Teammate Inaccuracy Blindness: When Information Sharing Tools Hinder Collaborative Analysis

Ruogu Kang[1], Aimée A. Kane[2], Sara Kiesler[1]

Human Computer Interaction Institute[1]
Carnegie Mellon University
5000 Forbes Ave., Pittsburgh, PA 15213
{ruoguk,kiesler}@cs.cmu.edu

Palumbo-Donahue School of Business[2]
Duquesne University
600 Forbes Avenue, Pittsburgh, PA 15282
kanea@duq.edu

ABSTRACT

Asynchronous collaborative analysis is important in many fields, but information sharing can be a bottleneck. Tools for annotating, organizing, and summarizing information can help, but their value will likely depend on the accuracy of teammates' information. To document this claim, two experiments examined participants' performance on a complex detective task when they asynchronously received information from a teammate in a collaboration tool or received no such information. We found that receiving a progress report containing accurate information was associated with improved performance (vs. no information shared in a tool) but worse performance when the information was inaccurate. Teammates were evaluated as helpful even when they were not. Our findings point to a phenomenon of *teammate inaccuracy blindness* that arises when teammates provide inaccurate information. We propose some strategies for helping collaborators avoid or lessen this effect.

Author Keywords

Asynchronous collaboration, collaborative analysis, information sharing, teams, groups, sensemaking, problem solving

ACM Classification Keywords

H.5.3. Group and Organizational Interfaces: Asynchronous interaction, Computer-supported cooperative work.

INTRODUCTION

Remote asynchronous collaboration is increasingly common as analysts in domains ranging from science to intelligence tackle problems with large amounts of complex data [4, 11]. However, information sharing can be challenging for them. Remote asynchronous collaboration reduces the kinds of spontaneous communication that help analysts resolve discrepancies [2], and it also provides fewer opportunities for analysts to integrate one another's expertise [31]. If so, analysts may fail to effectively share information and collaborative performance may suffer.

Researchers have developed IT-enabled tools to help analysts make sense of large amounts of complex data [29]. Increasingly, these tools support collaborative information sharing among analysts. Annotation tools, for example, enable analysts to highlight specific pieces of evidence for one another [16]. Organizing and visualizing tools enable analysts to display linkages across dispersed pieces of evidence, helping them point out patterns [2, 10, 13].

Researchers have not yet examined how the quality of the information conveyed through different tools affects collaborative analysis. Teammates' information could be highly accurate and useful, but it also could be inaccurate, erroneous, or irrelevant. In other domains, when teammates share inaccurate information, it can send them off in the wrong direction [12]. This possibility is especially critical in teams working asynchronously. Team members may assume others on the team offer help only when the help is useful. Not having continual communication with each other may hinder collaborators from spotting and ruling out misunderstandings [23]. In this paper, we tackle this issue in the context of asynchronous collaborative analysis of a complex crime problem.

This research addresses whether team members will distinguish between inaccurate and accurate information conveyed in information sharing tools. We experimentally examine how receiving information that varies in its accuracy affects the problem solving performance of a remote team member. We find evidence that the recipient may fail to recognize that teammates' information is inaccurate. In our studies, failure occurred when inaccurate information was shared in a progress report tool. We coin the term *teammate inaccuracy blindness* to highlight the influence that inaccurate information can have on problem-solving performance and on inflated evaluations of teammates.

Levels of sensemaking in collaborative analysis

Our consideration of information accuracy in asynchronous collaborative analysis builds on the concept of "sensemaking." Sensemaking refers to a process of making data intelligible by identifying evidence, schematizing and perceiving patterns and anomalies, and, at the highest level, developing insights, hypotheses, and conclusions [28]. Pirolli and Russell [29] have pointed to an emerging recognition that sensemaking is increasingly conducted by teams. They argue research is needed to understand how to

facilitate the transfer of useful sensemaking information from one collaborator to another.

Analysts can use collaboration tools to share sensemaking information [16]. Generally, these tools should be more useful to collaborators if the information provider is on the right track and uses the tool to share accurate sensemaking information. But there is no guarantee that an analyst teammate is on the right track [12]. The usefulness of information sharing tools may depend on the accuracy of the sensemaking information the teammate conveys.

For purposes of our argument, we simplify the process of sensemaking by positing three key levels: (1) identifying evidence, (2) using categorization or schemas to organize evidence, and then (3) generating hypotheses and conclusions, and marshaling evidence to support them [28]. We propose that when analysts work together, annotation or tagging tools afford collaborative information sharing at the first level of sensemaking. Organizing worksheets, visualizations, and other ways of identifying categories and patterns afford collaborative information sharing at the second level of sensemaking. Summary report tools that convey hypotheses or conclusions, and give evidence in support of them, afford collaborative information sharing at the third level of sensemaking.

Because any sensemaking information added to complex data analysis potentially contributes cognitive load to an analysts' work [29], it is important to consider whether this information provides help to analysts beyond what they can glean from the underlying data. We now consider the possibly variable quality of the information collaborators may share through tools at each level of sensemaking.

Annotating, tagging, and highlighting
The sensemaking framework suggests theoretical reasons to believe that an initial analyst's annotations or tags of data may be helpful to a teammate when the information space is complex. Shared annotations may be helpful because they direct attention to different parts of a problem or evidence, thus helping analysts decide what parts of a problem are important or how to decompose the problem [24,37]. For instance, a police detective in one department who is familiar with the geography of a city might tag a crime location report with a comment that the neighborhood has experienced similar crimes.

Nonetheless, there is a risk associated with sharing inaccurately annotated information. Annotations may be irrelevant notations that only increase cognitive load and time spent in analysis. Annotations also may direct a problem solver's attention to an incorrect strategy. Sloppy or incorrect annotations could guide collaborators to erroneous conclusions [16].

Organizing, categorizing, and creating schemas
The second level of sensemaking information represents knowledge categories and schemas that can help collaborators. The organization of information is itself a shared information space that may be even more useful than the information itself [8]. Researchers have proposed strategies such as creating new categories to help solve insight problems [17, 21]. Categorizing and creating schemas highlights associations across data; distant associations can trigger innovative thinking and insights [22]. For example, when detectives group crimes according to the weapon used, their teammates may more readily detect a pattern, if one is present. Along related lines, visualization tools that display otherwise hidden linkages across dispersed information have improved complex problem solving [2, 10,13, 33].

There is a risk of sharing inaccurate organizing information. Categories and schemas are higher-level representations of information, and using them may require expertise in the domain [5]. Teammates may misunderstand or misuse the categories or schemas created by collaborators. Patterns can reveal relationships that look interesting but are innocent (false positives). If teammates highlight inaccurate categories or schemas and insufficiently link them to sources in the data, they could be misleading.

Summarizing and pointing to solutions
The highest level of representation in sensemaking is the process of creating hypotheses and drawing inferences, and connecting evidence to those ideas [28]. Analysts who communicate their hypotheses and conclusions, with evidence in their favor, provide shortcuts for their colleagues. Information sharing tools that afford sharing these developed hypotheses and conclusions should be a powerful aid to analysts and help their efficiency. The first person to close in on a demonstrably correct solution could share this knowledge and save teammates' time and effort [23].

There is a risk of sharing inaccurate hypotheses and conclusions. If the shared hypothesis or conclusion is the wrong one, is weakly supported by the evidence, or is poorly explained, then the teammate could be sent in the wrong direction. It would also magnify confirmation bias, that is, searching only for evidence in favor of a particular hypothesis [25]. Using an information-sharing tool to exchange summary hypotheses or conclusions may be worse than not sharing information at all.

In sum, we argue that the accuracy of different levels of sensemaking information shared through tools by an analyst teammate will affect the subsequent problem solving of a remote analyst. In this research, we examine the use and effects of three existing tools intended to aid each level of sensemaking. We compare the effects of an initial analyst who shares no information, other than his or her case data (no tool control), with the effects of an initial analyst's information conveyed in an annotation tool meant to help analysts identify evidence (sensemaking level 1), an organizing worksheet tool meant to help analysts share interconnections among categories of evidence (sensemaking level 2), or a summarizing report tool meant

to help analysts share hypotheses and evidence in support of hypotheses (sensemaking level 3).

Taking together our reasoning about the importance of information accuracy and the levels of sensemaking conveyed in a tool, we hypothesize the following:

H1a. When an asynchronous partner uses collaboration tools to share accurate information, the receiver of this information will show better problem solving than a control partner who receives no extra information.

H1b. When an asynchronous partner uses collaboration tools to share inaccurate information, the receiver of this information will show worse problem solving than a control partner who only receives no extra information.

H2. The different effects of accurate and inaccurate information will be greater when the partner uses the summarizing report tool because that tool conveys the highest level and most efficient sensemaking information.

EXPERIMENT 1

To test the hypotheses, we used a homicide detective task in which participants needed to cull through numerous police reports, crime statistics, and bus routes, and piece together clues to identify a serial killer [2]. Two remote analysts collaborated in a sequentially interdependent manner. Each of them used three tools for information sharing, an annotation tool, a worksheet organizing tool, and a progress report. Halfway through the task, initial analysts shared their cases and, in all conditions except the no tool control, information about those cases with final analysts using one of the tools (i.e., annotation, worksheets, or reports). The final analysts ultimately prepared a final summary report and provided an evaluation of their collaborator. We coded these summary reports for evidence that the analysts had identified the serial killer, which provided an outcome measure of effective problem solving. We measured the evaluations that analysts made of one another because we wondered if their evaluations would reflect the quality of information they had received.

Method
Participants and design
Fifty-nine pairs participated in the experiment. All participants were undergraduate or graduate students. Each participant was paid $15 for participating in the experiment for 1.5 hours.

Participants were instructed to solve as many of the eight homicide cases as possible and to write a summary report on behalf of their team. The pairs were randomly assigned to one of three tool conditions or a control condition. Within each pair, one participant was randomly assigned to be the initial analyst and the other participant to be the final analyst. In subsequent discussion we refer to these participants by their roles of initial or final analyst.

The final analyst could not solve the serial killer case without case data held by the initial analyst. These data were shared midway through the task with all final analysts. In the tool conditions, the initial analyst shared additional information about these newly-shared case data. We were mainly concerned with the accuracy of the information that the initial analyst provided and the subsequent performance of the final analyst. We tested the hypotheses by comparing the problem-solving performance of final analysts who had received information from the initial analyst in information-sharing tools against the performance of final analysts who received no extra information in tools from the initial analyst. We also examined how these final analysts evaluated their asynchronous collaborator.

Task
We adapted the homicide detective task involving a serial killer used by Balakrishnan et al. [2]. Analysts needed to review multiple, complex homicide case files and other data such as bus route information and maps to piece together disparate clues across five cases that identified a serial killer responsible for four homicides. To simulate the complexity of real-world data, in addition to the four homicides perpetrated by the serial killer, analysts also had information on two unrelated homicides and two homicides that could not be solved with the evidence given. Whether or not analysts solved the serial killer crime was the problem solving success criterion.

During the first phase of the task, serial killer related cases and material were distributed unequally between teammates to create a situation in which the final analyst might benefit from receiving information from the initial analyst about the serial killer crime. Each teammate initially received four different crime case documents and relevant supplementary materials such as bus route information and an area map. The initial analyst received three serial killer cases and one unrelated case. The final analyst received one unrelated case, two unsolvable cases, and one serial killer case. They each set about their task of crime analysis, trying to identify the perpetrator of each homicide.

Halfway through the period for analysis, final analysts received access to all case data, including data from the initial analyst's four cases. In the tool conditions, final analysts also received information provided by the initial analyst in one of the tools. As noted above, by design, the final analysts could not solve the serial killer crime with the information provided at the start of the experiment because they had only one serial killer case. The other case data provided by their teammate at the mid-point was essential for solving the serial killer crime because it included three additional serial killer cases and an unrelated case in which the serial killer appeared as a witness. Our question was whether the additional sensemaking information from the initial analyst would help the final analyst in solving the serial killer crime beyond what was provided in newly shared case data.

Tools

To help analysts share their insights about cases, all participants were encouraged to use the annotation tool, the organizing worksheet tool, and the summarizing report tool.

Annotation tool. All documents were in a main folder on a website for annotating documents (http://a.nnotate.com/). Participants created annotations and highlighted text directly on their documents (as shown in Figure 1).
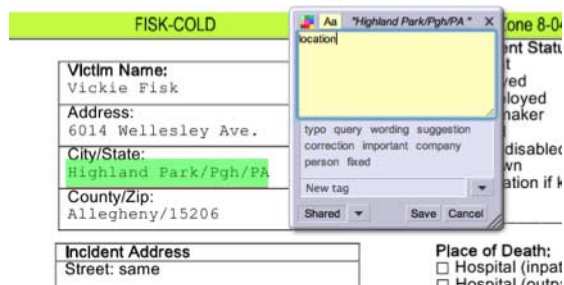

Figure 1. A screenshot of the annotation interface

Organizing worksheet tool. We used a "modus operandi" (MO) worksheet to help analysts organize evidence and clues in different categories. The worksheet listed all eight cases and provided categories for organizing evidence. As shown in Figure 2, participants entered data into this document and, optionally, added new rows in the worksheet to connect evidence across cases.

Figure 2. Organizing worksheet (italic added by a participant)

Report Tool. This tool was a simple form for conveying a progress report halfway through analysis. The progress report asked participants to provide an open-ended summary of their progress. The instructions reminded analysts their goal was to solve as many cases as possible and, in the report to "identify suspects and clues" (Figure 3).

Procedure

Each experimental session began with two practice cases as used in Balakrishnan et al. [2]. Next, the experimenter gave instructions for using the information sharing tools. Participants were instructed to use the tools to help them
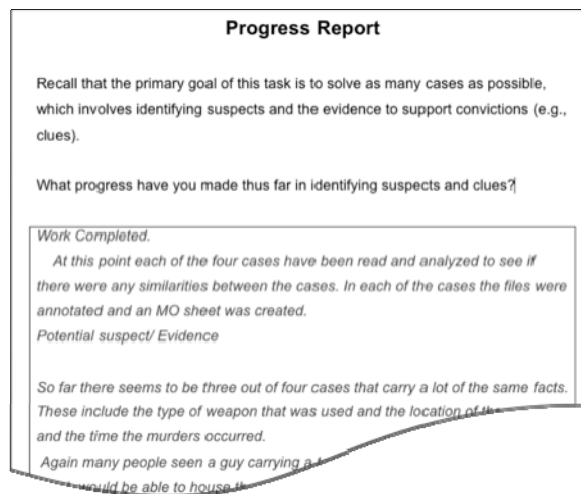

Figure 3. A progress report (italics added by a participant)

solve the cases and to communicate their ideas to their teammates, with whom they would share cases at the mid-point.

After training and instructions, participants were moved into two separate rooms. Each experimental session was split into two equal phases. During the first phase, participants worked alone on half of the total cases, and used the annotation tool and organizing worksheet tool to conduct their analysis. At the end of this phase, participants were asked to write a progress report.

At the midpoint, final analysts received the four crime case documents that initial analysts had worked on and either (1) the initial analyst's annotations, (2) the initial analyst's organizing worksheet, (3) the initial analyst's mid-point progress report, or (4) no information in any tool from the initial analyst (control condition). Crime documents and annotations were shared via annotate.com. MO worksheets and progress reports were saved as PDF documents by an experimenter and were shared with participants via Dropbox (www.dropbox.com).

The final analysts then worked on all eight cases with access to all crime documents. This period ended with final analysts writing a final report that identified suspects and evidence. Finally, the participants completed an online survey to assess analyst evaluations of their teammate.

Measures

Information accuracy was assessed with a proxy measure obtained from coding the initial analysts' midpoint progress reports. If the initial analysts identified the serial killer in their progress reports, the information they provided to the final analysts (in their annotations, worksheet, or report) was considered to be accurate. On the contrary, if they identified no suspect or an incorrect suspect as the serial killer in their progress reports, the information they provided to the receivers was considered to be inaccurate.

We observed more cases of final analysts receiving inaccurate than accurate information, particularly in annotation and the worksheet tool conditions. These results are shown in Table 1, below.

	Accurate information	Inaccurate information
Annotation tool	5	10
MO worksheet tool	5	10
Report summary tool	8	7
No tool control	14	

N = 45 in six cells of 2x3 design, 59 in all seven cells

Table 1. Experiment 1 information accuracy in each condition

Problem-solving performance was measured as a binary variable that reflected whether the final analyst solved the serial killer crime, evidenced by a final report correctly identifying by name the perpetrator of the four serial killer homicide crimes.

Collaborator evaluation was measured with a continuous variable obtained from the post-analysis survey in which final analyst teammates rated how much the initial analyst teammate had helped them solve the homicide cases they received at the midpoint. On the survey, they used a 5-point scale ranging from 1 ("not helpful at all") to 5 ("very helpful") to evaluate their teammate's helpfulness on each homicide case. We averaged these ratings across the three serial killer cases provided by the initial analyst teammates to form a composite rating of collaborator helpfulness on the serial killer problem.

Results

Our statistical analysis used a 2 (information accuracy: accurate, inaccurate) × 3 (information shared in tool: annotation, organizing worksheet, or summary report) design with an additional control group (no information shared in a tool). As noted below, we considered the control group to be a benchmark condition against which to evaluate the efficacy of information shared in a tool.

Problem-solving performance

In Hypotheses 1a and 1b we predicted that final analysts who received accurate information in tools would do better than final analysts in the no tool control condition (H1a) and that final analysts who received inaccurate information in tools would do worse than the controls (H1b). Because the dependent variable is a discrete variable (1: solved serial killer crime, 0: did not solve serial killer crime), we conducted a logistic regression analysis that estimated the effect of receiving accurate information in a tool and the effect of receiving inaccurate information in a tool. Each of these situations was compared to the benchmark no tool
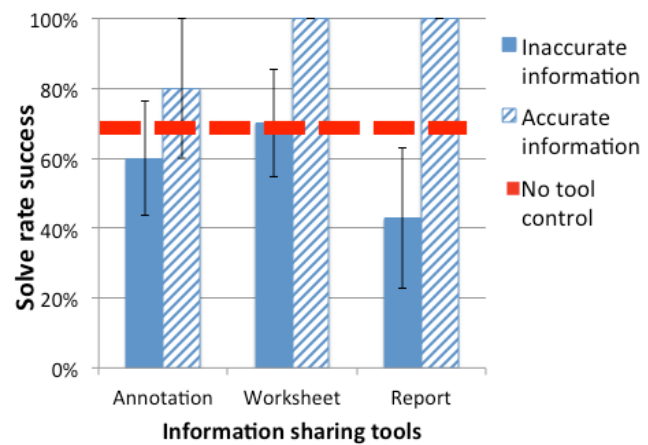

Figure 4. Experiment 1 solve rate success when teammate provides accurate or inaccurate information using one of three tools, compared with the no tool condition.

control, which served as the statistical reference group in the analysis.

As shown in Figure 4, the analysis revealed a significant main effect of information accuracy (X^2 [n=59] = 8.00, p = .02). As predicted in Hypothesis 1a, final analysts performed better when they received accurate information in a tool (i.e., annotations, worksheet, or report) from the initial analyst than when they received no information in a tool. The odds of solving the serial killer crime were 4.30 times greater for analysts receiving accurate information than for analysts in the control condition (X^2 [n=59] = 6.32, p = .01). As predicted in Hypothesis 1b, analysts performed worse when they received inaccurate information in a tool than when they received no information in a tool. The odds of solving the crime were .37 times lower for analysts receiving inaccurate information than for their counterparts in the control condition (X^2 [n=59] = 5.30, p = .02).

Hypothesis 2 posited that the effect of information accuracy would be most pronounced when the information is received in a progress report because this tool conveys the highest level of sensemaking. We used nested-by-values effects coding to estimate an additional logistic regression (see [32; 34, p. 391]). This supplementary analysis allows use of all the observations in the 2 (information quality) x 3 (information tools) cells, but alters the parameterization with one parameter for each of the three tools conditional on information accuracy. This analysis reveals that the odds of solving the case were 7.30 times greater for analysts receiving a report tool containing accurate rather than inaccurate information (X^2 [n=45] = 7.52, p = .006), but the odds of solving the case did not differ as a function of information accuracy for annotation (X^2 [n=45] = 1.68, *ns*) or worksheet tools (X^2 [n=45] = 1.68, *ns*).

Collaborator evaluation

Those in the control condition who received no information in tools rated their teammate comparatively poorly (M = 1.67, SD = 1.76) when contrasted with those who did

receive information in tools (M = 3.07, SD = 1.45, t [57] = 3.02, p = .004). However, the 2 (information accuracy) x 3 (information sharing tool) ANOVA yielded just one significant main effect for the type of information sharing tool (F [2, 41] = 6.14, p = .005). Final analysts evaluated the collaborator more positively when they received information conveyed in the summary report tool (M = 4.07, SD = 1.22) than when they received information conveyed in an organizing worksheet tool (M = 2.67, SD = 1.43) or an annotation tool (M = 2.49, SD = 1.19). There was no overall accuracy effect.

Because significant problem solving differences due to information accuracy were found only for the report tool, we conducted an ANOVA of teammate evaluations comparing the no tool control condition with the accurate and inaccurate summary report conditions. The analysis showed an overall effect (F [2, 26] = 12.16, p < .0002). A planned contrast comparing the evaluations of the accurate teammate (M = 4.75, SD = .46) with the evaluations of inaccurate teammate (M = 3.29, SD = 1.38) showed a marginally more favorable evaluation of the accurate teammate (F [1, 26] = 3.92, p = .059), and the planned contrast comparing the inaccurate teammate with the control teammate (M = 1.67, SD = 1.76) showed a significantly more favorable evaluation of the inaccurate teammate (F [1, 26] = 5.98, p = .02).

In sum, analysts evaluated asynchronous teammates more positively when information was conveyed in a report tool, and they evaluated these teammates more positively than teammates who provided no report even when they received inaccurate information in the report and performed worse than those receiving no additional information. Consistent with this finding, we did not find a significant association between analysts' problem-solving performance and their evaluations of their teammate (r = -.28, $n.s.$).

Discussion of Experiment 1
Our results suggest that analysts may be misled by inaccurate sensemaking information provided in a high-level information-sharing collaboration tool. Overall, final analysts receiving accurate information in a tool (annotation, worksheet, or report) performed better than their counterparts who received no sensemaking information from analyst teammates (hypothesis 1a). This increment in performance suggests that sharing accurate information can be helpful. By contrast (hypothesis 1b), final analysts receiving inaccurate information in a tool (annotation, worksheet, or report) performed worse than their counterparts who received no sensemaking information from analyst teammates.

The decrement in performance due to inaccurate information was driven primarily by the result of information sharing in the progress report condition (hypothesis 2). We found limited support for hypothesis 2. As shown in Figure 4, analysts performed better when they

received accurate rather than inaccurate information, but the increment was only statistically significant for the report tool. The lack of statistical significance for annotations and worksheets cannot be interpreted to suggest that the accuracy of the information conveyed in these tools does not affect asynchronous collaborative analysis. It is possible, for example, that the smaller sample of accurate than inaccurate information made it difficult to detect differences. Nonetheless, the results do suggest that receipt of accurate information in a summarizing report tool can help analysts, whereas receipt of inaccurate information in a report tool can harm analysts.

We also discovered that analysts receiving sensemaking information in tools (annotations, worksheets, or reports) evaluated their collaborator more positively than analysts receiving no information in tools. This positivity was the case even for analysts who not only received inaccurate information in the report tool but also performed worse than analysts receiving no information in tools. These results showing inflated evaluations of a teammate who is associated with a decrement in performance raises a question about how accurately asynchronous analysts can evaluate one another's contributions.

A major limitation of this study is that we measured rather than manipulated the quality of the information conveyed in the information sharing tools. Given the complexity of the task, fewer than half of the initial analysts had solved the case by the midpoint, which resulted in fewer observations of accurate than inaccurate information sharing. This imbalance was not an issue in the summary report tool condition, but the small sample size (8 accurate reports vs. 7 inaccurate reports) warrants replication.

EXPERIMENT 2
The main purpose of Experiment 2 was to replicate and extend Experiment 1. Because the effects were greatest in the summary report tool conditions of Experiment 1, we examined the effects of the accuracy of the information conveyed in a summary report. We addressed a key limitation of Experiment 1 by manipulating the accuracy of the information conveyed in the reports. We also increased the sample size twofold.

We extended our examination to an asynchronous collaboration with more than one other teammate and source of information because this structure is increasingly common (e.g., [12]). We increased the team size to three. Final analysts received either two reports from their teammates that conveyed inaccurate information (inaccurate-inaccurate report condition) or they received one report conveying inaccurate information and another report conveying accurate information (inaccurate-accurate report condition). This design allowed for replication of Experiment 1 in a team of three and an examination of whether an accurate report could offset the detrimental effects of an inaccurate report.

Method
Participants and design
We recruited 38 native English-speaking participants for the experiment. All participants were undergraduate or graduate students. Each participant was paid $10 for participating in the experiment for 1 hour.

All participants took on the role of final analyst. As in Experiment 1, they could not solve the case without case data from teammates. Participants were randomly assigned to either the inaccurate-inaccurate report condition or the inaccurate-accurate report condition, resulting in 19 observations per condition.

Task
We used the same experimental homicide detective task involving a serial killer as in Experiment 1 with a few modifications. The procedure was altered by simulating the first phase of the sequentially interdependent collaboration with initial analyst reports compiled from progress reports provided by participants in Experiment 1. Participants were led to believe that two teammates had been working on the cases during an initial phase and had each prepared a progress report. With access to these two initial analyst progress reports and all underlying case data, participants began their analysis. They were instructed, as in Experiment 1, to solve as many of the cases as possible and to write a summary report on behalf of their team.

Procedure
Each experimental session lasted approximately one hour and followed a procedure similar to Experiment 1 with the exception that participants did not go through an initial phase before they were provided with teammates' sensemaking information conveyed in two summary reports.

We manipulated information accuracy. Participants were randomly assigned to one of the information accuracy conditions. To create the inaccurate-inaccurate and inaccurate-accurate report conditions, we compiled three reports of equal length (214 words) from the progress reports provided by participants in Experiment 1. Each report conveyed the third level of sensemaking information (e.g., suspect hypotheses) about three of the four serial killer cases and one unrelated case. Two of these progress reports were inaccurate, mentioning different, wrong suspects and an unrelated case suspect, and one report was accurate, mentioning the serial killer and an unrelated case suspect.

Measures
The measures of problem-solving performance and collaborator evaluation were the same as in Experiment 1.

Results
Problem-solving performance
As seen in Figure 5, analysts receiving an inaccurate report and an accurate report performed better than their counterparts whose teammates provided them with two
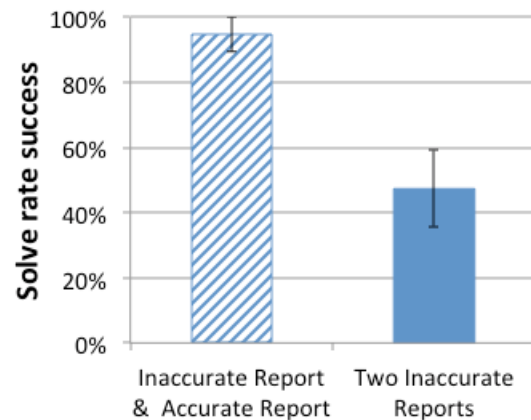

Figure 5. Experiment 2 solve rate success when teammates provide analysts with an inaccurate and an accurate report or two inaccurate reports.

different inaccurate reports (X^2 = 10.36, p = .001). Whereas the vast majority of the analysts receiving an inaccurate and an accurate report solved the case (94.7%), only half of the analysts receiving two inaccurate reports solved the serial killer case (47.4%). The latter result is comparable to the performance of analysts receiving one inaccurate report in Experiment 1.

Collaborator evaluation
Analysts in the inaccurate-accurate report condition evaluated their teammate who provided an accurate report (M = 4.08, SD = .80) as significantly more helpful than their other teammate who provided an inaccurate report (M = 2.39, SD = 1.08, t (16) = 4.42, p = .0004). Analysts in the inaccurate-inaccurate report condition whose teammates provided them with two different inaccurate reports did not evaluate one teammate as more helpful than another (Ms= 3.08 and 3.13, SDs = 1.08 and 1.00, t (15) = 0.27, ns). However, analysts receiving two inaccurate reports rated their two teammates, on average, just as positively (M = 3.10, SD = .99) as did analysts receiving an inaccurate report and an accurate report (M = 3.24, SD = .50, t (31) = -.48, ns) even though the former were half as likely to solve the case as the latter.

Discussion of Experiment 2
Experiment 2 accomplished its primary goals of replicating and extending Experiment 1. First, it confirmed, in a larger sample, that inaccurate sensemaking information from teammates conveyed in a report tool has a detrimental effect on problem-solving performance, and that analysts erroneously evaluated these collaborations as helpful. The manipulation of information accuracy establishes that these findings are independent of other characteristics of reports or teammates. (In Experiment 1, information accuracy was measured rather than manipulated so we could not separate accuracy from other aspects of reports or teammates, such as writing style.) Second, we extended the study to a three-person team, and showed that accurate information from one teammate can cancel out inaccurate information from

another. This finding suggests the value of having more than one teammate. A third teammate increases coordination costs but also increases the chances of receiving accurate sensemaking information and may counteract the tendency for asynchronous collaborative analysts to accept a teammate's information, even when it is inaccurate.

LIMITATIONS

Some limitations of our studies are worth noting before we discuss their implications. Although our participants underwent a thirty-minute training session on collaborative analysis, they were not experienced detectives or intelligence analysts. Experts in these fields may be better able to discern the quality of their collaborators' contributions than were our naïve participants. Another limitation is that we compared a small set of text-based information sharing tools. We also focused on small team collaborations rather than on larger-scale analysis situations. In future work, it would be worthwhile to expand the populations, tools, and groups tested.

IMPLICATIONS FOR COLLABORATIVE SENSEMAKING

When they function effectively, teams of collaborative analysts can be especially useful for solving analysis problems because they leverage the insights of teammates [9]. When they function poorly, collaborative performance is even worse than when analysts work alone [35].

Our research simulated information sharing in the context of remote asynchronous collaboration—a collaboration mode increasingly used by teams facing complex problems [4]. In two experiments, we examined whether participants could identify a serial killer from unsolved homicides after they received raw case data and, in all conditions except the Experiment 1 no tool control condition, additional sensemaking information from a teammate. These experiments help to identify the benefits and costs of asynchronously sharing sensemaking information in a controlled setting.

Experiment 1 measured information accuracy and compared an annotation tool meant to help analysts share evidence attached to a problem (sensemaking level 1), an organizing worksheet tool meant to help analysts share interconnections among categories of evidence (sensemaking level 2), and a summarizing report tool meant to help analysts share new associations and hypotheses (sensemaking level 3). Experiment 2 replicated and extended Experiment 1 with a manipulation of the accuracy of the information conveyed in two summarizing reports.

Taken together, results suggest that the degree to which information sharing tools help collaborative analysis depends on the level and the accuracy of the sensemaking information they convey, as follows:

1. Sensemaking information from a teammate at the third level (summarizing report containing hypotheses and/or conclusions) was associated with significant differences in problem-solving performance, either positive if the information was accurate, or negative if the information was inaccurate (Experiment 1; replicated in Experiment 2).

2. Analysts evaluated information sharing collaborators, especially those conveying the third level of sensemaking information, as helpful even when they were not (Experiment 1; replicated in Experiment 2).

3. Accurate sensemaking information at the third level from an additional teammate mitigated the negative effects of inaccurate sensemaking information at the third level from another teammate (Experiment 2).

These findings contribute to an understanding of the kind of information sharing that underlies effective asynchronous collaborative analysis. They expose a problem we call *teammate inaccuracy blindness* whereby people take a teammate's contributions at face value without looking closely at their accuracy. Although previous sensemaking literature has addressed issues that may arise from an analyst's own confirmation bias [25, 28], only recently have researchers considered the accuracy of sensemaking information contributed by others [12, 36]. Our work highlights how challenging it can be for analysts to recognize when utilizing teammates' sensemaking information at the highest level might harm, rather than aid, their problem-solving performance. These results have implications for the many situations in which collaborators share high-level sensemaking information that varies in accuracy.

Similarities exist between teammate inaccuracy blindness and other group phenomena, such as in-group favoritism [1] and groupthink [14]. Although in each case group members make poor use of information provided by others, there are some differences. Groupthink and in-group favoritism stem from group identity and conformity to the group, and predict globally positive evaluations of insiders. These phenomena harm decision quality, in large part, because group members reject accurate information from outsiders [1, 14]. By contrast, our notion of teammate inaccuracy blindness stems from misguided cognitive efficiency; it harms problem-solving performance because people, perhaps mindlessly, incorporate inaccurate sensemaking information from teammates. (A related phenomenon, called "social recognition memory," is documented in forensic analysis and false eyewitness testimony [38].) Adopting sensemaking information at the highest level is cognitively efficient. As a result, teammate inaccuracy blindness occurs most at this level, and analysts evaluate teammates providing this level of sensemaking information more positively than teammates sharing lower level or no sensemaking information at all.

It is noteworthy we uncovered teammate inaccuracy blindness in the context of teammates sharing data and information critical to solving a complex analysis task with a correct solution. Research indicates that it is rare for

solutions to be so highly demonstrably that "truth wins" [19]. Demonstrability can be limited, in part, because people tend to lack sufficient information and cognitive resources to distinguish inaccurate from accurate information. These limiting factors provide insights into why teammate inaccuracy blindness occurred for participants receiving two different inaccurate reports but not for participants receiving an inaccurate and an accurate report—the accurate report would have directed attention to the serial killer and caused the participant to see the pattern revealed by his behavior. This logic suggests that teammate inaccuracy blindness is likely to be more prevalent in lower demonstrability situations characterized by greater problem complexity and information overload.

OVERCOMING TEAMMATE INACCURACY BLINDNESS

Even given the limitations of our studies, our results provide some insights for avoiding or mitigating teammate inaccuracy blindness. In the ideal case when teammates have developed accurate hypotheses or conclusions, it seems advisable to use information sharing tools, especially those that convey higher-level representations such as summary reports. However, when teammates lack expertise or the solution is not very demonstrable, high-level sensemaking tools could be harmful for subsequent analysts, especially because analysts tend to perceive their teammates as helpful even when, objectively, they are not.

Experiment 2 and research in CSCW and organizational science suggest some strategies for overcoming teammate inaccuracy blindness. Experiment 2 showed that adding an accurate report to an inaccurate report raised performance. This finding suggests the value of tools that facilitate identifying and adding competent asynchronous teammates through organizational expertise directories [7], work logs to identify structurally similar others [30], or reputation-rating systems for collaborators [9].

One study of collaborative peer production systems shows that displaying controversy or inconsistent contributions tends to increase accuracy [36]. To the extent these findings generalize to asynchronous collaborative analysis teams, software might be used to evaluate the overlap in teammates sensemaking information and display the similarity or dissimilarity of their contributions. It is possible that a simple indicator of a multiplicity of views could cause receivers to put more mindful thought into their judgments and to avoid teammate inaccuracy blindness.

CONCLUSION

Tools for summarizing information can assist analysts in their sensemaking process, but the value of these tools for sharing information will be limited by the accuracy of teammates' information. Accurate information improves performance but inaccurate information can cause performance to fall below the level obtained when no information is shared. This work uncovers teammate inaccuracy blindness and identifies a few ways to overcome it.

REFERENCES

1. Abrams, D., M. Wetherell, S. Cochrane, M.A. Hogg, J.C. Turner. 1990. Knowing what to think by knowing who you are: Self-categorization and the nature of norm formation, conformity and group polarization. *British Journal of Social Psychology* 29, (1990), 97-119.

2. Balakrishnan, A.D., Fussell, S.R., and Kiesler, S. Do visualizations improve synchronous remote collaboration? In *Proc. of CHI 2008*, (2008), 1227–1236.

3. Bier, E.A., Card, S.K., and Bodnar, J.W. Entity-based collaboration tools for intelligence analysis. *Visual Analytics Science and Technology, 2008. VAST'08. IEEE Symposium*, (2008), 99–106.

4. Bruns, H. Working alone together. Coordination in collaboration across domains of expertise. *Academy of Management Journal 56,* 1 (2013), 62-83.

5. Chase, W. G., & Simon, H. A. Perception in chess. *Cognitive Psychology*, 4 (1973), 55-81.

6. Chung, H., Yang, S., Massjouni, N., Andrews, C., Kanna, R., and North, C. VizCept: Supporting synchronous collaboration for constructing visualizations in intelligence analysis. *Visual Analytics Science and Technology (VAST), 2010 IEEE Symposium on, IEEE* (2010), 107–114.

7. Cross, R., & Baird, L. Technology is not enough: Improving performance by building organizational memory. *Sloan Management Review* (Spring), (2000), 69-78.

8. Fisher, K., Counts, S., Kittur, A., and Ave, F. Distributed sensemaking : Improving sensemaking by leveraging the efforts of previous users. *Proc. of CHI 2012*, ACM Press (2012), 247-256.

9. Furtadoa, V., Ayresa, L, de Oliveirac, M., Vasconcelosb, E., Caminhaa, C., D'Orleansa, J., and Belchiora, M. Collective intelligence in law enforcement – The WikiCrimes system. *Information Sciences 180*, 1 (2010), 4-17.

10. Goyal, N., Leshed, G., and Fussell, S. Effects of visualization and note-taking on sensemaking and analysis. In *Proc. of CHI 2013*. ACM (2013), 2721-2724.

11. Hackman, J.R. *Collaborative intelligence: Using teams to solve hard problems*. Berrett-Koehler Publishers, San Francisco, (2011).

12. Hansen, D. L., Schone, P. J., Corey, D., Reid, M., & Gehring, J. Quality control mechanisms for crowdsourcing: peer review, arbitration, & expertise at

familysearch indexing. In *Proc. CSCW 2013*, ACM Press (2013), 649-660.

13. Heer, J., Viégas, F.B., and Wattenberg, M. Voyagers and voyeurs: Supporting asynchronous collaborative visualization. *Communications of the ACM 52*, 1 (2009), 87–97.

14. Janis, I. L. *Victims of groupthink*. Boston: Houghton Mifflin, (1972).

15. Kane, A. A., Argote, L., & Levine, J. M. (2005). Knowledge transfer between groups via personnel rotation: Effects of social identity and knowledge quality. *Organizational Behavior And Human Decision Processes 96*, 1 (2005), 56-71.

16. Kang, R. and Kiesler, S. Do collaborators ' annotations help or hurt asynchronous analysis? In *Proc. CSCW 2012*, ACM Press (2012), 123–126.

17. Knoblich, G., Ohlsson, S., and Haider, H. Constraint relaxation and chunk decomposition in insight problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition 25*, (1999), 1534-1555.

18. Laughlin, P.R., B.L. Bonner, A.G. Miner, P.J. Carnevale. Frames of reference in quantity estimations by groups and individuals. *Organanizatioal Behavior and Human Decision Processes 80*, 2 (1999), 103-117.

19. Laughlin, P.R., A.L. Ellis. Demonstrability and social combination processes on mathematical intellective tasks. *Journal of Experimental Social Psychology 22*, 3 (1986), 177-189.

20. Lojeski, K.S., Reilly, R., and Dominick, P. The role of virtual distance in innovation and success. *Proc. Conf on System Sciences (HICSS'06),* (2006), 25c–25c.

21. McCaffrey, T. Innovation relies on the obscure: A key to overcoming the classic problem of functional fixedness. *Psychological Science 23*, 3 (2012), 215-218.

22. Mednick, S. The associative basis of the creative process. Psychological Review 69, 3 (1962), 220.

23. Mesmer-Magnus, J.R. and DeChurch, L.A. Information sharing and team performance: A meta-analysis. *Journal of Applied Psychology 94*, 2 (2009), 535.

24. Munro, A.J., Höök, K., and Benyon, D.R. *Personal and social navigation of information space*. Springer-Verlag London, (1999).

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

26. Nobarany, S., Haraty, M., & Fisher, B. Facilitating the reuse process in distributed collaboration: a distributed

cognition approach. In *Proc. of CSCW 2012*, ACM Press (2012), 1223-1232.

27. Pioch, N.J. and Everett, J.O. POLESTAR: collaborative knowledge management and sensemaking tools for intelligence analysts. *Proceedings of the 15th ACM international conference on Information and knowledge management*, ACM (2006), 513–521.

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

29. Pirolli, P. and Russell, D.M. Introduction to special issue on sensemaking. *Human Computer Interaction*, (2011), 1–8.

30. Reagans, R., & McEvily, B. Network structure and knowledge transfer: The effects of cohesion and range. *Administrative Science Quarterly 48*, 2 (2003), 240-267.

31. Ren, Y. and Argote, L. Transactive memory systems 1985–2010 : An integrative framework of key dimensions, antecedents, and consequences. *Academy of Management Annals 5*, 1 (2011), 189–229.

32. SAS Knowledge Base Two-Factor Logistic Model Using Dummy and Effects Coding http://support.sas.com/kb/24/447.html#ex3

33. Stasko, J., Goerg, C., and Liu, Z. Jigsaw: Supporting investigative analysis through interactive visualization. *Information Visualization 7*, 2 (2008), 118–132.

34. Stokes, M.E., David, C. S., Koch, G. G. *Categorical data analysis using the SAS system* (2nd Ed). Cary, NC: SAS.

35. Tindale, R. S., Smith, C. M., Dykema-Engblade, A., and Kluwe, K. Good and bad group performance: Same processes--different outcomes. *Group Processes & Intergroup Relations 15*, (2012), 603–618.

36. Towne, W. B., Kittur, A., Kinnaird, P., & Herbsleb, J. Your process is showing: controversy management and perceived quality in wikipedia. In *Proc. of CSCW 2013*, ACM Press (2013), 1059-1068.

37. Weng, C. and Gennari, J.H. Asynchronous collaborative writing through annotations. In *Proc. of CSCW 2004*, ACM (2004), 578–581.

38. Wright, D. B., & Mathews, S. A., & Skagerberg, E.M. Social recognition memory: The effect of other people's responses for previously seen and unseen items. *Journal of Experimental Psychology: Applied 11*, 3 (2005), 200–209.