Establishing Collaborative Option Awareness during Crisis Management

J. L. Drury
The MITRE Corporation, Bedford, MA
jldrury@mitre.org

M. S. Pfaff
Indiana University, Indianapolis, IN
mpfaff@iupui.edu

G. L. Klein
The MITRE Corporation, McLean, VA
gklein@mitre.org

S. O. Entezari
Indiana University, Indianapolis, IN
sentezar@iupui.edu

ABSTRACT

This paper presents empirical results of the use of a novel decision support prototype for emergency response situations, which was designed to enhance the understanding of the relative desirability of one potential course of action versus another. We have termed this understanding "option awareness." In particular, this paper describes the process employed by pairs of experiment participants while performing emergency responder roles using different types of "decision space" visualizations to help them collaborate on decisions. We examined the decision making process via a detailed analysis of the communication between the cooperating team members. The results yield implications for design approaches for visualizing option awareness.

Keywords

Decision making, robust decision making, option awareness, decision space, situation space, chat.

1. INTRODUCTION

Crisis management requires inter-organizational teams of decision makers to take a number of actions in rapid succession. They must assess the situation, determine the available options, choose an acceptable option, and implement it quickly enough to keep the emergency conditions from worsening. Response to a major earthquake might require decisions that are coordinated among local/regional first responders, hospital adminstrators, political leaders, and aid agencies such as the Red Cross. Even a single, major fire could require police, fire, and hospital administrators to cooperate on choosing and executing joint responses.

Gary A. Klein (1999) has extensively studied how emergency responders make decisions. He found that experienced decision makers build up a set of patterns or templates, match situational cues to the relevant template, select a decision option that previously resulted in a positive outcome, then run a mental simulation to ascertain whether the proposed option will likely be successful in the current situation. This "recognition-primed decision making" (Klein 1999) can occur quickly if the current situation can be related to prior experience and the mental simulation is feasible to run in the decision maker's head. Weick (2001) also notes the role of prior experience in "sensemaking," which aids in enabling decision makers to determine options based on experience.

Some disaster response efforts are so unique, such as Hurricane Katrina and the 2011 Tohoku (Japan) earthquake and tsunami, that it is unreasonable to expect that even the most experienced emergency responders have sufficient relevant experience. But even if response personnel are able to match the situation to a template in these unique cases, running a mental simulation to determine an option's acceptability becomes unreasonable. Instead, "decision support systems can be used to reduce the time needed to make crucial decisions regarding task assignment and resource allocation" (Thompson et al. 2006, p. 250) in the context of complex disaster response efforts.

Many recent research efforts are directed at providing decision support for emergency response, such as RimSim (Campbell et al. 2008), Zographos and Androutsopoulos's (2008) integrated hazardous materials routing and emergency response decision support system, and the DIORAMA disaster management/resource allocation system (Kondaveti & Ganz 2009). There are also commercially-marketed decision support systems for emergency responders, such as CoBRA (Defense Group, Inc., Undated). These systems concentrate primarily on providing facts about the situation to crisis managers. This is what Hall et al. (2007) termed the *situation space*. Perceiving, comprehending, and projecting the situation space information into the future results in *situation awareness* (Endsley 1988).

In contrast, our work focuses on the *decision space*, which consists of information about the options that are available to address the situation. In the decision space, *option awareness* (Drury et al. 2009) is the understanding of how options compare to one another and the underlying factors that contribute to the plausible outcomes of choosing any particular option. While others have reported research on visualizations to aid decision making, in general these efforts have either focused on the situation space (e.g., Dong & Hayes 2012), or have less comprehensively addressed the decision space and the relationship between the two spaces (e.g., Hoffman 2006).

The ultimate goal of our work is to develop better team-based decision support systems appropriate for emergency responders and crisis managers. To best assist designing such systems for emergency response teams, however, we need to understand how people interact and coordinate their actions using situation and decision spaces. Do they concentrate more on the situation space or the decision space? Are they confused by differences in decision spaces designed to assess options for individual emergency services versus those designed to compare options for the team as a whole? We answer these and related questions via an analysis of communications among teams of experiment participants playing police and fire/rescue roles. Their role-playing communications shed light on the fundamentals of providing option awareness via decision spaces in tactical decision making situations.

We believe this paper contributes to the literature that provides guidance for designing decision support systems for emergency response teams. The emergency events in this experiment were created by team members with domain experience, often based on actual emergency events. A few of the participants in this experiment were emergency responders, and earlier experiments (e.g., Drury et al. 2009) had a larger percentage of emergency responders (16 out of 41 participants). We have not found a difference in performance (decision making speed, accuracy, and confidence) based on participants' domain experience in any experiment so far. In addition, this research is in exploratory stages as we investigate the fundamental properties of decision space visualization and the features that most significantly impact cognition. Therefore, much of what we are examining does not yet require substantial domain expertise. Yet, because we have tried to keep the fundamental characteristics of the experimental events consistent with those encountered in the field, the results should be particularly pertinent to the design of emergency response decision support systems.

2. RELATED WORK

Some previous research has examined providing information to the emergency responder about options and outcomes, although this was not referred to as a decision space. For example, the ARGOS Accident Reporting and Operational Guidance System for chemical, biological, radiological, and nuclear incidents provides predictions for the consequences of the events and possible courses of action (Prolog Development Center 2010). Bonazountas et al. (2007) developed a fire fighting decision support tool that shows the costs for undertaking different firefighting strategies. These examples focus on supporting an individual responder.

We started our research by developing decision spaces for individuals using the exploratory modeling approaches of Bankes (2003) to support *robust decision making*: decisions that may not be optimal under the most probable set of assumptions, but will be acceptable under the widest range of plausible conditions that may occur (Chandrasekaran & Goldman 2007). These decisions take a broad range of uncertainty into account to determine a course of action that will be robust. We found that providing visualizations of a decision space in an interactive interface could support decisions that were more normatively correct and made with more confidence than decisions made by experiment participants who did not receive the decision space information (Drury et al. 2009).

Based on the results of developing robust decision spaces for individuals, we began to explore decision spaces for teams of crisis managers. We developed a testbed to explore joint decision making for crisis response managers. A few team-based crisis management decision support systems were developed previously, such as EMPROV (Mendonca et al 2001), which included both situation and decision space elements. The work presented in the current paper is unique, however, because our approach uses exploratory modeling to provide visualizations of the decision space, which facilitates robust, team-based decision making under uncertainty.

Our studies have used performance-based metrics such as decision accuracy, speed, and confidence to provide the means of assessing whether decision space information is helpful. Analyzing communications between cooperating emergency responders can provide another means of assessing whether decision makers are considering information provided by the decision space.

Although communication among emergency responders is frequently over radios, such communication is ephemeral and does not lend itself to easy analysis. For this reason, others have analyzed user-generated content of text-based chat and social media in the context of crisis management. For example, Pfaff (2009) examined chat to determine the effects of mood and stress on within-team communication behaviors in a simulated crisis management environment. DeLongueville et al. (2009) analyzed tweets regarding an actual forest fire. Starbird et al. (2010) analyzed tweet content pertaining to the flooding of the Red River Valley in the US and Canada in March and April 2009. However, there have not been many analyses conducted of the *content* of chat posts or email messages in crisis situations. Rather than directly studying the content of postings, many studies, such as one that studied the social media-generated response to the Virginia Tech shootings (Palen et al. 2007), relied heavily on interviews and questionnaires to assess the phenomenon of social media in crisis situations.

This paper documents for the first time our analysis of the decision making process that occurred among computer-mediated team members while considering a collaborative decision space. Each team consisted of two experiment participants who were provided with a chat channel as their sole intercommunication mechanism. We analyzed the teams' chat conversations to identify the facts and options the participants jointly considered during the decision making process.

3. DECISION SPACE VISUALIZATION

Decision space visualization of options can be generated by computer-based forecasting models that can assess dozens of options with hundreds or thousands of variations due to uncertainty. Uncertainty arises when there are variables outside of the decision makers' control. For example, the chances for successfully dousing a fire with two fire trucks will be different if high winds arise to fan the fire's flames, or a drenching downpour occurs. Simulation models can run many "what if" variations to determine a landscape of plausible outcomes from choosing a particular option. The goal is to significantly reduce the cognitive load of conceiving and evaluating this large array of contingencies.

To aid the decision maker in comparing options, we devised a multi-attributed utility function (Keeney & Raiffa 1993) as an evaluative metric that is based on the costs incurred by each option. Consequences such as property damage, injury, and death are assigned monetary values (insurance actuarial tables can be used to assign a monetary value to death). The cost of each "what-if" case is computed by summing the execution-cost of enacting the option and the outcome-cost of consequences resulting from enacting the option, and any opportunity-costs or additional indirect-costs that might occur in the future due to having enacted the option.

Even for the same option, the costs vary depending on situational conditions beyond decision makers' control, such as whether fire trucks can respond quickly or traffic congestion delays their arrival. Thus there is a distribution of possible consequences for each option. Each distribution is a function of the uncertainty of the situation space (e.g., how big the fire actually is) and the uncertainty inherent in the decision option (e.g., what percent of fire trucks will get to the scene and when). Although an optimal plan would generate the highest expected return on investment, under deep uncertainty (Lempert et. al. 2003), where uncertainty is irreducible, optimal strategies lose their prescriptive value if they are sensitive to these uncertainties. In other words, selecting an optimal strategy is problematic when there are multiple plausible futures for each option. Instead, Chandresekaran (2005) and Chandresekaran & Goldman (2007) suggest shifting from seeking optimality to seeking *robustness*: acceptable outcomes across the broadest swath of plausible futures.

The results of the hundreds or thousands of cost evaluations for the many different plausible futures for each option can be summarized graphically by a box-plot (Tukey 1977). Although future research may determine a more nuanced visualization approach, box-plots provide a simple means of comparing the cost distributions of the options. We further simplified the box-plot visualization by eliminating the need to display so-called "outlier" data points.

Figure 1 shows the box-plots indicating the range of costs for each option in the example scenario. The top and bottom "whiskers" of the box-plot depict the maximum and minimum cost cases, respectively. The top and bottom sides of the box show the cost of the 75th and 25th percentile cases. The line bisecting the box indicates the median cost case. In Figure 1, the best choice is to send all three fire trucks to the scene because its box-plot has the lowest cost for the minimum case, the 25th percentile case (the lower bound of the box), and the median. Although its 75th percentile (the upper bound of the box) and maximum cases are the highest of any option, this option is still best for three out of five of the box-plot parameters and thus is the winner.

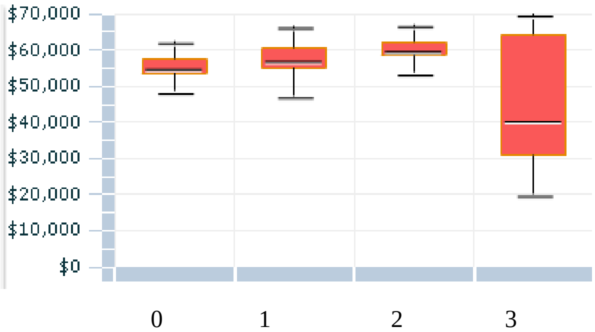

Figure 1. A decision space showing the relative costs of sending between 0 and 3 fire trucks to a fire

Note that the visualization of the options in Figure 1 is presented from an individual viewpoint of a single fire station, and assumes that no other responders will send assets to handle the emergency. But what if multiple fire stations cooperate, or police and fire plan a joint response?

The following example illustrates the often non-obvious synergy that can occur from such collaboration. Assume that a fire is reported when roads are clogged with people viewing an accident scene. An individual decision space for the police alone, assuming that the police are only concerned with the traffic incident, shows the most robust option as sending one squad car. Such an approach, however, ignores the possibility for the two emergency response departments to help each other. If the police department sends additional vehicles to clear traffic in favor of the fire trucks, the trucks can reach the fire more quickly. Consequently, the fire will be smaller upon the trucks' arrival, so fewer trucks would be needed to extinguish the blaze. Despite the need for more police cars, the total cost to the city would be lower when each police car is less costly to deploy than each fire truck, and since putting out the fire while it is still small results in less property damage. Thus, combined decision spaces that illustrate the synergy involved in a collaborative response show the most robust option for this scenario as sending two fire trucks (instead of three) and two police cars (instead of one).

Decision makers in the fire and police department would likely be unaware of the level of potential cost savings from this type of cooperation without being able to view a combined decision space. Although emergency responders frequently make tradeoffs in their heads, there are limits to human cognition when analyses involve many variables and high uncertainty, especially when decisions must be made quickly and under stress. The combined decision space can provide rapid understanding of the likely costs and consequences of collaborative decisions (Liu et al., 2011).

4. EXPERIMENT METHODOLOGY

This experiment employed a mixed design. For the between-subjects portion of the experiment, pairs of participants were divided into three groups with each receiving different visualizations for providing option awareness and decision support: one group was given both individual and combined decision spaces; another was given only individual decision spaces; and the third was given only combined decision spaces. For the within-subjects portion, all participants evaluated the same 20 emergency events, which varied in magnitude, impact, conflict level, type of resources allocated, and type of potential damage. These events were developed by project members with emergency response experience and were pre-tested prior to the experiment. Ten of these events required the pair of participants to allocate the same type of resource (fire, rescue, or police) from different stations to the same event: we call these *homogeneous* events. The other ten events were *heterogeneous* events that required the pair of participants to allocate different kinds of resources to two potentially related events within the same city. Five of the homogeneous events and five of the heterogeneous events pertained to injury or potential loss of life while the other events focused on potential property damage only.

Participants were recruited from undergraduate information technology courses at a large Midwestern university with 24 men and 6 women participating. Nineteen participants were between 18 and 24 years old, seven participants were between 25 and 31 years old, three participants fell into the 32 – 38 age group, and one person was between 53 and 59 years old. Sixteen participants reported being unfamiliar with box-plots prior to the experiment, whereas fourteen participants were familiar with box-plots.

Because we are investigating basic psychological responses to the decision space versus the situation space, domain expertise was not required, although three participants in this experiment did have some previous emergency response experience. Some of our earlier experiments had a larger percentage of participants with emergency response experience (e.g., Drury et al. 2009). However, in those experiments we did not find a difference in performance based on experience or the other cognitive traits that were assessed via post-evaluation survey instruments.

Participants first read a one-page introduction to the experiment and signed the informed consent form. They were then given a copy of a training manual and a list of frequently-asked questions. The manual included a step-by-step walkthrough of a practice event. This walkthrough and subsequent two additional practice events familiarized participants with the steps to complete the task, the user interface, level of detail provided by the event description, and guidelines for how to identify the best option for each emergency event. Participants received feedback during training.

Following the training, participants were presented with the 20 emergency events in a randomized order. For each event the participants saw a description of the current emergency and a log of previous related emergencies displayed in panel on the left half of the screen. A pop-up map of the area with emergency information was available by clicking on a title bar on the bottom of the page. This map highlighted one or more event-related locations and displayed the routes linking the emergency resource stations (e.g. police stations or fire/rescue stations) and the current emergency location.

Decision space information was displayed as sets of box-plots illustrating the six best-ranked options among all possible options in that decision space. Participants saw one set of box-plots being presented for each type of decision space in their experimental condition. Participants could visually sort their displayed options either by rank or by the quantity of resources to be sent. Participants were also free to choose an option that was not among the top six-ranked options.

Participants in the "individual only" condition saw their individual decision space as well as their partner's individual space side-by-side, but could not interact with their partner's space. The two individual spaces could be quite different because they each were calculated assuming that their partner will not take an action. Participants in the "combined" condition saw their own view of the combined decision space and a tab whereby they could see their partners' view of the combined decision space. The two views of the combined space always showed the same box-plots and only differed in appearance when the two partners chose to sort or weight the box-plot elements differently. Finally, participants in the "individual + combined" condition saw all of these sets of box-plots.

An embedded chat window on the right hand side of the display was set up to enable and record communication between the two partners. There was no other way for the partners to communicate. We erased previous chat text every time a new event was displayed. This approach reinforced the fact that the emergency events were all independent of one another instead of being parts of a continuously developing situation occurring sequentially in time. Each event required one decision from each of the two partners regarding how they would allocate their own resources.

Once participants submitted a decision, a pop-up screen appeared for them to confirm it. At this point they could return to the decision space tool, or they could continue by indicating their confidence in the decision with a slider control. If their partner had not yet made a decision, the participant saw a message asking him or her to wait for their partner's choice. Once the partner entered a decision, both participants were asked to rate their confidence in the *combined* decision, and then proceeded simultaneously to the next event.

After evaluating 20 emergency events, participants filled out a set of surveys assessing individual cognitive traits that previous research suggested would likely influence their interactions with the decision space tool. These traits included human-computer trust (HCT (Madsen & Gregor 2000)) and locus of control (ICI (Duttweiler 1984)). The survey also included basic demographic data on age and gender, and whether participants had prior experience with emergency management or reading box-plots. The experiment took approximately two hours to complete.

5. CHAT ANALYSIS AND RESULTS

We employed the Grounded Theory method of data analysis (Glaser & Strauss 1967) to develop the coding scheme. This starts with an open coding stage in which the data is grouped into categories based on references to common concepts. Categories are then refined so that all data can be coded in accordance with the categories.

While this method assumes that the data itself will suggest categories to the analyst, in practice the persons developing categories view the data through their particular knowledge and experience. We are aware that conversation analysis (Sacks et al. 1974) suggests examining the types of statements at the level of so-called "speech acts" (Searle 1969), such as requests or acknowledgements. Clark and Brennan (1991) suggest that communications be examined for evidence of grounding: indications that the communicator has not been misunderstood. However, we were interested in how participants thought about the resource allocation problem and came to a decision with the aid of situation and decision space information. Thus we examined chat postings to determine whether they pertained to the situation space or the decision space.

Coding followed Creswell's (2002) process for analyzing and interpreting qualitative data. Two coders independently examined the data and suggested categories, proposed definitions, and discussed them. At this stage, categories were added, dropped, or changed, examples were clarified, and categories were iteratively adjusted as the coders compared notes with one another to achieve a unified and clear definition for each category.

The final set of categories and their definitions are listed in Tables 1 and 2. Five categories address the situation space (Table 1) and eight categories describe decision space statements (Table 2). Both coders analyzed the same subset of the data (corresponding to approximately 20% of the data set). There was 93% inter-coder agreement across the 13 categories and an acceptable Cohen's (1960) Kappa (κ= .72). After attaining this level of agreement, in accordance with Landis & Koch (1977), coding was completed by one of the analysts.

The unit of analysis was the "event conversation": the series of chat statements that pertained to a single event and that led to a single decision (the number of resources to send to that event). There was a clean demarcation between one event conversation and the next, since the entire screen was blanked, including the chat pane,

Table 1. Situation Space Chat Categories

Code	Definition	Examples
S-S	**Severity**: A statement that makes a judgment regarding whether or not the event has a high magnitude (separate from the impact of the event's severity). Focuses on the size or quantity of the source(s) of the event and how localized it is.	"the fires are relatively small" "we have two groups that could be dangerous"
S-R	**Risk:** A statement that explicitly addresses the risk or probability of the event changing in size; e.g., uses the words "risk," "likely," "potential," or "chance" regarding the facts of the situation, but not the outcomes.	"it's likely that the fire in the bushes will jump to the café."
S-P	**Patterns:** A statement that comments on the relationships (or lack thereof) among the current and prior events described in the log.	"sounds like they are similar" "these all seem to be stemming from different issues"
S-D	**Distance:** A statement about how close a station is to the event, or a statement comparing the relative distances between the two stations and the event.	"I am almost 2X as far away"
S-M	**Map:** A statement that provides evidence that a participant is looking at or has used the map.	"I can not find bragg blvd"

Table 2. Decision Space Chat Categories

Code	Definition	Example
D-C	**Cost**: A statement that indicates awareness of the costs as described by the box-plots.	"if you look at overall the median cost is about a 1000 more and min cost is separated by almost 2000"
D-Re	**Reserves:** A statement that talks about keeping one or more resources in reserve in case of future emergencies.	"…saves me 1 to respond to other calls"
D-E	**Expending all resources:** A statement that explicitly mentions using *all* resources from one or both stations.	"I think we both need to send our max resources" or references to going "all out"
D-Ra	**Rank:** A statement mentioning the rank of one or more options, indicating that the participant is aware of the relative rankings of the options	"normal and min rank 1 is 5 cop cars"
D-Ri	**Risk:** A statement that uses words denoting uncertainty such as "risk," "likely," or "maybe," in the context of cost as expressed by the box-plots	"looks like either one of us sends one and takes a risk or we each send one sending the cost up more but less(en)ing the risk for max cost"
D-Sy	**Synergy:** A statement that recognizes or acknowledges the way that the two partners can help each other.	"I[t] will be necessary for me to control the traffic with my squad cars to open space for your firetrucks"
D-T	**Distrust:** A statement that calls into question the recommendations of the system itself.	"it's sayin send in 4 trucks; but I think that's too much a risk...."
D-Suf	**Sufficiency:** A statement that refers to a specific resource allotment as "enough" or "sufficient" for the scenario as opposed to the "correct" number to send.	"I am sending 2. that should be enough"

between events. Each event conversation could include one or several of the categories of statements listed in Tables 1 and 2.

The analysis showed that the conversation length per-event ranged from 15 to 1641 characters, with a very positively skewed distribution (*Median* = 226, *Mean* = 297.72, *SD* = 238.71). There was no significant variation in chat quantity between conditions or between homogenous and heterogeneous events.

The relative proportions of conversation dedicated to the situation space and decision space were compared to determine main effects and possible interactions between condition (combined decision space only, individual space only, and both spaces together) and between the homogenous and heterogeneous events. Due to the non-normal distribution of data, non-parametric tests were used. A Kruskal-Wallis analysis of variance revealed that participants with both the combined and individual spaces dedicated much more of their conversation to the decision space ($M = 0.70$, $SE = 0.04$) than the individual-only ($M = 0.55$, $SE = 0.03$) or combined-only ($M = 0.57$, $SE = 0.04$), $\chi^2(2) = 8.51$, $p < .05$. A Mann-Whitney U analysis of conversations showed that a significantly greater amount of decision space statements occurred for the heterogeneous events ($M = 0.67$, $SE = 0.03$) than the homogenous events ($M = 0.55$, $SE = 0.03$), $z = 3.15$, $p < .01$.

The analysis of differences between conditions was repeated, splitting out the homogenous and heterogeneous events, in order to detect any underlying interactions between the decision space visualizations and event type. The main effect of having both the individual and combined spaces is maintained across both event types (Homogenous: M = 0.66, SE = 0.05; Heterogeneous: M = 0.74, SE = 0.05), and the proportion of decision space statements are nearly identical when presented with the combined space only (Homogenous: M = 0.57, SE = 0.05; Heterogeneous: M = 0.58, SE = 0.05). However, there is a substantial difference between the way teams discussed and reasoned through each event type when presented with only the individual space. There is a significant drop in the proportion of decision space statements in this condition for homogeneous events (Homogenous: M = 0.66, SE = 0.05; Heterogeneous: M = 0.74, SE = 0.05), $\chi^2(2) = 8.51$, $p < .05$ (analysis of homogenous events). These results are shown in Figure 2.

Individual analysis of the 13 statement types for differences in the overall quantity of each type, again split between homogenous and heterogeneous events, showed significant differences for four statement types. For heterogeneous events, the only statement type showing a significant variation was Trust/Distrust of Rank (D-T; $\chi^2(2, N = 19) = 7.68$, $p < .05$).

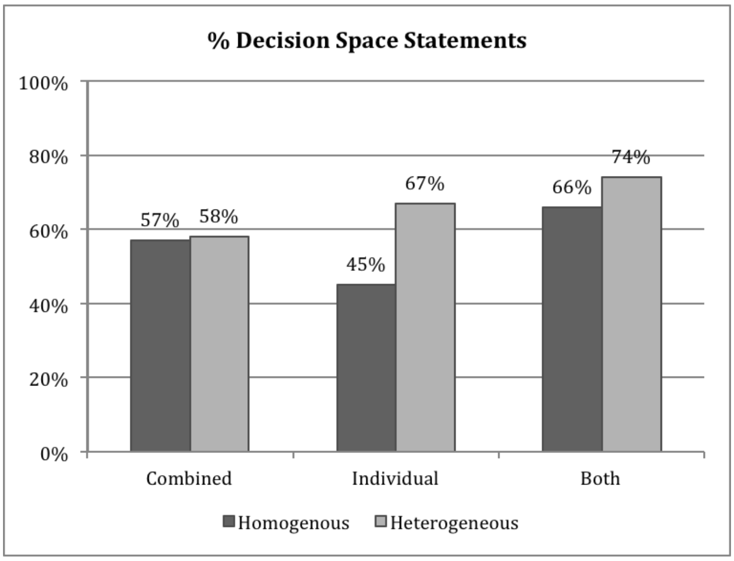

Figure 2. The proportion of statements dedicated to the decision space in each condition, split across homogenous and heterogeneous events.

For homogeneous events, three statement types showed significant variations: Severity (S-S; $\chi^2(2, N = 44) = 6.73$, $p < .05$), Distances (S-D; $\chi^2(2, N = 49) = 8.86$, $p < .05$), and Cost (D-C; $\chi^2(2, N = 29) = 9.17$, $p < .05$). These four results are shown in Table 3 and Figure 3.

Finally, pairwise correlations were performed across the 13 statement types to find whether any happened to co-occur frequently. A full sequence analysis of the conversations is beyond the scope of this paper, but frequently occurring statement pairs reveals trends in the nature of the teams' problem solving processes. Only three pairs stood out with Spearman's rank-order correlations greater than .25: S-M (Maps) and S-D (Distance), $r_s = .29$, $p < .0001$; D-Re (Maintaining Reserves) and S-P (Patterns), $r_s = .25$, $p < .0001$; and D-T (Trust/Distrust) and D-Ra (Ranks), $r_s = .30$, $p < .0001$.

Table 3. Counts of Statement Types Showing Significant Differences Between Conditions

Event Type	Statement Type	Combined-Only	Individual-Only	Both Combined and Individual
Heterogeneous	*D-T*	3	12	4
Homogenous	*S-S*	22	14	8
	S-D	10	26	13
	D-C	17	14	8

6. DISCUSSION

The goal of this analysis was to assess how the decision space visualizations and the event types influenced the nature of the conversations between partners, and particularly to understand the process by which they consider the facts of the situation and deliberate over their options for dealing with it. The results clearly show that the manner in which the decision space is presented to participants significantly changes how they process this information and problem-solve as a team.

Previous research analyses found that showing participants both individual and combined decision spaces appeared to introduce confusion, leading to degraded performance compared to those with only the combined display, but still better than those with only the individual display (Liu et al. 2011). In the current analysis, the significant increase in decision space-oriented communications in the condition with both decision spaces

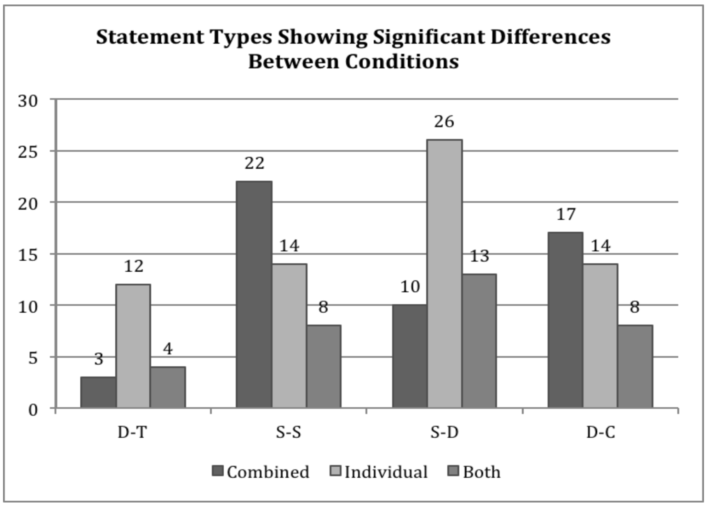

Figure 3. The counts of specific statement types showing significant changes across conditions. S-S (Severity), S-D (Distance), and D-C (Cost) are results for only the homogenous events; D-T (Distrust) is a result for only the heterogeneous events.

suggests that this apparent confusion spurred much conversation. Those in the other two conditions, for whom the conflict consequently was not explicitly apparent in the interface, discussed the decision space significantly less: about the same amount as they discussed the situation space.

The overall increase in decision space communication for heterogeneous events over homogenous events is not surprising given the substantial increase in cognitive complexity surrounding the potential synergy between the application of each team member's resources. This interpretation was also supported in previous research by the analysis of decision-making performance, with significantly increased performance for homogenous events (Liu et al. 2011). In the current analysis, in the individual-only condition, the proportion of decision space communication for homogenous events drops down to only 47% percent – much lower than the amounts for those with only the combined space (57%) and those with both spaces (66%). Assuming the increase in decision space communication for heterogeneous events in the latter two conditions is indeed attributable to the apparent increased complexity of the potential interactions between the team members' options, and that in contrast the participants with the individual-only display are least likely to be aware of those complexities, it makes sense that decision space communication would be lower under those circumstances.

The analysis of each statement type revealed more detail of what participants discussed and how that may have changed across conditions. The only significant change for heterogeneous events is the Distrust (D-T) category, in which participants mention something that calls into question the recommendations made by the system. Returning again to the complexity and conflict inherent in those events, it seems that participants are either not understanding or are reluctant to accept the synergistic solutions presented by the system, which contradict either their individual decision space (if they have one), or their intuitive assessment of the situation and the potential impact of each of their courses of action, or both. The significant correlation between D-T statements and D-Ra statements (regarding rankings) seems to confirm this.

Participants with only the individual decision space display in homogeneous events were far more likely to refer to distances than those with a combined decision space display. This result makes sense because in the homogenous events, the main considerations for participants in choosing a course of action are the relative distances from each participant's station and the event, as well as the relative number of their available resources. One potential conflict might be that the closest participant only has one or two resources to spare, while the more distant participant has several available. In resolving that conflict, it appears that without any kind of display explicitly showing the range of outcomes that would occur by the two working together, they focus on the most relevant parameters of the situation space. In this case, it appears to be the information regarding the time-to-target displayed in the interface. Interestingly, no corresponding increase is seen for these participants in discussing the map, or in discussing their available reserves.

The co-occurance of statements regarding maps and distance confirms that teams used the mapping part of the interface to understand the situation. However, the co-occurance of statements regarding patterns (which could be inferred from the map and the event description) and maintaining reserves reveals an interesting bridge between the situation-space and decision-space. This confirms that teams were detecting such patterns in the situation as indicative of the increased likelihood of future events and therefore taking them into account when deciding whether to reserve resources for future events.

The changes in the amount of chat postings regarding severity (S-S) and cost (D-C) show similar trends across the three conditions for homogenous events. With the combined-only display, participants discussed the severity of the event and costs displayed in the decision space visualization substantially more than those with the combined and individual displays. Complementing the combined display with the individual display has already been shown to decrease decision-making performance (Liu et al. 2011), and this change in communication behavior may help explain why. Event severity and the range of costs resulting from the various available courses of action are arguably among the most critical decision-making factors in this task. The individual display may be misleading participants into believing they know more about these factors than they really do, and therefore they have less reason to discuss it.

The results of the current research lead to two implications for designing decision spaces. First, this analysis confirmed our earlier results (Liu et al. 2011) that decision makers should be presented with the combined decision space, as opposed to both the individual and the combined spaces or the individual spaces alone. (See Liu et al. (2011) for additional performance-related results: the speed and accuracy of option choices and participants' confidence in their decisions.) Sometimes the content in the individual versus combined spaces is conflicting, leading to confusion and requiring extra discussion. Second, the chat analysis showed that there were times when the participants seemed unaware of the synergies that were possible between their two stations or services. Displays such as the combined decision space display that highlight synergistic effects of the combinations of resources could therefore be helpful.

7. FUTURE WORK

As noted in the introduction, because we have tried to keep the fundamental characteristics of the experimental events consistent with the characteristics encountered in the field, the results should be particularly pertinent to the design of emergency response decision support systems. Therefore, in the interest of bridging this research into practice, it would be worthwhile to increase the level of detail of the events and rerun the experiment solely with participants having emergency response or crisis management domain expertise.

Finally, future work should investigate whether the box-plot visualizations should be replaced with a visualization that is more appropriate for users performing emergency response or crisis management tasks. Performance measurements such as decision speed, accuracy, and participant's confidence, could determine relative differences in visualization efficacy. A chat analysis such as the one described in this paper could shed light on any differences that might be found in performance when using the alternative visualization approaches.

ACKNOWLEDGMENTS

We would like to thank Sung-Pil Moon and Yikun Liu for their work developing the testbed environments and assistance in running the experiment. We would also like to thank the following colleagues at Pennsylvania State U. for their earlier contributions to the project: Loretta Moore, Jon Becker, and David Hall. This work was partially funded by MITRE Corporation Innovation Project numbers 43MSR001-EA and 45MSR026-FA.

REFERENCES

1. Bankes, S. (1993). Exploratory modeling for policy analysis. *Operations Research*, 41, 3 (May – June), 435 – 449.
2. Bonazountas, M., Kallidromitou, D., Kassomenos, P., and Passas, N. (2007). A decision support system for managing forest fire casualties. *J. of Environmental Management*, 84 (2007), 412 – 418.
3. Chandrasekaran, B. (2005). *From optimal to robust COAs: Challenges in providing integrated decision support for simulation-based COA planning*. Columbus, OH: Laboratory for AI Research, Ohio State U.
4. Chandrasekaran, B. and Goldman, M. (2007). Exploring robustness of plans for simulation-based course of action planning. In *Proc. of the 2007 IEEE Symposium on Computational Intelligence in Multicriteria Decision Making,* Honolulu, HI, April 2007, 185 – 192.
5. Campbell, B., Mete, H. O., Furness, T., Weghorst, S., and Zabinsky, Z. (2008). Emergency response planning and training through interactive simulation and visualization with decision support. In *Proc. of the IEEE 2008 International Conference on Technologies for Homeland Security*. IEEE, Waltham, MA.
6. Clark, H. H. and Brennan, S. E. (1991). Grounding in Communication. In Resnick, L. B., Levine, J. M., and Teasley, S. D., *Perspectives on socially shared cognition*. Washington, D. C.: American Psychological Association, 127 – 149.
7. Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20, 37-46.
8. Creswell, J. W. (2002). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research.* Upper Saddle River, NJ: Merrill Prentice Hall.
9. Defense Group, Inc., undated. CoBRA: For the new breed of first responders. http://www.defensegroupinc.com/cobra/, accessed 22 February 2011.
10. Delongueville, B., Smith, R. S., and Luraschi, G. (2009). "OMG, from here, I can see the flames!": A use case of mining location based social networks to acquire spatio-temporal data on forest fires. In *Proc. of the 2009 International Workshop on Location Based Social Networks (LSBN '09),* November 2009.

11. Dong, X. and Hayes, C. C. (2012). Uncertainty visualizations: Helping decision makers to become more aware of uncertainty and its implications. *J. of Cognitive Engineering and Decision Making*, 6(1).
12. Drury, J. L., Klein, G. L., Pfaff, M. S., and More, L. D. (2009). Dynamic decision support for emergency responders. In *Proc. of the 2009 IEEE Technologies for Homeland Security Conference*, Waltham, MA.
13. Duttweiler, P. C. (1984). The Internal Control Index: A newly developed measure of locus of control, *Educational and Psychological Measurement,* 44, 209 – 221.
14. Endsley, M. R. (1988). Design and evaluation for situation awareness enhancement, In *Proc. of the Human Factors Society 32nd Annual Meetin*g, Santa Monica, CA, Human Factors Society.
15. Glaser, B. and Strauss, A. L. (1967). *The discovery of Grounded Theory: strategies for qualitative research*. Chicago: Aldine Publishing Co.
16. Hall, D. L., Hellar, B. and McNeese, M. (2007). Rethinking the data overload problem: Closing the gap between situation assessment and decision making. In *Proc. of the 2007 National Symposium on Sensor and Data Fusion (NSSDF) Military Sensing Symposia (MSS)*, McLean, VA.
17. Hoffman, J. R., Wilkes, M. S., Day, F. C., Bell, D. S., Higa, J. K. (2006). The roulette wheel: An aid to informed decision making. *PLoS Medicine*, 3(6): e137.
18. Keeney, R. and Raiffa, H. (1993). *Decisions with multiple objectives: Preferences and value tradeoffs.* Cambridge, UK: Cambridge University Press.
19. Klein, G. A. (1999). *Sources of power: How people make decisions*. Cambridge, MA: MIT Press.
20. Klein, G. L., Pfaff, M. and Drury, J. L. (2009). Supporting a robust decision space. In *Proc. of the 2009 AAAI Spring Symposium on Technosocial Predictive Analytics*, Stanford, CA.
21. Kondaveti, R. and Ganz, A. (2009). Decision support system for resource allocation in disaster management. In *Proceedings of the 31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS)*, Minneapolis, MN.
22. Landis, J. R. and Koch, G. G. (1977). Measuring nominal scale agreement for categorical data. *Biometrics*, 33, 159 – 174.
23. Lempert, R. J., Popper, S. W., and Bankes, S. C. (2003). *Shaping the next one hundred years: New methods for quantitative, long-term policy analysis*. Santa Monica, CA: RAND.
24. Liu, Y., Moon, S. P., Pfaff, M. S., Drury, J. L., and Klein, G. L. (2011). Collaborative option awareness for emergency response decision making. In *Proc. of the 8th International Conference on Information Systems for Crisis Response and Management (ISCRAM)*, Lisbon, Portugal.
25. Madsen, M., and Gregor, S. (2000). Measuring human-computer trust. *Proc. of the 11th Australasian Conference on Information Systems*, Brisbane, Australia, December 2000, 53-65.
26. Mendonca, D., Beroggi, G. E. G., and Wallace, W. A. (2001). Decision support for improvisation during emergency response operations. *Int. J. Emergency Management*, 1, 1, 30 – 39.
27. Palen, L., Vieweg, S., Sutton, J., Liu, S. B., and Hughes, A. (2007). Crisis informatics: Studying crisis in a networked world. In *Proc. of the Third International Conference on e-Social Science*, Ann Arbor, MI.
28. Pfaff, M. S. (2009). Emotional conversations in command and control: The impact of mood and stress on computer-mediated team communication behaviors. In *Proc. of the 53rd Annual Meeting of the Human Factors and Ergonomics Society*, San Antonio, TX, 414 – 418.
29. Phillips, J. L., Klein, G. A., and Sieck, W. R. (2008). Expertise in judgment and decision making: A case for training intuitive decision skills. In D. Koehler & N. Harvey (eds.), *Blackwell Handbook of Judgment and Decision Making*, pp. 297-315. Malden, MA: Blackwell Publishing Ltd.
30. Prolog Development Center (2010). ARGOS CBRN information system for emergency management. White paper version 0.8, June 2010. http://www.pdc.dk/Argos/downloads/ ARGOS_whitepaper.pdf, accessed 22 February 2011.
31. Sacks, H., Schegloff, E. A., and Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50. 696 – 735.
32. Searle, J. R. (1969). *Speech acts: An essay in the philosophy of language*. Cambridge, England: Cambridge University Press.
33. Starbird, K., Palen, L., Hughes, A. L., and Vieweg, S. (2010). Chatter on The Red: What hazards threat reveals about the social life of microblogged information. In *Proc. of the ACM Computer Supported Cooperative Work (CSCW) 2010 Conference,* Savannah, GA, February 2010, 241 – 250.
34. Thompson, S., Altay, N, Green, W. G. III, and Lapetina, J. (2006). Improving disaster response efforts with decision support systems. *IJ of Emergency Management*, 3, 4, 250 – 263.
35. Tukey, J. (1977). *Exploratory Data Analysis*. Reading, MA: Addison-Wesley.
36. Weick, K. E. (2001). *Making sense of the organization*. Malden, MA: Blackwell Publishers.
37. Zographos, K. G. and Androutsopoulis, K. N. (2008). A decision support system for integrated hazardous materials routing and emergency response decisions. *Transportation Systems Research Part C, Emerging Technologies*, 16, 6, 684-703.