Collaborative Option Awareness for Emergency Response Decision Making

Yikun Liu, Sung Pil Moon, Mark S. Pfaff
Indiana University - Indianapolis
{yikliu;sungmoon;mpfaff}@iupui.edu

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

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

ABSTRACT

We have been using exploratory modeling to forecast multiple plausible outcomes for a set of decision options situated in the emergency response domain. Results were displayed as a set of box-plots illustrating outcome frequencies distributed across an evaluative dimension (e.g., cost, score, or utility). Our previous research showed that such displays provide what we termed "option awareness" – an ability to determine robust options that will have good outcomes across the broadest number of plausible futures. This paper describes an investigation into extending this approach to collaborative decision making by providing a visualization of both collaborative and individual decision spaces. We believe that providing such visualizations will be particularly important when each individual's decision space does not account for the synergy that may emerge from collaboration. We describe how providing collaborative decision spaces improves the robustness of joint decisions and engenders high confidence in these decisions.

Keywords

Option awareness, decision support, robust decision-making, collaborative decision-making, collaboration.

INTRODUCTION

We developed **C**ollaborative **O**ption **A**wareness for joint a**ction**s—or COAction—to enable robust tactical joint decision making under challenging conditions. An example of such conditions occurred during the joint fire/rescue, hazardous materials (HAZMAT), and police response to the major explosion at the CAI, Inc. ink manufacturing plant in Danvers, Massachusetts in 2006. During this emergency, two dozen homes were seriously damaged and debris fell on approximately 90 buildings within a quarter-mile radius (Mishra, 2006). Emergency responders had to make safety-critical decisions such as the size of the evacuation area, the number of resources to deploy, and where to stage them for best effect.

Decision support systems have been identified as important tools for enabling emergency response decision makers to reduce the time needed to make crucial decisions for task assignments and resource allocation (Thompson et al. 2006). While there is a considerable body of recent research into emergency response decision support (e.g., Bonazountas, Kallidromitou, Kassomenos, & Passas, 2007; Campbell, Mete, Furness, Weghorst, & Zabinsky, 2008; Kondaveti & Ganz, 2009; Mendonca, Beroggi, & Wallace, 2001; Prolog Development Center, 2010; Yan, Liu, Zhang, and Zhou, 2010; Zographos & Androutsopoulos, 2008), none of these researchers are developing team-based systems for robust decision-making under uncertainty.

Prior to COAction research, we investigated individual option awareness, which was based on the distinction between the *situation space* and the *decision space* (Hall, Hellar, & McNeese, 2007). The situation space consists of facts about the environment such as the location and nature of an emergency situation, leading to situation awareness (Endsley, 1988). The decision space consists of information about the options that a decision maker might take. We have integrated these concepts with the exploratory modeling approach of Bankes (1993). The result is the ability to compare options and understand the underlying factors that contribute to outcomes, which we have termed *option awareness* (Drury, Klein, Pfaff, & More, 2009b). Our previous studies (Drury, Klein, Pfaff, & More, 2009a, 2009b; Drury, Pfaff, More, & Klein, 2009; Pfaff, Drury, Klein, & More, 2010; Pfaff et al., 2010) have shown that the addition of option awareness to situation awareness

Reviewing Statement: This full paper has been fully double-blind peer reviewed for clarity, relevance, significance, validity and originality.

can result in quicker, more accurate decisions that are made with more confidence.

While the results for individual option awareness are promising, many time-sensitive, safety-critical domains such as emergency response rely on teams to make important decisions. Thus we developed COAction to provide decision makers with a collaborative version of a decision space that enables joint option awareness.

Collaborative decision spaces require investigation because they are not necessarily formed as a combination of the top-ranked choices from the underlying individual decision spaces. The individual spaces do not account for the synergies or conflicts that may occur due to organizations with different abilities coming together to resolve an emergency situation. For example, picture the role that police have in effectively managing rush hour or special event traffic to ensure that fire vehicles can reach a fire quickly. Without such assistance, the fire keeps growing during delays in response and therefore is more difficult to extinguish when vehicles do arrive.

In our most recent experiment, we investigated questions such as:

- Will participants be less certain of their decisions if they see both their individual decision spaces and a combined space, compared to seeing only one of the two types of spaces?
- Will participants use different decision-making strategies in different types of situations?

This paper is the first to provide empirical results of an exploration of collaborative option awareness for decision support. As such, it provides both a baseline of performance and points the way towards future experimentation. To help in understanding the experiment, the next section provides additional background on the visual presentation of decision space information and the ways in which individual and combined decision spaces can differ.

DECISION SPACE VISUALIZATION

Computer-based forecasting models can assess dozens of options with hundreds of variations due to uncertainty, resulting in a landscape of plausible outcomes. A frequency format[1] approach to displaying the results (Gigerenzer & Hoffrage, 1995; Hoffrage & Gigerenzer, 1998) makes these results comprehensible (see Figure 1, explained further below). The goal is to provide the decision maker with an ability to compare options and ultimately understand the underlying factors that contribute to the outcomes.

Consider an example in which a fire breaks out in an historic building during a time when there is an automobile accident along the route from the fire station to the historic building. The four courses of action available to the fire chief would be to send between 0 and 3 fire trucks to the scene. To guide the decision-maker, we need a means of determining the relative desirability of the available options. We used the sum of the monetary costs implied by each option as an evaluative metric, so all consequences such as property damage, injury, and death are calculated and assigned monetary values[2]. The cost of each case in our example scenarios (which we call *emergency events*, or simply *events*) is computed by summing the cost of acting on the option, the costs of the direct consequences resulting from enacting the option, and any opportunity costs or additional costs that might occur in the future due to having enacted the option.

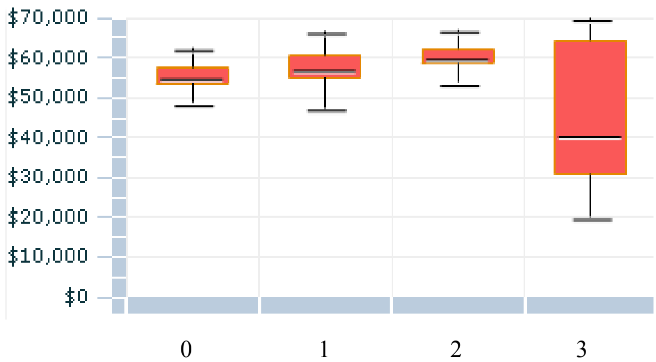

Figure 1. A decision space showing the relative costs of sending between 0 and 3 fire trucks to a fire, assuming no other response

Even for the same option, the costs vary depending on situational conditions beyond decision makers' control, such as whether strong winds fan flames or a drenching downpour quenches them. Thus for each option in Figure 1, there is a distribution of possible consequences. Each distribution is a function of the uncertainty of the situation space (e.g., how big is the fire) 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, Popper, & Bankes, 2003), where situation and execution uncertainty are irreducible, optimal strategies lose their prescriptive value if they are sensitive to these

[1] Displaying uncertainty information in terms of frequency distributions instead of probabilities.
[2] Note that insurance actuarial tables can be used to assign the cost of death.

uncertainties. That is, selecting an optimal strategy is problematic when there are multiple plausible futures for each option, as is the case in this example. Instead, Chandresekaran (2005) and Chandresekaran & Goldman (2007) suggest shifting from seeking optimality to seeking robustness for planning under deep uncertainty. Robust options are those that result in acceptable outcomes across the broadest swath of plausible futures.

The result of the cost evaluations for the range of plausible futures for a given decision option is summarized graphically by a box-plot for that option (Tukey, 1977). Although future research will be needed to determine a more nuanced visualization approach, we chose box-plots as a starting point to provide a simple means of comparing the cost distributions of the options. Besides their simplicity, box-plots are a common visualization of distributions that typical research subjects can be readily trained to read. We further simplified the box-plot visualization by eliminating outlier data points.

In Figure 1, the best choice is to send all three fire trucks to the scene. Because the traffic slows down the response time, the fires will be bigger than it would have been under lighter traffic conditions, and so all three trucks will be needed to combat the flames.

Sending three fire trucks is considered the top-ranked choice because its corresponding 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.

The "best three out of five" rule is a simple, reasonable, easily taught heuristic for determining the top-ranked option, but it assumes an equal weighting of each of the five box-plot distribution parameters. In real-world situations this strategy is not necessarily the best fit in all cases. We could imagine people concerned about the worst-case scenario choosing options that minimize the maximum cost, e.g., in situations where loss of life seems likely (we placed a high value on loss of life commensurate with insurance actuarial tables). We termed this weighting scheme *emphasize-maximum*. Another example is a normalized weighting scheme (called *normal*) that places the most emphasis on the median cost, and the least emphasis on the maximum and minimum cost cases due to the lower likelihood that they will occur. This was the default weighting used when the distributions were initially ranked and presented. Thus it is reasonable to assume that emergency response decision makers may weigh the distribution parameters differently under different conditions. This context-weight interaction means that the same box-plot graph may result in different rankings of options and different choices depending on the context that generated it and how that context is reflected in the decision maker's weighting strategy.

We experimented with providing other weighting schemes besides *equal* in Pfaff, Drury, Klein, More, et al., (2010) and determined that the percentage of cases in which participants chose the top-ranked option and their confidence in the decision increased if they were allowed to choose from among pre-set weighting schemes or create their own *custom* scheme.

Note that the visualization of the options in Figure 1 is presented from the 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? Clearly, collaboration is needed and hence we extended option awareness to assist multiple decision makers.

COLLABORATIVE OPTION AWARENESS

In the example case, the fire is reported just as the roads are clogged with bystanders viewing an accident scene. Normally police would send one squad car to handle the traffic and fill out accident reports for an accident of this magnitude. Such an approach, however, ignores the possibility for the two emergency response departments to help each other. If the police department sends additional vehicle(s) to clear traffic in favor of the fire trucks, the extra police presence can help the fire trucks reach the fire more quickly. 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.

Figure 2 shows the decision space for the police, which is based on the assumption that the police are only concerned with the traffic incident. As expected, the figure indicates that the most robust option is to send one squad car. A different picture emerges from Figure 3, however, which illustrates the combined decision space for a collaborative response. The most robust combined option is to send two fire trucks and two police cars.

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 uncertainty, especially when decisions must be made quickly and under stressful conditions.

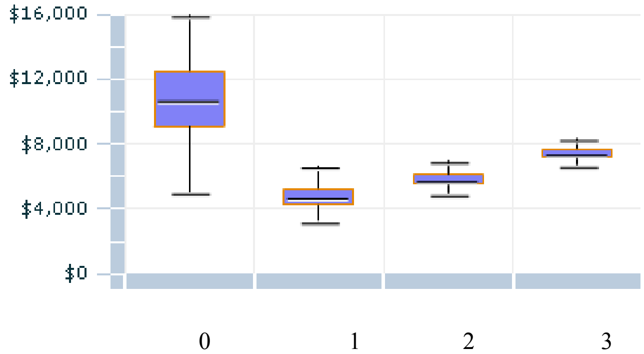

Figure 2. A decision space showing the relative costs of sending between 0 and 3 police cars to a traffic accident.

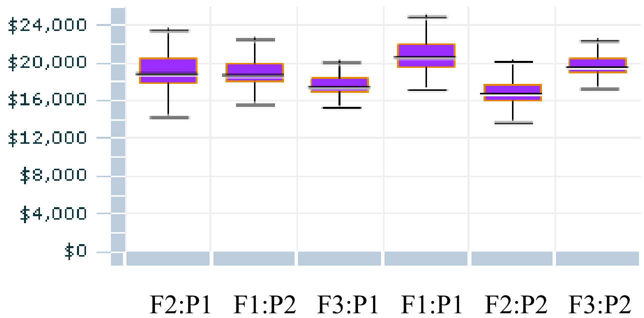

Figure 3. A combined decision space showing the relative costs due to the synergy of sending combinations of fire and police vehicles.

Legend: F*x*:P*y* = *x* fire trucks and *y* police cars.

The purpose of the combined decision space is to provide rapid understanding of the likely costs and consequences of collaborative decisions.

In the fire/accident event discussed above, the two most robust options in the individual fire and police decision spaces taken together (three fire and one police) are not the same as the most robust option in the combined decision space (two fire and two police). Because of this difference, the combined decision space is *conflicted*. Sometimes, however, the most robust option in the combined decision space will simply be the sum of the most robust options in the individual spaces. We call this an *unconflicted* combined decision space.

METHODOLOGY

This experiment employed a mixed design. For the between-subjects portion of the experiment, the two-person teams of participants were divided into three groups and given individual and combined, individual-only, or combined-only decision spaces as a mechanism for providing option awareness and decision support. 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. Ten of the events used in this experiment were also used in a previous experiment; these events required the two-person teams 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 are called “heterogeneous” events because they involved two-person teams of participants allocating different kinds of resources to two potentially related events within the same city. For example, one person decided the number of fire trucks to send to an incident while the other decided how many police squad cars to allocate. 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.

<table>
<tr><th rowspan="4"></th><th colspan="4">Collaborative Decision Space</th></tr>
<tr><th colspan="2">Yes</th><th colspan="2">No</th></tr>
<tr><th colspan="2">Individual Decision Space</th><th colspan="2">Individual Decision Space</th></tr>
<tr><th>Yes</th><th>No</th><th>Yes</th><th>No</th></tr>
<tr><th rowspan="2">Types of Synergy (Within Subjects)</th><td>Homogeneous Resources – Load Sharing</td><td>Homogeneous Resources – Load Sharing</td><td>Homogeneous Resources – Load Sharing</td><td></td></tr>
<tr><td>Heterogeneous Resources – Complementary Option Elements</td><td>Heterogeneous Resources – Complementary Option Elements</td><td>Heterogeneous Resources – Complementary Option Elements</td><td></td></tr>
</table>

Table 1. Experimental design indicating the three between-subject manipulations used and the two within-subject categories of emergency events.

Procedures

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, which included a list of frequently-asked questions. The manual included a step-by-step walkthrough of the user interface based on the first of three practice events users encountered before beginning the actual experiment. This walkthrough and subsequent practice events familiarized participants with the steps required to complete the task, the functionality of the user interface, and guidelines for how to identify the best option for each emergency event. During training (only), participants received feedback regarding the current magnitude of the event, its potential impact, and the potential for future situations to occur.

Following the training, participants were presented with the 20 emergency events in a randomized order. A description of the current emergency event and a log of previous events was displayed in panel on the left half of the screen. A map of the area 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 event location. To maximize the screen area available to display event information and provide interactive controls, participants could show or hide the map as needed. An embedded chat window on the right hand side of the display was set up to enable (and record) communication between the partners in the two-person teams. Note that this paper reports performance results in accordance with the hypotheses presented below, and does not include results from analyzing the content of the chat data.

Decision space information was displayed as sets of box-plots representing the six best-ranked options among all possible combinations of actions (with a rank of 1 being the best-ranked option), with one set of box-plots being presented for each type of decision space according to the condition. A set of controls for manipulating the weighting of the box-plot elements (and hence the rankings of the options) was associated with each of the participants' own decision space(s) to facilitate individual exploration. As participants modified the weighting strategy, the ranks of the available options were recalculated in real time. Those options whose rank changed as a result of the last change in strategy were highlighted in red. Participants could visually sort the 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. 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, with different weightings potentially resulting in different rankings being assigned to these same box-plots. Finally, participants in the "individual + combined" condition saw all of these sets of box-plots.

Once participants submitted a decision of which option they chose, a pop-up screen appeared for them to confirm the decision. This screen reminded them of their decision and what weighting strategy they had last selected. 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, enter a brief textual justification of their decision, and proceed to the next event at the same time.

After evaluating 20 emergency events, participants filled out a set of surveys assessing individual traits likely to 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.

Hypotheses

We defined seven hypotheses, as follows.

H1. For the heterogeneous events, the team's two choices will, when taken together, constitute the highest ranked option less often when they have individual spaces (only) compared to when they also—or only—have the combined decision space.

Rationale for H1: Participants will not accurately forecast the impact of any synergistic effects without the aid of the combined decision space, which was designed to take these effects into account.

H2: For the individual only condition, participants will choose the combination of the highest ranked options in each individual space more often than any other choices.

Rationale for H2: Since participants will not have any other decision spaces they will tend to each choose the individually highest ranked option.

H3: Participants in the individual only and combined only conditions will have higher confidence in their decisions than those in the individual + combined condition.

Rationale for H3: Individual and combined decision spaces may conflict, so if participants see only individual spaces they will not, by definition, see any conflicts. In the absence of conflicts, they will be more confident in their decisions.

H4: When participants choose 7th- or higher-ranked options, we hypothesize that these options are close variations of the top-ranked options in terms of an “option distance.” This is a two-dimensional metric identifying the numeric differences between pairs of choices (e.g., one fire truck from Station A and two from Station B) compared to the pair representing the top-ranked option (e.g., two fire trucks from Station A and one from Station B).

Rationale for H4: We are hypothesizing that some poorly-scored choices are actually the result of trying to develop a creative variation of a top-ranked option, without understanding that the slight variation will likely yield unexpectedly large changes in cost.

H5: For the homogeneous events when participants have combined decision spaces only, the team’s two choices will, when taken together, constitute the highest ranked option less often than the comparable cases in the previous experiment.

Rationale for H5: Note that this hypothesis compares results obtained in this experiment with corresponding results obtained in an earlier experiment (Pfaff, Drury, Klein, More, et al., 2010). The homogeneous events are a subset of the events used in the that experiment (in which individuals had the equivalent of combined decision spaces), but decision spaces were more likely to have been in line with participants’ expectations in the last experiment because the participants set the model’s input parameters in that experiment but not in this one.

H6. When the suggested outcome of an event is death or injury, participants will be more likely to select the emphasize-maximum weighting strategy than when the suggested outcome is only property damage (ten of the 20 scenarios suggest death or injury as an outcome).

Rationale for H6: The emphasize-maximum weighting strategy is more likely to avoid options that include cases that could result in the highest cost, which are where deaths are more likely to occur. With two decision makers the death factor may be more salient and less likely to be ignored than if the two participants were working in isolation

H7: The lesser the visual variability of the options in the decision-space visualization, the more weighting strategies participants will consider.

Rationale for H7: If the options displayed are too visually ambiguous to easily evaluate, we predict more vigorous exploration of the weighting strategies. Participants will have a greater desire for confirmatory feedback and will use the weighting strategies as second, third, and fourth (if applicable) opinions.

Participants

The experiment took place 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 whereas fourteen participants were familiar with box-plots. Three participants had some previous emergency response experience.

RESULTS AND DISCUSSION

H1 (decision accuracy will be worse in the individual-only condition) was supported. For heterogeneous events, participants in the individual-only condition selected the highest ranked combined option only 9.76% of the time, compared to 45.12% in the combined-only condition, and 21.79% in the individual+combined condition,

$\chi^2(12, N = 242) = 34.89$, $p < .001$. The results for the combined-only and individual-only conditions were significantly different at $p < .01$. Note that these results are only for the heterogeneous events. The picture for the homogeneous events is quite different. For homogeneous events, participants in the individual-only condition selected the highest ranked combined option only 28.00% of the time, compared to 43.96% in the combined-only condition, and 45.95% in the combined + individual condition, $\chi^2(12, N = 265) = 26.51$, $p < .01$. However, these three results were not significantly different from each other.

H2 (participants in the individual-only condition will select the highest-ranked individual option more often than any other choices) was supported. Participants only seeing the individual decision space selected the top-ranked individual option 47.25% of the time, more than double the frequency of selecting any other options.

H3 (participants in the individual-only and combined-only condition will have higher confidence in the combined decision than those seeing *both* the individual and combined decision spaces) was partially supported. A Kruskal-Wallis test showed that participants with only the individual decision space had a significantly higher mean confidence ($M = 6.34$, $SE = .09$) than those with the combined decision space only ($M = 5.80$, $SE = .09$) and both the combined and individual decision spaces ($M = 5.89$, $SE = .09$), $H(2) = 22.71$, $p < .001$. Post-hoc Mann-Whitney tests indicated a significant difference between the individual-only and combined-only conditions ($U = 10985$, $r = -.27$) and the individual-only and individual+combined conditions ($U = 11909.5$, $r = -.12$).

H4 (participants selecting options ranked 7th or greater are actually choosing close variations of the top ranked option) was not supported. The error distance calculated for those options was not only the greatest of all ranks, but also had by far the greatest variance.

H5 (participants in the combined-only decision space condition will select the highest-ranked option less often than for the same homogeneous events used in the prior experiment, in which individuals saw the same combined decision space) was supported. Individual participants in the prior experiment selected the highest-ranked option 67.91% of the time, compared to paired participants selecting the highest-ranked option only 43.96% in the present experiment, $\chi^2(6, N = 695) = 67.52$, $p < .001$.

H6 (participants will select the emphasize-maximum strategy more often when death or injury is indicated in the event description than when the outcome is only property damage) was not supported. There was no significant difference in the use of weighting options between those event types.

H7 (participants will use more weighting strategies as visual variability decreases) was not supported. There was no significant relationship between the visual variability of the options displayed and the usage of weighting strategies. Nor was there any change in the amount of text-based chat, which may have accounted for some other means of exploring their options in lieu of the weighting.

CONCLUSIONS

The contribution of this work is based on the fact that it is the first formal experiment comparing decision making performance using individual or combined (or both) decision spaces. Further, this work is unique in its explicit emphasis on visualizing the effects of collaboration synergy.

It is clear from the results that the decision space visualization helped decision making, but the nature of this support is now much more complex under collaborative decision making than in previous experiments with a single decision maker. In the heterogeneous events, the nature of the synergy between the participants' resources was the most complex. In the simpler homogeneous events, collaboration involved mostly comparing the availability and distance from the event of the resources of the two decision makers. In the heterogeneous events, the actions of the other participants' resources impacted performance in a more complex fashion. For example, the police could reduce congestions and improve the fire department's time to get to the event. Another impact could be for the police to control crowds and improve the fire department's firefighting effectiveness. The results suggest that this complexity made it more difficult for participants to understand the nature of the decision space. Participants who received only the individual decision spaces were obviously unaware of this complex synergy, even though they had been warned of the possibility during their conceptual training. Their lack of joint option awareness resulted in their discovering the most robust joint option only about 10% of the time, and even choosing a high-cost joint option ranked 7th or greater (meaning, at least six options were ranked more favorably) about 40% of the time. Participants who saw the combined decision space did significantly better. However, even though those who only saw the combined decision space selected the most robust option about 45% of the time, they also chose a high-cost option ranked 7th or greater about 22% of the time. Moreover, adding the individual decision space view to the combined decision space view seemed to introduce confusion. It not only reduced by half choosing the most robust option displayed (down to about 22%) it also doubled the frequency of choosing a high-cost option to 43% of the time.

We hypothesized that when participants selected options ranked 7th or greater they would do so by falling into the trap of thinking that minor variations to the most robust option would achieve their objectives with minimum additional cost. However, the results showed that the participants did not confine themselves to minor variations. This seems again to demonstrate a lack of understanding of the nature of the decision space.

This lack of understanding is also demonstrated by the participants' relatively high confidence that they had chosen well, even when they mostly chose incorrectly especially in the individual space only condition. Indeed, that condition showed the highest confidence – perhaps because the decision space seemed simpler, even though it truly was not.

In the simpler homogeneous event conditions where the combined space involved mostly adding the resources of the two decision makers, we find a very different result. Here the individual decision space seems to add rather than detract from understanding the event. This can be seen in the improved performance of the individual-only participants, who now select the most robust option almost three times as often (28% vs. 10%) and selecting the 7 or greater ranked option now only 25% of the time. This is also clear from comparing the combined-only condition with the individual+combined condition. For homogeneous events, the participants in the individual+combined condition selected the most robust option about 46% of the time – equivalent to that of the combined-only condition. Moreover, they now selected 7th or greater ranked options only about 11% of the time.

In summary, the research has demonstrated that collaborative decision making is improved by collaborative decision space support. Providing a visualization of combined decision space is especially effective when that space is not formed merely as a combination of the top-ranked choices from the underlying individual decision spaces – as was the case in the heterogeneous event conditions. Not only was providing the individual space alone not very helpful, but in these heterogeneous event conditions the individual space actually detracted from the help provided by the combined space.

LIMITATIONS AND FUTURE WORK

Our approach assumes that all consequences can be evaluated in terms of a single metric, which we have selected to be cost. We acknowledge that it can be difficult to quantify all consequences in this way. To illustrate this point, think of the long-term effects of a landslide. The cost of such an event may depend on the future economic conditions that will drive land values and the likelihood that this piece of land would have been used for a low-yield economic purpose (e.g., farming) versus a high-yield purpose (e.g., a multi-story shopping/office/residential complex), had it not fallen away and become unstable. Development of a truly all-purpose evaluative metric for crisis management represents an opportunity for future work. However, for the present work using novice decision makers, such complexity, while authentic, can interfere with the assessment of the fundamental cognitive and behavioral underpinnings of interactions with the decision space.

Another limitation we imposed on the study was the examination of decision points as discrete episodes. In other words, each event represented a snapshot in time and was not connected in any way to the event that was presented previously or subsequently. Further, except for providing feedback on participants' assessment of the initial event characteristics during training, we did not provide feedback that would enable participants to learn from their experiences. We do not know if results would differ if participants saw the events unfold in a coherent sequence and were able to use their prior experience with the decision space to influence their future choices. Obviously, this approach represents another opportunity for future work.

We assumed in this study that decision makers would want to make their own choice regarding the best weighting scheme to use in any particular situation and to see the ranges of costs that characterize the top six options. But what if they do not have the time or spare cognitive capacity at that moment to deliberate in this fashion? Future work could investigate the desirability of providing automated recommendations to decision makers regarding the most appropriate weighting strategy and therefore the top-ranked option. Would decision makers trust the recommendation, and would their confidence in the recommendation be as high as the confidence in their own decisions when they manipulate the weights themselves?

We used relatively simple experimental events, yet their synergies seemed to be cognitively difficult to understand. Obviously, events will be more complex in the real world. We found that the collaborative decision space provided at least a minimum understanding of the synergies in the experimental events. However, achieving more than the minimum understanding likely will require additional experience with collaborative decision-spaces, so that joint option awareness can become part of the decision maker's repertoire. Further research will be needed to determine if providing such experience will indeed improve performance, and to determine the cognitive limits to understanding and making robust decisions in complex spaces.

This paper presents data on participants' performance using the decision spaces, and does not include an analysis of the process by which the arrived at their decision (other than to note their exploration with the different weighting schemes prior to making a decision). The next task we have set for ourselves is to analyze team members' chat postings. Since their only way to communicate with each other was through the chat tool, such an analysis should provide insight into their decision making process. Further, other issues regarding team dynamics were not addressed such as gender-matching within teams, which is known to impact collaborative processes especially in college student populations (Williams, Ogletree, Woodburn, & Raffeld, 1993). With a sample that was predominantly male in this study, future work needs to pursue more balanced samples when possible.

Finally, we noted previously that the decision space visualization may need to be tailored to better meet the needs of decision makers. While box-plots provide a summary of the distribution of consequences for each decision option, they do so in a static manner, meaning that they only show the consequences if an action was taken at that point in time. But what would be the consequences of taking an action at time 0 versus time 6 (for example)? Other visualizations would be needed to show time-based information. Even if it was sufficient to show the ranges of costs for decisions made at that moment, there may be simpler or more elegant visualizations that support performance that is as good or better than that attained when using box-plot-based visualizations.

ACKNOWLEDGMENTS

We thank the anonymous participants in the experiment. This work was supported in part 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 manageing forest fire casualties. *Journal of Environmental Management*, *84*, 412-418.
3. 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, 176-180.
4. 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, The Ohio State University.
5. Chandrasekaran, B., and Goldman, M. (2007, 1-5 April 2007). *Exploring robustness of plans for simulation-based course of action planning: A framework and an example.* IEEE Symposium on Computational Intelligence in Multicriteria Decision Making, Honolulu, HI.
6. Drury, J. L., Klein, G. L., Pfaff, M. S., and More, L. D. (2009a). *Data visualizations for dynamic decision support*. Paper presented at the Human Interaction with Intelligent and Networked Systems Workshop, Intelligent User Interfaces Conference (IUI 2009), Sanibel Island, FL.
7. Drury, J. L., Klein, G. L., Pfaff, M. S., and More, L. D. (2009b). Dynamic decision support for emergency responders. *Proceedings of the 2009 IEEE Conference on Technologies for Homeland Security (HST '09)*, 537-544.
8. Drury, J. L., Pfaff, M. S., More, L. D., and Klein, G. L. (2009). A principled method of scenario design for testing emergency response decision making. *Proceedings of the 6th International Information Systems for Crisis Response and Management (ISCRAM 2009)*, Göteborg, Sweden.
9. Duttweiler, P. C. (1984). The Internal Control Index: A newly developed measure of locus of control, *Educational and Psychological Measurement,* vol. 44, pp. 209-221.
10. Endsley, M. R. (1988). Design and evaluation for situation awareness enhancement. *Proceedings of the 32nd Annual Meeting of the Human Factors and Ergonomics Society*, 97-101.
11. Gigerenzer, G. and Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats, *Psychological Review*, *102* (4), 684-704, American Psychological Association.
12. Hall, D. L., Hellar, B. and McNeese, M. (2007). Rethinking the data overload problem: Closing the gap between situation assessment and decision making, *Proc. of the 2007 National Symposium on Sensor and Data Fusion (NSSDF) Military Sensing Symposia (MSS)*, McLean, VA.

13. Hoffrage, U. and Gigerenzer, G. (1998). Using natural frequencies to improve diagnostic inferences. *Academic Medicine*, *73*(5), 538-540.
14. 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, September 2009.
15. 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*. RAND MR-1626. Santa Monica, Calif.: The RAND Corporation.
16. Madsen, M., and Gregor, S. (2000). Measuring human-computer trust. *Proceedings of the 11th Australasian Conference on Information Systems*, 53-65.
17. McLain, D. L. (1993). The MSTAT-I: A new measure of an Individual's tolerance for ambiguity, *Educational and Psychological Measurement, 53*(1), 183-189.
18. 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.
19. Mishra, R. (2006). A 'Thanksgiving miracle:' No deaths, few injuries as Danvers blast levels plant, leaves hundreds homeless, *The Boston Globe,* 23 November 2006. [Online]. Available: http://www.boston.com/news/local/massachusetts/articles/2006/11/23/a_145thanksgiving_miracle146/. [Accessed: 2 June 2010]
20. Pfaff, M. S., Drury, J. L., Klein, G. L., and More, L. D. (2010). Decision support for option awareness in complex emergency scenarios. *3rd International Conference on Applied Human Factors and Ergonomics (AHFE)*, Miami, FL.
21. Pfaff, M. S., Drury, J. L., Klein, G. L., More, L. D., Moon, S. P., and Liu, Y. (2010). Weighing decisions: Aiding emergency response decision making via option awareness. *Proceedings of the 2010 IEEE International Conference on Technologies for Homeland Security (HST)*, 251-257.
22. 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.
23. Thompson, S., Altay, N, Green, W. G. III, and Lapetina, J. (2006). Improving disaster response efforts with decision support systems. *International Journal of Emergency Management*, *3*(4), 250-263.
24. Tukey, J. W. (1977). *Exploratory Data Analysis.* Reading, Mass: Addison-Wesley.
25. Weber, E. U., Blais, A.-R. and Betz, N. E. (2002). A domain-specific risk-attitude scale: Measuring risk perceptions and risk behaviors, *Journal of Behavioral Decision Making, 15*, 263-290.
26. Williams, S. W., Ogletree, S. M., Woodburn, W., and Raffeld, P. (1993). Gender roles, computer attitudes, and dyadic computer interaction performance in college students. *Sex Roles*, *29*(7), 515-525.
27. Yan, Y., Liu, H., Zhang, Y., and Zhou, Z. (2010). A fuzzy-QFD approach to design decision support system for emergency response of hazardous materials road transportation accidents. In *Proceedings of the 2010 IEEE International Conference on Automation and Logistics*, Hong Kong and Macao, August 2010.
28. 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.