. Fort Leavenworth, KA: School of Advanced Military Studies, United States Army Command and General Staff College.

Van de Ven, A., Delbecq, A., & Koenig, R. (1976). Determinants of coordination modes within organizations. *American Sociological Review, 41*, 322–338.

Veerasamy, A., & Belkin, N. (1996). Evaluation of a tool for visualization of information retrieval results. *Proceedings of SIGIR 1996*. ACM Press, 85–92.

Viégas, F., Wattenberg, M., & Dave, K. (2004). Studying cooperation and conflict between authors with history flow visualizations. Proceedings of the *CHI 2004*. New York, NY: ACM Press, 575–582.

Viégas, F.B. and Wattenberg, M. (2006). Communication-Minded Visualization: A Call to Action. *IBM Systems Journal, 45*(4).

Walsh, J. P., Kucker, S., Maloney, N. G., & Gabbay, S. (2000). Connecting minds: Computer-mediated communication and scientific work. *Journal of the American Society for Information Science*, *51*(14), 1295–1305.

Wang, H. C., Cosley, D., & Fussell, S. R. (2010). Idea Expander: Supporting group brainstorming with conversationally triggered visual thinking stimuli. In *Proceedings of CSCW 2010*, ACM Press, 103–106.

Wattenberg, M. (1999). Visualizing the stock market. Proceedings of the *CHI 1999* New York, NY: ACM Press, 188–189.

Wattenberg, M., & Kriss, J. (2006). Designing for Social Data Analysis. *IEEE Transactions on Visualization and Computer Graphics, 12*(4), 549–557.

Weick, K.E. (1993). The collapse of sensemaking in organizations: the Mann Gulch disaster. *Administrative Science Quarterly*, *38*(4), 628–652.

Weick, K. E. (1995). *Sensemaking in organizations*. Thousand Oaks, CA: Sage Publications, Inc.

Weldon, E., & Gargano, G. M. (1988). Cognitive loafing: The effects of accountability and shared responsibility on cognitive effort. *Journal of Personality and Social Psychology*, *14*, 159–171.

White, R. W., & Roth, R. A. (2009). Exploratory search: Beyond the query-response paradigm. In G. Marchionini (Ed.), *Synthesis Lectures on Information Concepts, Retrieval, and Services*. San Rafael, CA: Morgan & Claypool, 1–98.

Wickens, C. D., & Carswell, C. M. (1995). The proximity compatibility principle: Its psychological foundation and its relevance to display design. *Human Factors*, *37*, 473–494.

Willett, W., Heer, J., & Agrawala, M. (2007). Scented widgets: Improving navigation cues with embedded visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 1129–1136.

Willett, W., Heer, J., Hellerstein, J., & Agrawala, M. (2011). CommentSpace: structured support for collaborative visual analysis. *Proceedings of CHI 2011*, ACM Press, 3131–3140.

Wright, W., Schroh, D. Proulx, P., Skaburskis, A., & Cort, B. (2006). The Sandbox for Analysis: Concepts and Methods. *Proceedings of CHI 2006*, ACM Press, 801–810.

Wood, J., Wright, H., & Brodlie, K. (1997). Collaborative visualization. *Proceedings of Vis 1997*. IEEE Press.

Woolley, A. W. (1998). Effects of intervention content and timing on group task performance. *The Journal of Applied Behavioral Science*, *34*(1), 30.

Woolley, A.W., Gerbasi, M.E., Chabris, C.F., Kosslyn, S.M., & Hackman, J.R. (2008) Bringing in the experts: How team composition and collaborative planning jointly shape analytic effectiveness. *Small Group Research*, 39, 352-–.

Yamane, D. (1996). Collaboration and its discontents: Steps toward overcoming barriers to successful group projects. *Teaching Sociology*, 24(4), 378–383.

Zaccaro, S. J., & McCoy, M. C. (1988). The Effects of Task and Interpersonal Cohesiveness on Performance of a Disjunctive Group Task1. *Journal of Applied Social Psychology*, *18*(10), 837–851.

Zawacki, R. (1994). Do IS and teams mix? Retrieved from http://www.computerworld.com/news/1994/story/0,11280,1593,00.html

Zhang, J., & Norman, D. (1994). Representations in Distributed Cognitive Tasks. *Cognitive Science*, *18*(1), 87–122.

APPENDIX A: TAGGING NODES AND

RELATIONSHIPS

My "tagging feature" allows users to add information directly to their visualization, and I designed and developed a prototype visualization tool that includes two types of tagging: nodes and relationships.

A.1. IMPLEMENTATION

To evaluate tagging as a mechanism for improved coordination and sensemaking processes, I designed and implemented a prototype tag-enabled visualization tool[3]. The prototype visualization tool was built in order to add two tagging features: tagging nodes and relations.

System overview

The tool was implemented via Java programming language. This visualization tool depicted all of the characters in the seven fictional cases as well as their relationships to one another. This information and layout was identical to the visualization tool used in all three studies, NetDraw. Nodes could be moved around on the screen. However, there was no search capability; for example, users could not search for characters with similar attributes.

Figure A. illustrates a high-level overview of the tagging features built into the prototype visualization tool. The first feature allows users to tag a node, which corresponds to a fictional character, with any keyword (See Figure). When a user wants to annotate a particular individual with a particular piece of information, the user can right click on the desired node and a pop-up window appears. The pop-up window includes an input field wherein the user can add a tag about that individual. To review tags for a node or to include additional tags, users can right-click the node again.

---

[3] This prototype visualization tool was implemented by James D. Williams, an undergraduate research assistant.
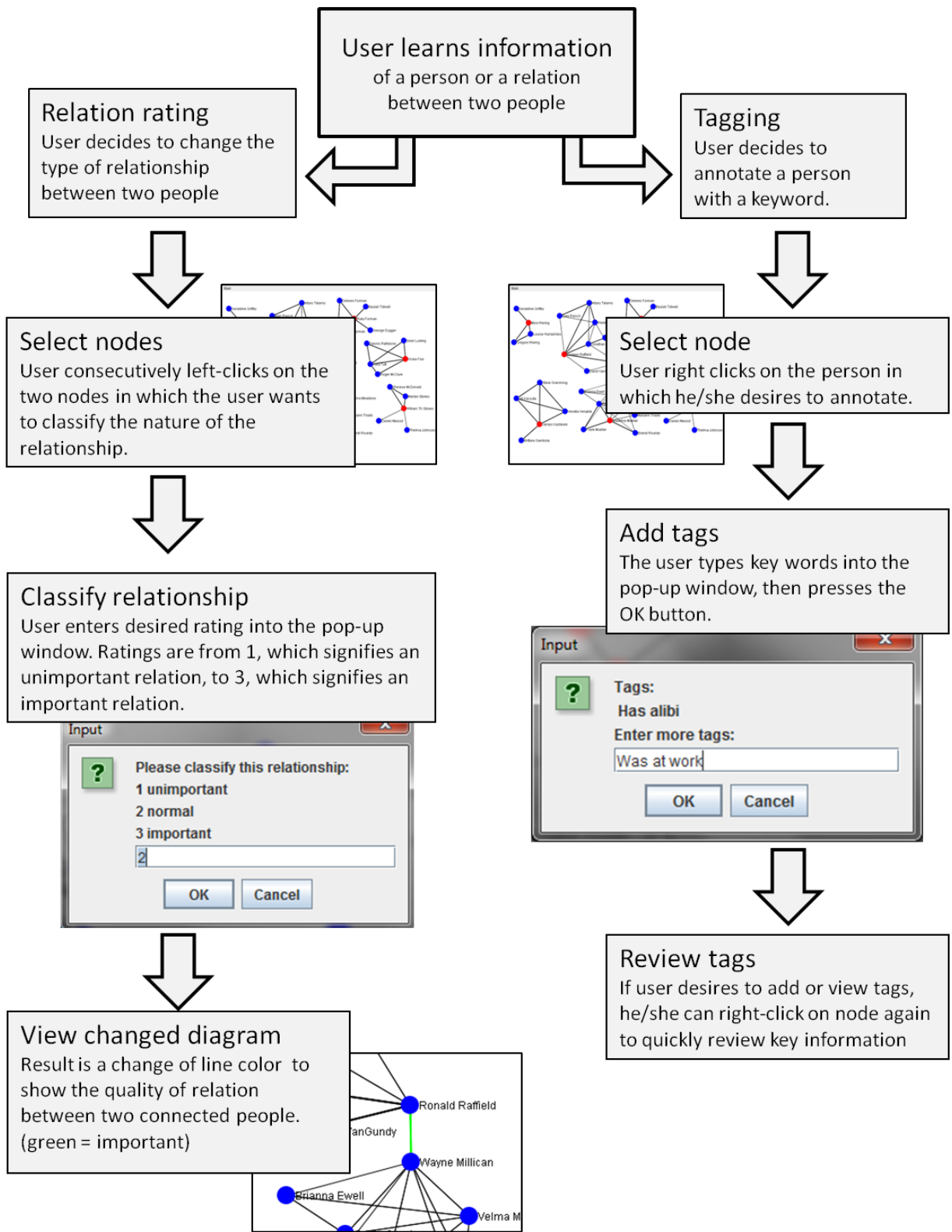
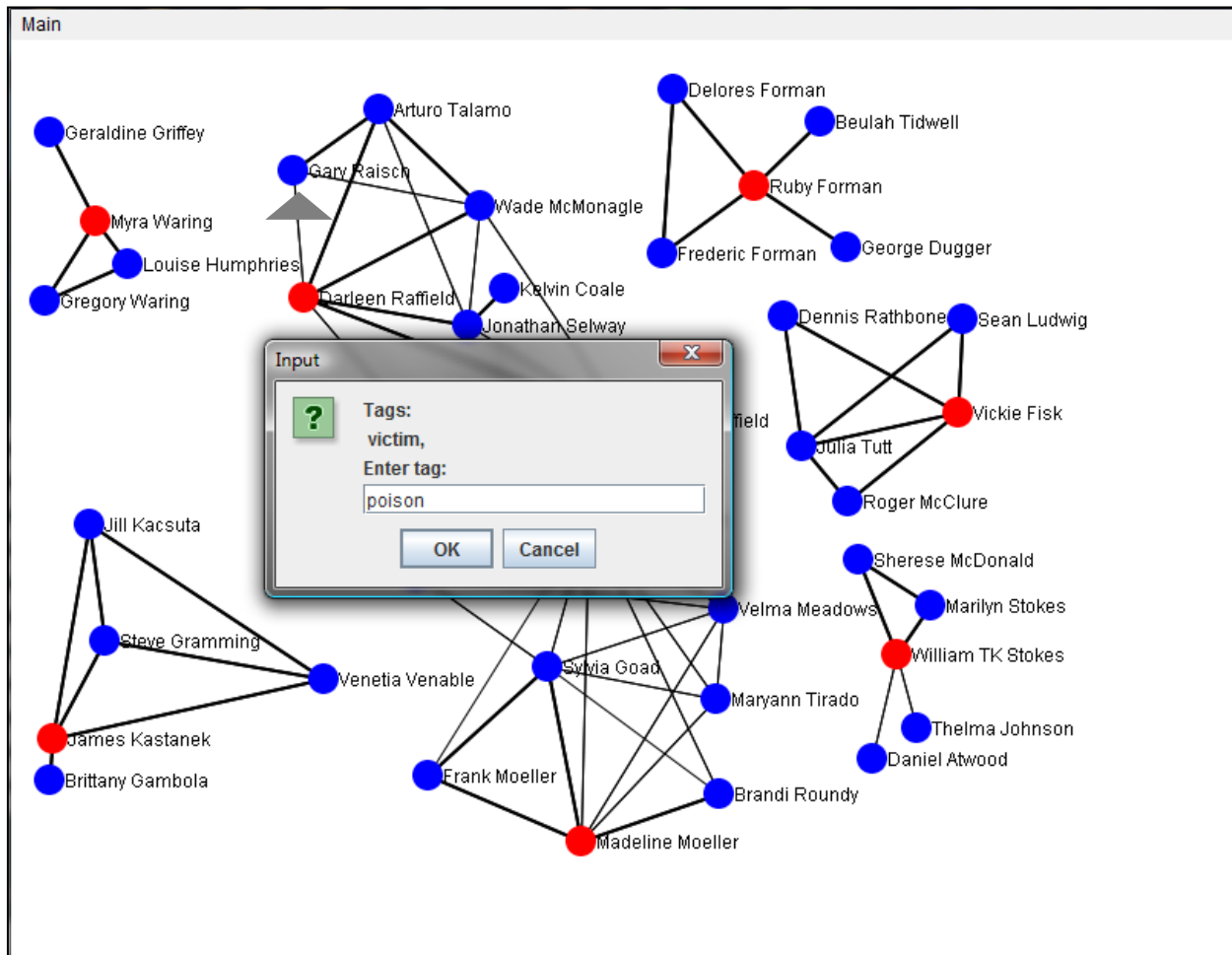**Figure A.1. Overview of the tag-enabled visualization tool.**

Figure A.2. Screenshot of a tagging feature in a prototype visualization tool.

Left clicking the node brings up a pop-up input window. This window shows any existing tags associated with the individual and also contains an open-text area for adding new tags. In this screenshot, a tag input window has been opened for Darlene Raffield. This character has already been tagged with the word *victim*, and she's currently having a new tag, *poison* (i.e., her cause of death), added to her node..
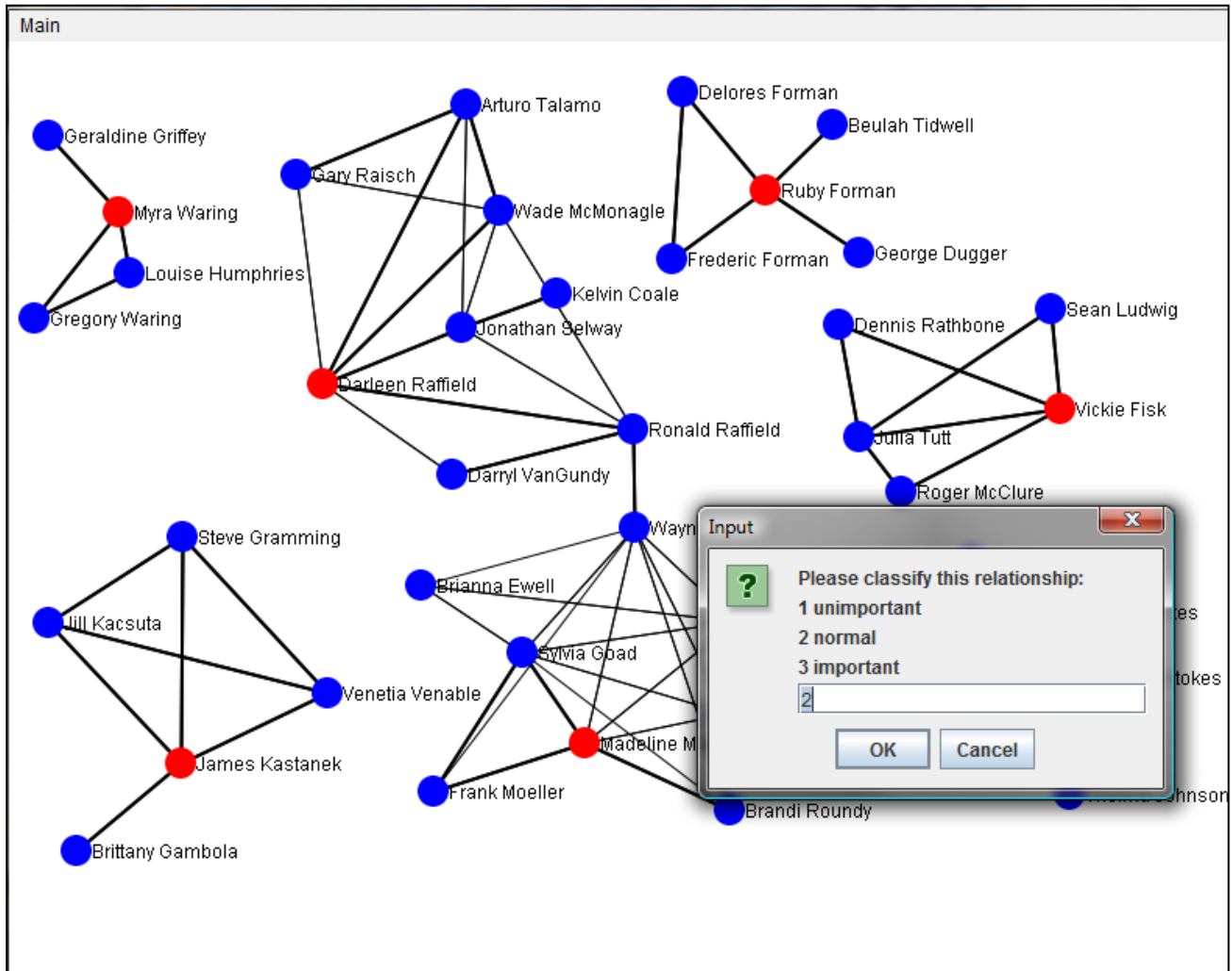
Figure A.3. Screenshot of the relationship tagging feature in the prototype visualization tool.
Clicking on two individuals opens up a pop-up input window that allows the user to classify the relationship between those two individuals as unimportant, normal, or important. A *normal* relationship is represented by a black line. Marking a relationship as *unimportant* changes the line connecting the two individuals from black to gray, and marking the relationship as *important* changes the line from black or gray to green.
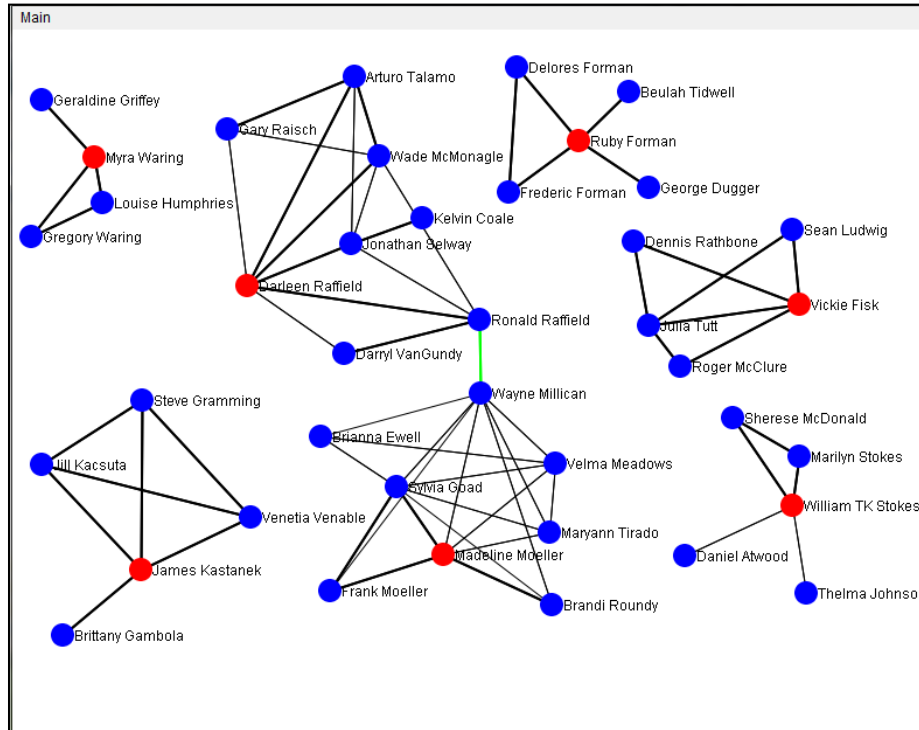
Figure A.4. Screenshot of a relationship marked at important.

The green line connecting the characters Ronald Raffield and Wayne Millican indicates this connection has been marked *important*.
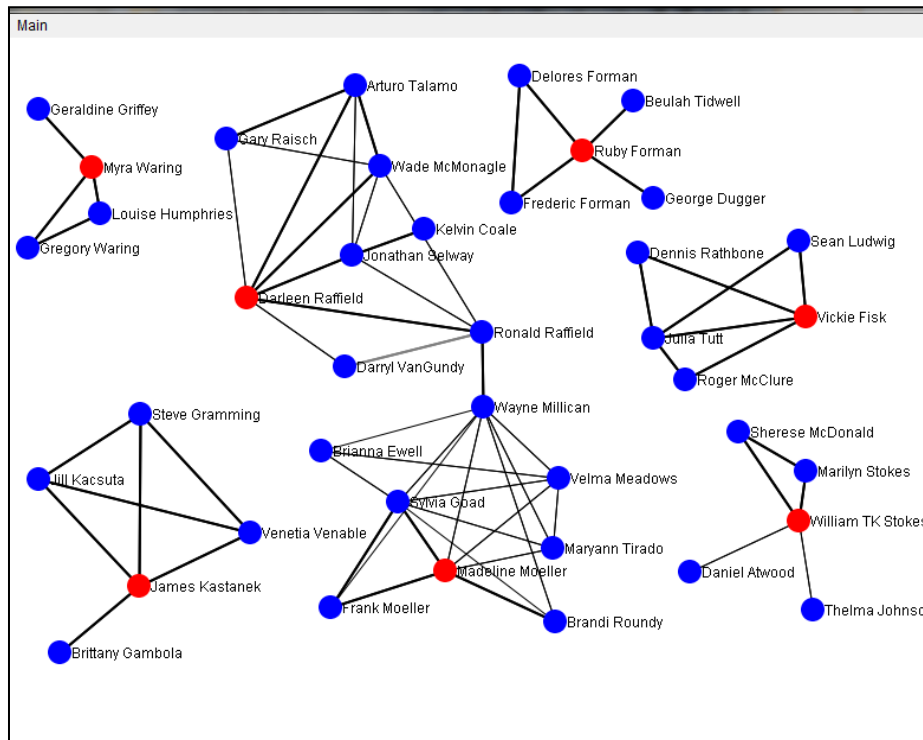

Figure A.5. Screenshot of a relationship marked as unimportant.

The gray line connecting the characters Ronald Raffield and Darryl VanGundy indicates that this connection has been marked *unimportant*.

The second tagging feature of the prototype is its ability to classify a relation or a relationship between two characters, which is represented by a line that connects the two nodes (See Figure). Users have the option to rate a relation between two characters as *unimportant*, *normal*, or *important*. In the default state, all relationships are categorized as *normal* and are designated by the color black. To change the status, the two desired nodes are clicked consecutively and a pop-up window appears. The pop-up window asks the user to input a new rating, from 1 to 3, where one represents *unimportant*, two represents *normal*, and three represents *important*. If the link is changed to *unimportant*, the line color changes to gray, whereas a link marked *important* changes to green (See Figure and Figure). While this is not a tag in the traditional meaning of the word,, I refer to it as a tag because it serves a similar purpose, albeit visual. Rather than annotating an item with a keyword, this feature is a visual annotation of a linkage; i.e., a type of visual information scent (Willett, Heer, & Agrawala, 2007).

A.2. EVALUATION

I performed four think-aloud user evaluations to assess the usability of the tag-enabled visualization prototype.

Method

Four participants were recruited to a voluntary think-aloud session to test the prototype tagging visualization tool (2 female, 2 male, 75% US born; age range 25-35). Participants were recruited via online resources. As in previous studies, fluency in English was a requirement to participate. Participants were not paid for their participation and were told the experiment would last 20 minutes.

The design of the think-aloud user studies aimed at an assessment of the usability and likeability of the tagging features. Participants were told they were a part of pretesting for a new visualization tool. Participants worked alone and only one think-aloud was conducted at a time. They were briefed on the detective mystery study and given access to all the seven cases. Participants were trained on how to use the original visualization used in all three studies (NetDraw), and also trained to use the prototype visualization tool. Participants were explicitly asked to tag character nodes and classify relationship links based on information they gathered from the documents. They were also told that their tags would be shared with a partner. An experimenter sat beside the participant during the study and encouraged them to say out-loud

what they were looking at, thinking, doing, and feeling, as they went about the task of tracking down a serial killer using the prototype tool. After fifteen minutes, participants were stopped and asked a series of questions regarding their overall experience with the prototype.

Because the tool was a prototype, I used "Wizard-of-Oz" techniques to better explore the collaborative properties of the tagging feature. For example, two participants were given access to the prototype tool in a "clear" state, meaning that were no existing tags on the nodes or links. The other two participants received a pre-tagged prototype and were told that the prototype had been tagged by a collaborator. This pre-tagged version contained five node tags and three relationship tags.

Findings
Overall, users found both tagging features useful. They pointed out several improvements that would increase the feature's utility.

All four users successfully added tags to several nodes and changed the classification of links. Two users commented that the tagging node feature was useful as a note keeper for important information they found during their investigation. All four users expressed a positive reaction to the relation rating feature. One user exclaimed, "Cool! The green really sticks out and [it] is obvious that it [the relationship] is important." This color change also served as a form of feedback for users; they knew that the system had stored their rating. With respect to the node tags, one user was unsure if the tag had been saved and clicked the node again to make sure that the tag had been stored successfully.

All four participants considered their collaborator as they tagged. For the two participants who received the tool without existing tags, thoughts on a collaborator were minimal. The two participants who received a pre-tagged tool, on the other hand, found they could easily find important links, but also found it difficult to find existing tagged nodes. One participant tagged a node as "blunt inst" and later noticed that "his partner" had tagged a different node with the keywords "blunt instrument." Realizing that these tags were referring to the same concept, the participant wanted to link the two concepts, which was not possible within the prototype because of their alternate spelling.

When asked about their experience using the tagging features after their trial, all four users said that tags were an appropriate function for the type of information and task being performed. One user succinctly said, "It makes sense." But the lack of certain capabilities made the tagging features less useful for certain tasks. Specifically, all four users expressed a need for a hover feature, such that when a user places their mouse's cursor over a node, a pop-up would appear with existing tags. This would allow users to easily review and find tags and also serve an important feedback function that many felt it was missing. One participant pointed out that "It is cool that someone working with me could see my tags and any changes but they wouldn't know why I had made those changes or how I found that information."

A.3. DISCUSSION

These evaluations suggest that tagging features could highlight opportunities for discussion, facilitate activity awareness, and promote searching for relevant information. Participants who were given the prototype already seeded with tags noticed their partner's work, which is a form of activity awareness. Additional cues could be provided to help users find previously added tags. Participants could also easily see relations that had been marked due to visual feedback. Similarly, nodes could have a comment flag automatically attached to them if a user added a tag to that node. This would facilitate activity awareness in the same manner as the tagging relationships feature.

Participants were also curious as to why their partner had annotated specific items, which could instigate opportunities for discussion. If a user wanted to understand their partner's reasoning for adding a tag, they could directly ask their partner. Essentially, users wanted to see the evidence behind the added tag. One way to build this into the visualization tool would be to directly connect the tag to its information sources, in this case to the crime documents.

Participants wanted to see existing tags more easily as a way to better search for visual facts. Tags were a step in the right direction, but the limited functionality restricted searching to the tool itself. One of the major benefits of tags is their capability to review and find items that are tagged with similar concepts. However, this prototype did not have that kind of search functionality. The addition of a simple hover pop-up window, as suggested by a participant, and the capability to highlight similarly tagged items, would greatly facilitate visual search.

Another potential issue with tags is that people are allowed to use any terminology they choose, resulting in different keywords or tags for the same concept (i.e. "blunt instrument" versus "blunt inst"). To mitigate this, a tag window could be included to show users tags that had already been used, and also to facilitate searching for items with common tags. Getting collaborators to use a common set of tags would reduce individual efforts for creating new keywords; at the same time, it would simultaneously encourage a common mental model among collaborators (Willett et al., 2011).

While certain benefits were suggested from user evaluations, it remains unclear whether tagging features would definitely increase visualization use or support finding patterns in the data. The evaluation was not a controlled laboratory study and users were specifically told to add tags. Left to their own devices, users may have chosen not to add tags. The evaluation also did not have performance measures. Participants only performed the task for a minimal amount of time and their focus was on the usability of the design tool rather than whether or not these tools lead users to complex problem solving success. Consequently, although some users were positive about their experience, it is unclear whether the tagging features could support the integration of visual insights.

A.4. SUMMARY

Visualization tools can improve coordination by highlighting opportunities for discussion and facilitating division of labor through activity awareness. Visualization tools can also improve collaborative sensemaking by encouraging visualization use, promoting search for important facts, and supporting finding patterns in the data.

In this particular design concept, a visualization tool enhanced with tagging features allowed users to annotate nodes with keywords, as well as to classify relationships as either important or unimportant. Users engaged with the tool and found that it fostered opportunities for discussion and activity awareness along with facilitating visual search. This design concept is not meant to be an exhaustive example of possible features—far, far, from it. But it does serve to demonstrate the potential for simple features, such as tagging, to better facilitate collaborative problem solving through aiding coordination and sensemaking processes.