12TH ICCRTS

"Adapting C2 to the 21st Century"

Title: The PRIM: Extracting Expert Knowledge for Aiding in C2 Sense & Decision Making

Tracks: Sensemaking, Social Processes, & Small Group Decision Making

Authors: Hudson Graham
Graduate Student
Humans & Automation Laboratory
Massachusetts Institute of Technology

Gilles Coppin
Visiting Professor
Humans & Automation Laboratory
Massachusetts Institute of Technology

M.L. Cummings
Assistant Professor
Humans & Automation Laboratory
Massachusetts Institute of Technology

Contact: Hudson Graham
Humans & Automation Laboratory
Massachusetts Institute of Technology

77 Massachusetts Avenue, Building 35-220
Cambridge, MA 02139

Work (617) 253-0993
Cell (719) 238-1713

hgraham@MIT.edu

Abstract: With the rapid and continual flow of information in today's battle spaces, the volume, tempo, and uncertainty of incoming information in the command and control (C2) environment can overwhelm decision makers, particularly those in complex mission planning teams. This paper will discuss a decision support design intervention that aids in the formulation of recommended courses of action for a command staff. This includes the gathering of expert knowledge and opinions in order to provide a vehicle for team dialogue and negotiations for alternative solutions. The key focus is accurately capturing the experts' decision making processes via a computer interface, so that differing points of view can be shared and can eventually converge towards one recommended consensual course of action. This research investigates whether the proposed Probability and Ranking Input Matrix (PRIM) allows for efficient and clear user input as well as to promote understanding of variable relationships and their likelihood of occurrence.

1 Introduction

In today's age of information, where massive data is continuously generated and updated, a complete and accurate picture of the environment can be difficult to maintain. Battlefield commanders are not exempt, as they rely on their staff and electronic assistance to keep a current and true picture of what is occurring in the battle space. For this complex environment, there is a need to generate helpful software that can aid battlefield commanders to synthesize information.

The focus of this research is to create a computer-based interface and software that can capture expert knowledge from a battlefield commander's staff and generate a recommendation based on the staff's knowledge. The interface presented in this paper is the Probability and Ranking Input Matrix (PRIM).

The intention of this paper is to show that with a focus on direct perception-interaction design, PRIM allows experts to input a variable's importance, effect, and likelihood of occurrence for possible courses of action. In completing a pilot usability study, a basic utility model is used to illustrate how information can be extracted from a user-focused interface design.

2 Design Considerations

Before viewing PRIM, there are some specific considerations that must be discussed concerning previous research. The first is how human decision makers process information. Klein describes "skilled problem solvers and decision makers" as "chameleon[s]" in that they "can simulate all types of events and processes in their heads" (Klein, 1999). In other words, humans relate through stories. In designing any group decision support tool, a key concept is to encourage the commander's staff to relate their knowledge through a story-telling environment, while weighing and linking together relevant variables in order to causally support one possible decision outcome.

An additional point discussed by Kahneman and Tversky is of the impact of the recency effect. If an individual has a recent experience with an event and an event outcome, then he or she assumes it is most likely to occur that way, if that event occurs again, which can significantly bias decision making. These authors also discuss anchoring and adjusting biases, in which humans have a tendency to rate specific points like the mean or extremes (the maximum or minimum) and then adjust other ratings from those points (Kahneman & Tversky, 1973, 1979). All of these numerical and rating biases must be remembered in creating a group decision support tool such as PRIM.

One factor to consider in scoping the problem and variables is that research shows humans only process a few alternatives when thinking about solutions (Gigerenzer *et al.*, 1999). As a result, PRIM's design only allows the commander to dictate a few alternatives that the staff can consider in their comparison models. Another factor is that humans do not process through all attributes of a decision alternative when making a decision (Slovic *et al.*, 1978; Wiener & Nagel, 1988); humans characteristically focus on three to four attributes when making a decision (Wickens *et al.*, 2004). In fact, experts typically assign variables of a decision into three to four categories and work from there (Klein *et al.*, 2004; Wiener & Nagel, 1988). As will be illustrated, because of these points, PRIM only allows the commander to dictate four categories within which the staff can work. By limiting the number of alternatives and number of categories for the variables to only a few, the problem is immediately scoped to a common level for all the members of the team.

Another design consideration is that no added accuracy comes from having decision models with more than ten variables. The only added benefit for additional variables beyond ten is more confidence in the output (Oskamp, 1965; Wickens *et al.*, 2004). Because there is no added accuracy beyond ten variables, each of the categories created in PRIM is limited to ten variables.

PRIM is designed for expert decision makers. Designing for experts is unique because of how they approach problems. Beyond the ability to immediately categorize problem variables, there are three advantages experts have in problem solving (Clemen, 1996). First is a specialized knowledge base. A military example is that a commander's staff has field experience that gives them expertise on equipment and tactics, as well as the limitations and capabilities of the troops. Second, experts often make multiple predictions. Similarly, the commander's staff is more senior, with many years of experience in the field, which likely provided numerous opportunities to make evaluations. Third, most experts receive immediate feedback. In the military, with each event, experts are meticulous about briefing, before and after mission events, and learn quickly from their experiences. These experiences are what produce the unique expertise upon which PRIM is trying to draw.

The team aspect of PRIM is another area of interest. The team aspect of mission planning is critical, and PRIM tries to play on the advantages of a collective conscious (Klein, 1999). The tool is meant to help collect ideas about key factors and advantages or disadvantages to alternative courses of action by generating dialogue from all team members. Sometimes a collective conscious can slow down the decision making process, but within the military structure, where ultimately the commander makes the final decision, it should be helpful because it will allow the commander to see and hear all the possibilities.

A social aspect of the team to be considered is that members will want to put in solutions that will be accepted by the team (Gigerenzer et al., 1999); this can be both good and bad. In one sense this may cause team members to appease a norm. However, this may be an advantage in that it will force the members to assume accountability for their solutions. Each team member's input must be supportable, because he or she will have to justify inputs to the rest of the team. An added benefit of an automated decision support tool with a distributed team focus is that the chain of command influences can be somewhat removed, since input is not influenced by rank or position. Thus, everyone has somewhat of an equal position in the discussion in a remote, virtual environment, and thus can share their knowledge more freely.

3 PRIM Components

To limit the problem solving scope for PRIM, it is expected that commanders or a designated individual will have already put in the categories and variables to consider in comparing alternative outcomes to a decision regarding different proposed courses of action. These categories are a way of organizing key components of the decision. The variables are those elements that influence each of the different categories. By predefining categories and variables (which could be done by a computer), the focus then shifts to the staff's expertise of importance, effect, and likelihood of these variables.

The PRIM is designed to walk the user through three information input stages before providing an output recommendation for the commander. The three input stages and how they progress are shown in Figure 1. The process presented in Figure 1 is carried out for each alternative that must be considered. In the first stage, the user ranks a variable's importance in a decision, a variable's effect on an alternative, and a variable's subjective likelihood of occurrence for each category of variables. Figure 2a shows the interface where the user inputs this information. The next stage is where the individual categories are shown side by side for an individual's model, as seen in Figure 2b, which allows the user to make minor adjustments to the ranking of a variable's importance across all categories. The final input stage, as seen in Figure 2c, compiles individual models in order to show a team model.

The individual category stage is completed on the interface shown in Figure 3. It is comprised of four components. The bottom right shows what alternative is under consideration. Again, there will only be a few alternatives that are under consideration, and a separate model will be created for each alternative. The top left shows what category of variables the individual is currently considering. Below the category label is the variable selection bank. This bank contains all of the pre-selected variables for consideration in the category. Again, as discussed in the previous design considerations section, the bank is limited to ten variables. Each variable is on a puck which can slide over (through clicking and dragging) onto the ranking matrix. The puck is then placed in one of five bins (from low to high) to indicate how important the individual feels the variable is in their decision. This free form of sliding and placing allows users to indicate slight variations in the importance of two variables. For instance, a person may feel two variables are of high importance to the decision, but one is slightly more important than the other. In this case, the user would put them both in the High bin, but the most important would be in the top half of the bin, while the other would be in the bottom half of the bin. The dashed lines in each bin divide the bin in half and allow for these finer distinctions to be shown.

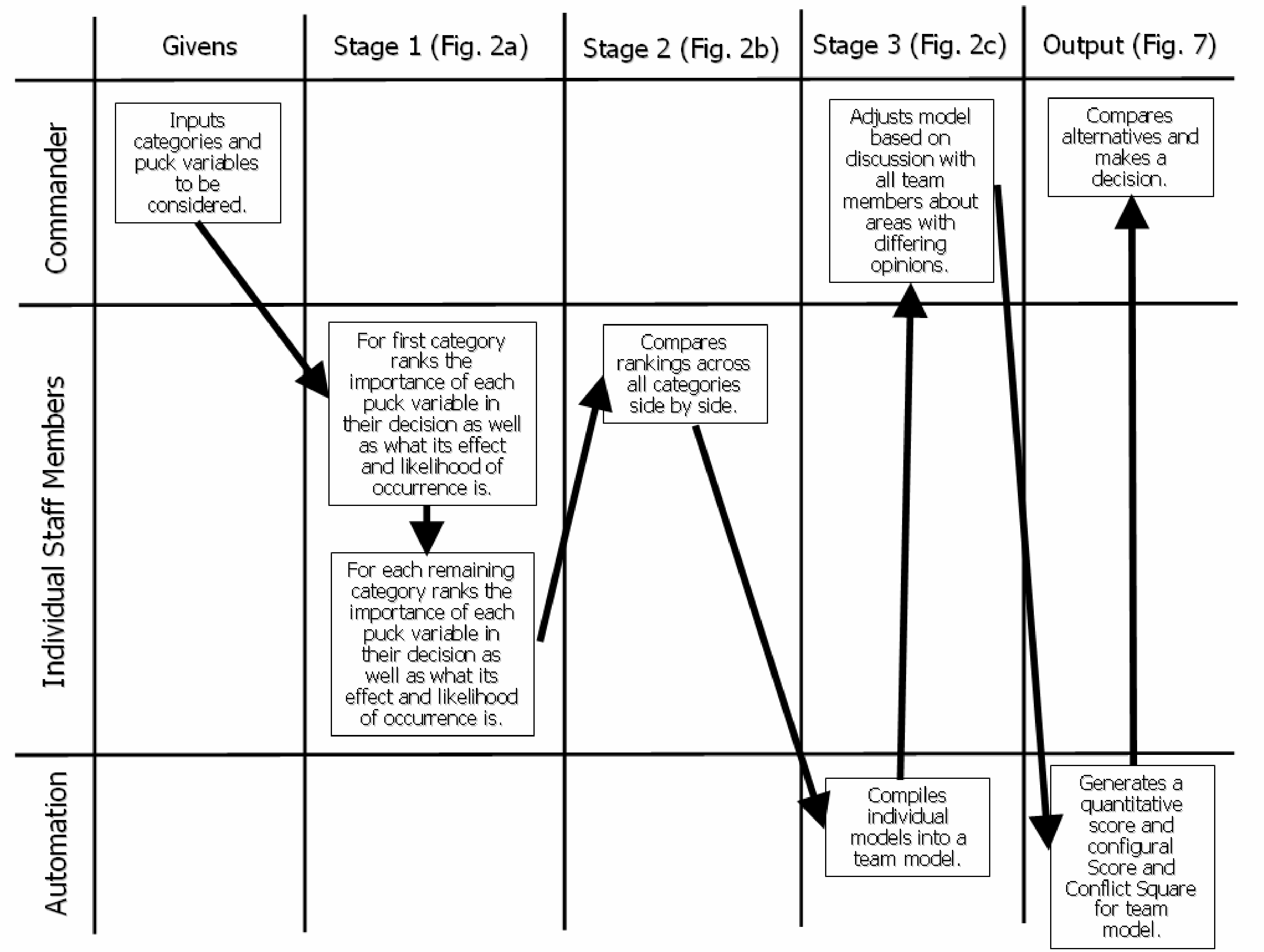

Figure 1: Process of inputting expert knowledge for one alternative PRIM.

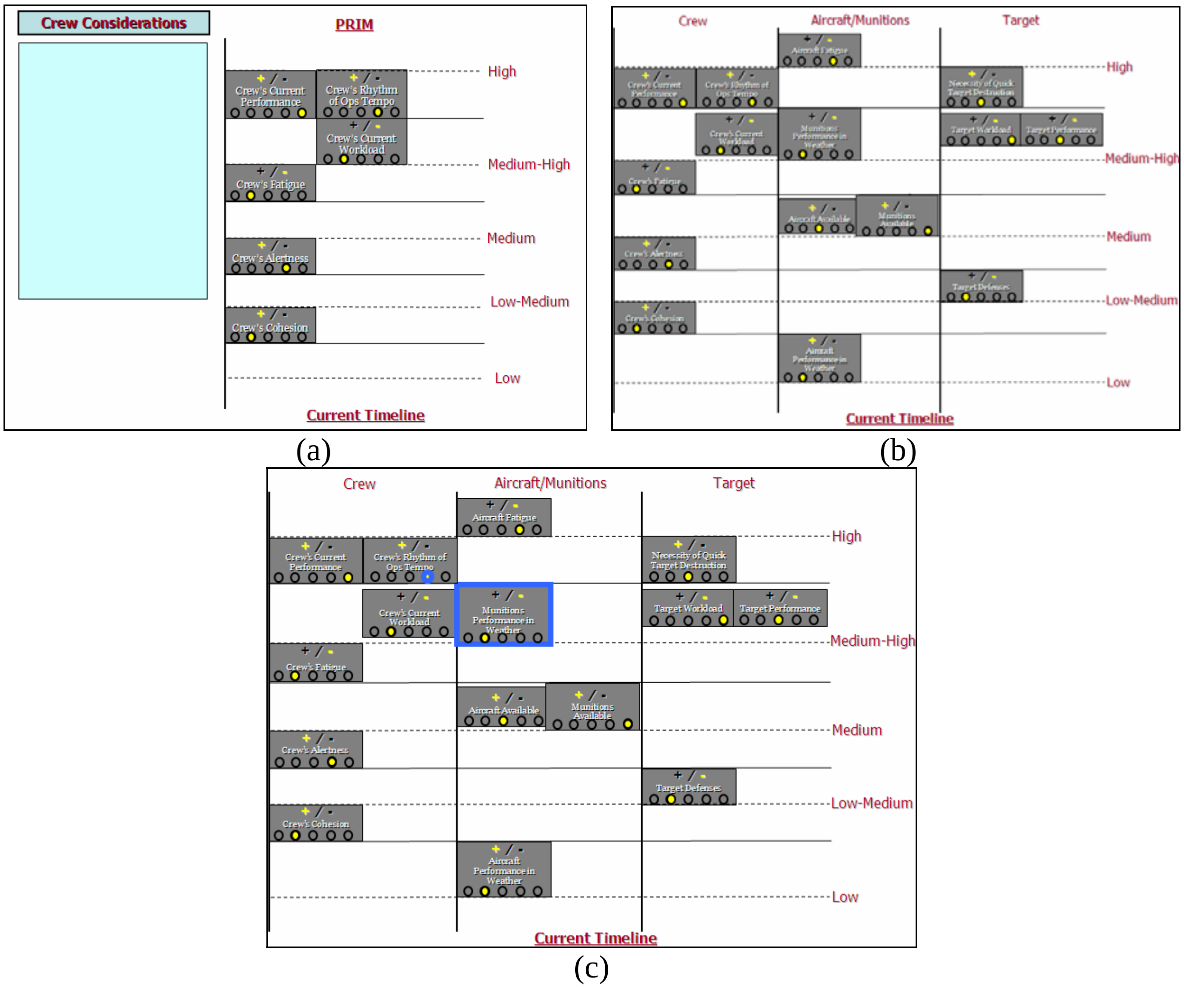

Figure 2: (a) A stage 1 individual PRIM category display, (b) The stage 2 individual PRIM cross category comparison display, and (c) The stage 3 team PRIM display.

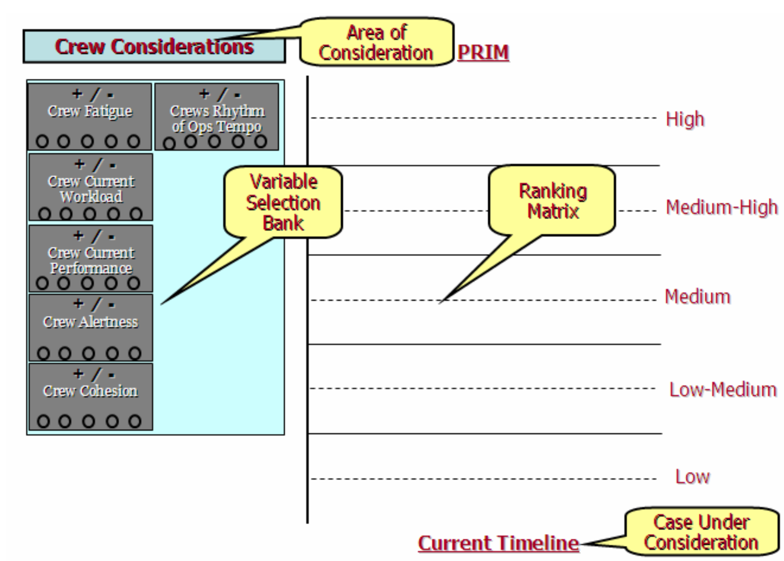

Figure 3: Individual PRIM category display.

After placing the puck on the ranking matrix, the individual must then indicate its effect on the overall alternative. This effect is indicated by either clicking on the positive or the negative sign on the top of the puck; see Figure 4 for a detailed picture of the puck. For example, in this case the user may click on the negative sign because of crew fatigue from weeks of continual work, and the user may believe this is detrimental to their performance. The next selection the individual must make is how likely the variable is to have an influence on the possible course of action. Again, because the crew has worked continually for many weeks, the individual will probably choose very likely, which is the far right bubble of the five bubbles along the bottom of the puck. These bubbles give five bins of probability, from left to right, as follows: very unlikely, unlikely, neutral, likely, and very likely. There are only five choices because humans have difficulty with absolute judgments, but are better with relative judgments (Sanders & McCormick, 1993).

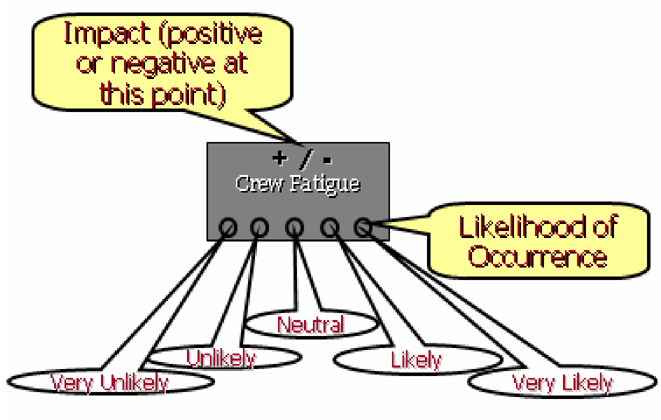

Figure 4: PRIM variable puck.

4 Walkthrough Scenario

A scenario was generated to explore the concepts of PRIM, which assumes the person using PRIM is a member of an Air Force battle staff tasked with reviewing a future Air Tasking Order (ATO). In the scenario, a decision must be made whether to launch as scheduled tomorrow or delay for 24 hours to accomplish a mission to destroy three key industrial factories that are supporting the resupply of the enemy's Unmanned Aerial Vehicles (UAVs). These UAVs are linked with the enemy's ability to maintain continual surveillance of friendly troop movements. The staff member's commander has three hours to make a decision and has given the staff member a list of key variables for three categorical areas the commander believes should be considered: the crews' readiness, the aircraft and munitions' readiness, and the target. The staff member must give the commander input of the importance, effect, and likelihood of these variables on the two alternatives—launching tomorrow on the current timeline or launching 24 hours later on the delayed timeline.

Other details of the scenario include the following considerations. The crews have been working twelve hour shifts, seven days a week, since the war began six weeks ago. The weather is forecasted to be overcast with scattered thunderstorms for the next eighteen hours. The clouds may block the precision munitions. After the thunderstorms, the forecast is for clear blue skies for the next week. Over the last seven days, 95 percent of all missions have been deemed a success (the set targets or objectives were achieved). Currently 30 percent of assigned aircraft are inoperative. Half of those aircraft have gone into maintenance during the last three days.

5 PRIM Input Walkthrough

In applying this scenario, an individual ranks importance, effect, and likelihood of each variable for each category, and then the different categories with their respective variables are compared as seen in Figure 2b. This comparison allows the user to make minor adjustments between the rankings of importance across all categories. For instance, if Fatigue is the primary concern in the decision, this allows the user to ensure it is the highest variable across all three categories on the ranking matrix scale.

From the individual categorization phase, the model moves to the team phase, where the individual results are compiled into an aggregate representation, as in Figure 2c. Summary results should be compiled to show puck placement, variable effect, or probability ratings in the mode position. Because the model is designed to give everyone an equal say, mode is the best team representation, as it shows the majority opinion of all of the individuals on the team. In addition, the ability to highlight the puck or probability bubble could be made available in order to flag inputs with individual inputs that are greater than two bin placements away from the mode, which represents extreme opinions. These highlighted inputs would be considered points of conflict within the model. Examples of flagged variables in Figure 2c include the Munitions Performance in Weather puck and the Crew's Rhythm Ops Tempo probability bubble. The point is to show that there may be a significant level of conflict across team members regarding this placement, which may generate dialogue between the commander and his staff. At any point, however, the commander can go forward with the model to see the resulting recommendation.

6 PRIM Computation Walkthrough

After gathering experts' knowledge, a computer-aided decision support tool was designed to extract information from the various inputs and generate a recommendation. To this end, a basic utility function is applied to model a solution generator. Utility is taken from the scale of how the experts rank the importance of a decision, because as experts, their ranking of importance can be construed as how critical the variable is within the alternative. Utility is gathered from the ranking matrix, which is divided into ten slices (each bin divided in half) on a scale of 0.05 to 0.95 with 0.10 intervals; see Table 1 for the specific utility assignments. The utility is then given a positive or negative sign based on the variable designation on the puck. The subjective probability is taken from the likelihood classifications, which are given values of 0.05, 0.25, 0.50, 0.75, and 0.95, along a scale from very unlikely to very likely. The bounds for the utility and probability are 0.05 and 0.95, as opposed to 0 and 1.0, because of the previously discussed bias that people tend to anchor on the mean and maximums and minimums and then adjust from there. Using these values for the extremes "essentially admits that there is a remote possibility that the uncertain quantity could fall beyond these assessed points," but it is remote (Clemen, 1996).

Table 1: Utilities assigned to each bin.

Importance Bin	Bin Portion	Utilitity
High	Upper	0.95
	Lower	0.85
Medium-High	Upper	0.75
	Lower	0.65
Medium	Upper	0.55
	Lower	0.45
Medium-Low	Upper	0.35
	Lower	0.25
Low	Upper	0.15
	Lower	0.05

In addition to the decision of how to scale the utility and probability values, consideration was given to weight the subjective utility higher than the subjective probability in the final calculations. In general, humans, and especially experts, have a better understanding of utility and classifying importance than probabilities, because they develop the knowledge to internally know the importance and key points within their decisions. The knowledge required for determining likelihoods of occurrence has many external influences that these experts may not entirely understand. In order to represent this in the model, because the probability is a decimal value between 0.05 and 0.95, it is raised to 1.5 before multiplying it by the utility to generate a score. Raising it to 1.5 causes it to affect the model less than the utility. Thus, the following equation is used to generate an overall score for each alternative that PRIM assesses:

$$Score = C \bullet \sum_{i=1}^{j} [\, p(x_i)^{1.5} \bullet U(x_i)] \qquad \text{(Equation 1)}$$

where i is the puck under consideration, j is total number of different pucks for x_i through x_j, $p(x_i)$ is the likelihood (probability) rating of that puck variable occurring, $U(x_i)$ is the importance (utility) rating of the puck's variable in the decision, and it is multiplied by C (a constant) to normalize the score to a scale of -100 to 100.

Using Equation 1 and a pilot study of the previously described sample scenario, scores were generated for the proposed courses of action of either staying with the current timeline or delaying the timeline (9.788 and 4.830, respectively). While these numerical solutions indicate a quantitative ranking for the alternative, they do not convey all the information available from PRIM beyond the fact that the first course of action scored slightly higher than the second on a scale of -100 to 100. In order to convey to decision makers not only how much one solution differed from another but why, a visualization was created, which will be described in the next section.

7 PRIM Output Display

To convey further information, an output configural display was designed. The Score and Conflict Square (SCS), Figure 5, is meant to provide a pictorial display of the score as it relates to all the alternatives in a decision. Moreover, a critical part of this display that goes beyond the quantitative ranking is the ability to depict conflict within that alternative. Conflict exists when a variable is rated by any team member greater than two bin placements away from the team's mode rating, and a consensus has still not been reached. The display moves beyond simply giving a numerical score for each alternative, and allows decision makers to quickly view qualitatively how each alternative's score compares to the other alternatives' scores and the level of conflict within each alternative.

The higher the score is, the more the score triangle fills with green. If it is the best score, it is completely full. For the other alternatives, it fills to an equivalent area of the ratio of the alternative score and the best score. Similarly, the conflict triangle fills up with a yellow triangle which represents increasing conflict. It is filled based on a ratio of how many of the options are left in conflict on the model (Figure 2c). If there are 15 variables (each having a utility, effect, and probability assigned), there are 45 possible options that could be in conflict. If 15 of these are left in conflict, then a third of the conflict triangle area will fill with a yellow triangle. There will also be a scroll over function, so the user can check what percentage of conflict exists. Figure 6a shows an SCS that has the best score and very little conflict, while 6b shows an SCS with a score that is a third as good as the best score and approximately a third of its variable options in the model are left in conflict.

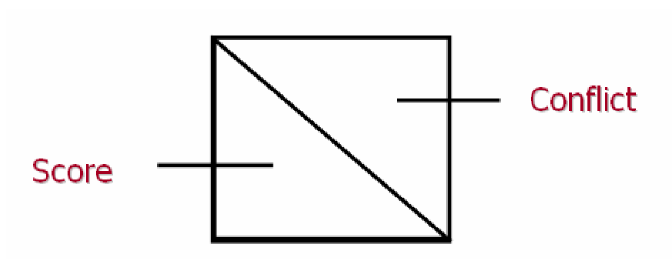

Figure 5: The Score and Conflict Square (SCS).

(a) **(b)**

Figure 6: (a) An SCS for an alternative with the best score (full green triangle) and very little conflict (yellow triangle), and (b) An SCS for an alternative with a score only a third as good as the best score and approximately a third of its model's variable options in conflict.

The final output of PRIM could be given on a display in a fashion similar to Figure 7. This now gives a quantitative score with the best option highlighted, and a qualitative representation of how it relates to the other scores and how much conflict still remains in the model.

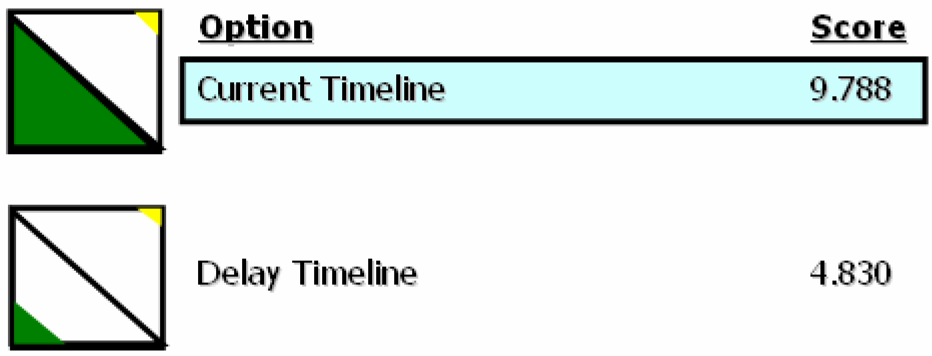

Figure 7: PRIM results page.

8 Conclusion

PRIM is a design of a possible decision support tool in response to a growing problem for commanders in today's battle spaces: attempting to gain consensus from a group for a possible course of action. Indeed, this problem is also an issue for other commercial and industrial domains, where distributed groups of people need to make time-critical decisions. We propose that it is possible to have experts input their knowledge in a story-telling environment that can then be used to extract specific values for a quantitative model that powers a qualitative display.

Other benefits recognized in this research are that while an interface like PRIM may not be readily applicable in giving a commander a comprehensive final solution, it can help to generate and focus discussion within a staff. As previously mentioned, it allows all group members the ability to voice their options and share knowledge. Because of both the individual and team compilations, the commander is given information on what the staff thinks are the most important variables involved in the decision and which variables have the greatest amount of uncertainty in effect or outcome. Thus, PRIM can guide and encourage informed negotiation.

Acknowledgments

This research was sponsored by the Office of Naval Research through a small business technology transfer program lead by Perceptronics Solutions, Inc. The views and conclusions presented in this paper are those of the authors and do not represent an official opinion, expressed or implied, of the Office of Naval Research or Perceptronics Solutions, Inc.

References

Clemen, R. T. (1996). *Making hard decisions: An introduction to decision analysis* (2nd ed.). Pacific Grove, CA: Brooks/Cole Publishing Company.

Gigerenzer, G., Todd, P. M., & Group, T. A. R. (1999). *Simple heuristics that make us smart*. Oxford, NY: Oxford University Press, Inc.

Kahneman, D., & Tversky, A. (1973). On psychology of prediction. *Psychological Review, 80,* 237-251.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*(2), 263-292.

Klein, G. (1999). *Sources of power: How people make decisions*. Cambridge, MA: The MIT Press.

Klein, G., Phillips, J. K., Rall, E. L., & Peluso, D. A. (2004). *A data/frame theory of sensemaking*. Fairborn, OH: Klein Associates Inc.

Oskamp, S. (1965). Overconfidence in case study judgments. *Journal of Consulting Psychology, 29*, 261-265.

Sanders, M. S., & McCormick, E. J. (1993). *Human factors in engineering and design*. New York: McGraw-Hill.

Slovic, P., Fischhoff, B., & Lichtenstein, S. (1978). Behavioral decision theory. *Annual Review of Psychology, 28*, 1-39.

Wickens, C. D., Lee, J. D., Liu, Y., & Becker, S. E. G. (2004). *An introduction to human factors engineering* (2nd ed.). Upper Saddle River, New Jersey: Pearson Education, Inc.

Wiener, E. L., & Nagel, D. C. (1988). *Human factors in aviation*. San Diego, CA: Academic Press.