Comparing Uncertainty Visualizations for a Dynamic Decision-Making Task

Ann M. Bisantz
Dapeng Cao
Michael Jenkins
Priyadarshini R. Pennathur
University at Buffalo, State University of New York

Michael Farry
Charles River Analytics

Emilie Roth
Roth Cognitive Engineering

Scott S. Potter
Jonathan Pfautz
Charles River Analytics

ABSTRACT: Supporting complex decision making requires conveying relevant information characteristics or qualifiers. The authors tested transparency and numeric annotation for displaying uncertainty about object identity. Participants performed a “missile defense” game in which they decided whether to destroy moving objects (which were either threatening missiles or nonthreatening birds and planes) before they reached a city. Participants were provided with uncertain information about the objects’ classifications. Uncertainty was represented through the transparency of icons representing the objects and/or with numeric annotations. Three display methods were created. Icons represented the most likely object classification (with solid icons), the most likely object classification (with icons whose transparency represented the level of uncertainty), or the probability that the icon was a missile (with transparency). In a fourth condition, participants could choose among the representations. Icons either were or were not annotated with numeric probability labels. Task performance was highest when participants could toggle the displays, with little effect of numeric annotation. In conditions in which probabilities were available graphically or numerically, participants chose to engage objects when they were farther from the city and had a lower probability of being a missile. Results provided continued support for the use of graphical uncertainty representations, even when numeric representations are present.

KEYWORDS: decision making, displays, uncertainty

ADDRESS CORRESPONDENCE TO: Ann M. Bisantz, University at Buffalo, SUNY, Industrial Engineering, 438 Bell Hall, Amherst, NY, 14260-2050, bisantz@buffalo.edu.

Journal of Cognitive Engineering and Decision Making, Volume 5, Number 3, September 2011, pp. 277-293. DOI: 10.1177/1555343411415793. ©2011 Human Factors and Ergonomics Society. All rights reserved.
Introduction

An important question in the design of displays to support decision making and performance in complex environments is how to convey relevant information characteristics or qualifiers. The qualifiers, or meta-information (Bisantz et al., 2009), can include characteristics such as the age of information, its source, or its associated degree of uncertainty. For instance, in military command-and-control tasks, individuals may have to rely on information from sensors that have known uncertainties, from human sources of varying reliability, and that may be hours or days old.

Researchers and designers from the scientific visualization and geographical information research areas have suggested a variety of methods representing one form of meta-information: uncertainty (Bisantz et al., 2009). These methods include those that are intrinsic (an integrated component) or extrinsic (an annotation) to the information being displayed (Gershon, 1998).

For instance, intrinsic representations include the use of graphical variables, such as aspects of color (e.g., hue, saturation, brightness; Howard & MacEachren, 1996; Pfeiffer, 2002; Slocum, Cliburn, Feddema, & Miller, 2003), blurriness (Bisantz, Marsiglio, & Munch, 2005; Finger & Bisantz, 2002; MacEachren, 1992), animation (Brown, 2004; Ehlschlaeger, Shortridge, & Goodchild, 1997), texture (Botchen, Weiskopf, & Ertl, 2005; Interrante, 2000; Uness, Interrante, Marusic, Longmire, & Ganapathishubramani, 2003), transparency (Drecki, 2002; MacEachren et al., 2005), or object size and shape (Aigner, Miksch, Thurnher, & Biffl, 2005; Pang, Wittenbrink, & Lodha, 1997; Wittenbrink, Saxon, Furman, Pang, & Lodha, 1996). For example, Slocum et al. (2003) represented more uncertain geospatial information about water availability with lighter colors, as did Aerts et al. (2003) on a display showing urban growth predictions.

Some studies have evaluated the impact of these representations. For instance, Bisantz et al. (2005) and Finger and Bisantz (2002) studied the use of blurred and degraded icons to represent object identity. In both cases, objects could assume one of two categories (e.g., a hostile or friendly aircraft, or a high-quality or low-quality stock). Probability of category membership was represented by the degree to which the icons were crisp representations of the category (e.g., an upward arrow for a high-quality stock) or were blurred and combined representations (e.g., an overlay of blurred upward- and downward-facing arrows to show a 50% chance of being in either category). Finger and Bisantz found that participants could sort, order, and rank icons created using this blurriness scheme.

Additionally, both Finger and Bisantz (2002) and Bisantz et al. (2005) tested the use of different uncertainty representations in dynamic tasks in which participants had to use the information to either identify the objects as hostile or friendly, or to make stock purchases. Other representations, such as linguistic expressions, numbers, colored icons, and blurred or solid icons annotated with numeric probabilities were also tested. Results from these studies showed few differences in performance attributable to uncertainty representation, indicating support for the use of graphical representations of uncertainty in category membership.
Nadav-Greenberg, Joslyn, and Taing (2008) used hue to indicate predicted wind speeds (either maximum or ranges, along with median values) on a map and compared these representations to box plots. In the conditions that involved hue, participants had to reference two maps (a map showing median speeds and one showing ranges). In the box plot condition, the boxes served as a graphic that combined both range (extent of the box) and median speed (center point). However, the color maps allowed participants to see range and magnitude data across geographic regions, whereas the box plot data applied only to the point where the box was located.

The research showed that both novice and expert participants were able to judge the level of uncertainty in the forecast (defined as the range in wind speed values) using the color representation. Participants were better at matching the need for wind speed advisories to predicted speeds when they used the box plot representation. The authors also suggested that because participants were less confident in their decisions when using the color representations of ranges, such displays might be useful in reducing overconfidence.

Bisantz et al. (2009) also studied the use of color to display uncertainty about regions on maps. They found that participants could reliably rank order and rate the level of uncertainty (e.g., about the chance of a thunderstorm) associated with regions when uncertainty was coded by saturation, brightness, and transparency. Additionally, they found that the mapping of levels of uncertainty to levels of the color variable depended not only on the contrast of the colored regions with the background but also on the relevance of information to the task. Thus, in a task in which it was important to highlight regions of uncertain information, participants would assign that region a color saturation that had high contrast with the background (e.g., bright green against a gray background or gray against a bright green background). If more-certain information was more relevant to the task, the order of assignments would be reversed.

Thus, although a variety of methods for representing uncertainty have been implemented by researchers and designers in fields such as geographic information science, statistics, and scientific visualization, evaluation of these concepts, particularly, their impact in dynamic decision-making tasks, has been more limited. The present research extends these studies by using transparency to represent category membership in a dynamic task. Whereas past research studied techniques that blurred and combined icons to represent membership in one of two categories, this study compares strategies for representing uncertainty about an object’s identity when there are more than two potential classifications of the object.

Additionally, the study compares representations that do and do not include numeric annotations that specify the probability of category membership. Numeric probabilities provide precise, rather than vague, information about uncertainty (Wallsten, 1990) and therefore could be thought to provide the most information to decision makers. Some previous research, however, has suggested that providing numeric uncertainty representations may not lead to improved performance compared with some graphical representations (blurred
or degraded icons representing category membership; Bisantz et al., 2005; Finger & Bisantz, 2002), perhaps because people use similarly vague or fuzzy mental representations of both concepts (Bisantz et al., 2005; Wallsten & Budescu, 1995). Numeric representations also may increase information density on the screen, leading to performance decrements attributable to clutter, and may require more attentional resources than an integrated representation (Wickens & Andre, 1990).

This research explicitly compared representations that combined graphical and numeric representations of uncertainty with conditions in which uncertainty was represented only graphically or only numerically to provide further empirical evidence about the value of graphical compared to (or in combination with) numeric uncertainty display. Finally, this research investigates user strategies when they can choose among different representations.

**Method**

**Participants**
Participants with self-reported normal or corrected-to-normal vision were recruited from the university community and were screened with the Ishihara test for colorblindness (2 participants failed the screening and were excused). Ten female and 14 male participants ranging in age from 19 to 32 (mean of 25.25) completed the study. Participants were compensated and a bonus was offered to the highest-scoring participant in each between-subjects condition to provide motivation.

**Experimental Stimuli**
Participants were shown an 800 × 600 pixel area on a 19-in. LCD monitor (set to 1,280 × 1,024 pixel resolution) that displayed task resources and score information as well as a number of moving 20 × 24 pixel icons representing either missiles, planes, or birds. Icons provided information regarding the classification of an object into one of the three categories. This information was probabilistic, emulating a situation in which sensors provide uncertain (and sometimes inaccurate) information regarding object classification. The format of the icons depended on the experimental condition. Additionally, because the probabilities were dynamic, icons changed throughout the task.

**Independent Variables**
Two independent variables were manipulated to display the classification probabilities: display format (within subjects) and probability visibility (between subjects).

*Display format.* There were four display format conditions corresponding to three methods for displaying entities and one condition in which participants could switch between the three methods during the task (see Figure 1). The display
conditions varied the method by which the classification probability was shown. In
the first condition (labeled most likely solid), entities were displayed with an opaque
icon representing the category for which the probability of classification was high-
est. For instance (as shown in Figure 1), if an entity (at a particular point in time)
had an 80% chance of being a plane, a 10% chance of being a bird, and a 10%
chance of being a missile, an opaque plane icon would be shown.

In the second condition (most likely transparent), entities were again displayed
with an icon representing the category for which the probability of classification
was highest; however, the transparency of the icon reflected the probability. As the
probability decreased, the icon became more transparent (effectively, less saturated
against the background). Transparency was varied corresponding to changes in
probability such that the level of probability corresponded to the percentage opac-
ity (lower probabilities were more transparent). Transparency ranged from 0% to
99% (changing by single percentages), corresponding to changes in probability. In
previous research, participants could successfully rank order and rate regions with
a similar range (but fewer levels) of transparency (Bisantz et al., 2009).

For the third condition (missile transparent), all entities were displayed with
the missile icon (the most threatening category, in the context of the task), with
the level of transparency corresponding to the probability that the icon was clas-
sified as a missile. Thus, for the entity shown in Figure 1, the icon would reflect,
across the three conditions, either an opaque plane, a slightly transparent plane,
or a very transparent missile. For the fourth display condition (toggle), partici-
pants could switch among the three other conditions as they wished throughout
the task by pressing a key.

**Numeric probability visibility.** Icons either were or were not annotated with a
number showing the probability that entities were classified as the category
shown by the icon. Thus, in Figure 1, the probability shown for most-likely-solid and most-likely-transparent conditions is $p(\text{plane})$, whereas for the missile-transparent condition, it is $p(\text{missile})$.

**Experimental Design**

Display format (within subjects) and probability visibility (between subjects) were crossed to form eight conditions. Twelve participants were randomly assigned to each probability visibility condition. We balanced display format order using a modified Latin square design, with the exception that all participants performed the toggle condition last (to assess their use of different options after they had experienced all three display manipulations). Participants performed two trials in each display condition before moving to the next display condition.

**Experimental Tasks and Procedure**

The experimental task consisted of a “missile defense” game in which participants selected and destroyed entities (with uncertain identities as planes, birds, or missile) that were approaching a city. The goal of the game was to destroy only those entities that were truly missiles, while not allowing missiles to reach the city. Note that the game was designed to test the uncertainty display concepts described previously and was not an analog to a realistic missile defense scenario. For instance, the only information participants had about object identity was the icon; other information, such as object speed or origin, was either randomly assigned or unavailable.

Eight scenarios corresponding to the eight trials were created. For a pair of scenarios corresponding to one display condition, the number of “true” missiles was controlled to be 100 (out of the total entities across the two trials). Between 30 and 70 missiles were assigned (randomly) to one scenario, and the remainder (out of 100) were assigned to the second. The remaining entities in each scenario were randomly assigned to be either birds or planes. There were between 117 and 140 entities in each scenario. Entities entered the screen at randomly assigned times throughout the scenario, from random locations around the edge of the screen. Entities traveled at one of five randomly assigned speeds in a straight path from their entry point, toward the city at the center of the screen, and then (for birds and planes) off the screen. Missiles that passed over the city (i.e., those that were not destroyed) exploded over the city. Speeds were unrelated to the true entity type (e.g., a missile did not always travel at a particular speed). The fastest entities took 5 s to cross the screen and reach the city, and the slowest took 20 s.

Regardless of the true identity of entities, each entity was randomly assigned an initial classification probability of being a missile, plane, or bird (probabilities summed to 1.0). These probabilities changed as the entity moved, changing at five evenly spaced times (intervals were dependent on entity speed) before the entity reached the city. Thus, depending on the display condition, the icons, transparency, and/or numeric annotations associated with an entity changed as the probabilities changed. At each interval, the classification probability that the
entity was its true identity had a 70% chance of increasing and a 30% chance of decreasing, a probability between 0.1 and 0.15. Thus, over time, the classification probabilities tended to reflect the true identity but with some randomness.

Participants were told during training that the information tended to improve over time, as it could in a real task when more information was collected. The scenarios were scripted such that classification probabilities were capped at 0.00 and 0.99 and so that all participants saw entities with the same behaviors.

An overall score as well as score components were shown on the screen. The score was composed of points for destroying a missile (200 points) and penalties for destroying a bird or plane (30 points) or allowing a missile to reach the city (60 points). Because the number of missiles differed in each scenario, the maximum possible score differed. Therefore, scores were normalized with respect to the maximum score for analysis.

To perform the task, participants used a mouse to point at entities they believed to be threatening and used the left mouse button to fire at the entity. The number of shots was limited (equal to 1.5 times the number of missiles in a scenario), and the number of shots remaining was displayed at the bottom of the screen. A “reload” delay of 2.0 s was imposed before another shot could be fired. The time remaining until participants could fire was indicated graphically with a shrinking bar at the bottom of the screen. Firing at an entity did not always destroy it: The probability of a successful hit decreased as the entity moved closer to the city. If a shot was successful, the entity disappeared; if it was unsuccessful, the word miss appeared next to the entity.

The experiment took about 1.75 hr, on average. After participants gave informed consent and were screened for color deficiency, they were given a demonstration of the task and a description of the four conditions and the task scoring scheme. They were told that more transparent icons indicated more classification uncertainty. Participants were told that their goal was to maximize the score and that icons showed information regarding the classification probability of an entity but that the information was uncertain and could change. They were told there was sufficient ammunition to engage all of the missiles (but not every entity) but were not told the specific proportion of actual missiles, birds, and planes in the scenarios. They then completed a short training scenario in which they could toggle among display conditions and practice the task. All participants performed eight task trials (two in each of the four display format conditions). After each trial, they completed the NASA Task Load Index (NASA-TLX; Hart & Staveland, 1988) subjective workload assessment and were offered a short break.

The task software automatically logged participant actions with respect to the entities, type of icon displayed (for the toggle condition), and the probabilities and locations of the entities at the point participants fired at the entities.

Results

A repeated-measures analysis was performed with display format (four levels) as a within-subjects factor and probability visibility (two levels) as a
between-subjects factor. Four dependent variables were analyzed: normalized score, NASA-TLX score, probability of missile at point of engagement, and distance at point of engagement.

**Normalized Score**

There was a significant effect of display format on score, $F(3, 138) = 3.53, p = .017$. Participants generally performed better with the most-likely-transparent and toggle conditions (see Figure 2). Post hoc tests (with Bonferroni correction) indicated a significant difference between the toggle and most-likely-solid condition (two-tailed $p = .007$). There was no significant effect of probability visibility or any significant interaction.

**Subjective Workload**

We assessed subjective workload using the NASA-TLX with subscales weighted evenly. Workload was higher for the toggle and most-likely-solid conditions, but the effect was not significant, $F(3, 138) = 2.437, p = .067$. Although workload was lower when numeric probabilities were visible, the effect was not
significant, $F(1, 46) = 1.847, p = .181$; the interaction was also not significant. (See Figure 3.)

Further analysis of the individual subscales did show some significant effects of display condition on the Physical Effort subscale, $F(3) = 4.317, p = .006$, with the toggle condition requiring the highest physical effort. There was no effect of probability visibility on any of the individual subscale ratings.

**Probability of Missile at Time of Engagement**

We recorded the probability of an entity's being classified as a missile at the time participants first fired at it to understand the effects of the display manipulations on when participants chose to engage a target (Figure 4). Note that if a participant shot at the entity and missed it, additional attempts could be made. Only the first engagement was analyzed because it reflected participants’ decision making regarding an entity’s identity.

There were 8,793 instances of engagements analyzed across all conditions. Because the data did not meet assumptions for normality, a Friedman test was performed separately on each between-subjects group (numeric probability visible group, 4,395 instances; numeric probability hidden group, 4,398 instances) as well as to compare the numeric probability visibility conditions. There was a significant effect of display format for both the probability visible group, Friedman test $\chi^2(3) = 44.95, p \geq .000$, and the probability hidden group, Friedman test $\chi^2 = 43.22, p \geq .000$. In both cases, the probability the object was a missile was greatest in the most-likely-solid group. There was no significant difference between the probability visibility conditions.
Distance From City at Time of Engagement

We also recorded the location of entities at the time they were first engaged, measured as the distance from the city (center of the screen; see Figure 5). Similar to probability of a threat, the data did not meet assumptions of normality, so non-parametric tests were used. There was a significant effect of display format only for
the probability visible group, Friedman test $\chi^2 = 7.66, p = .054$. Additionally, there was a significant effect of probability visibility, Friedman test $\chi^2 = 17.34, p \geq .000$. In general, participants fired at entities sooner (when they were farther from the city) when numeric probabilities were displayed; however, this effect was reduced when entities were displayed with solid icons.

**Display Choice at Engagement**

Finally, to understand participants’ choices in the last condition, when they could toggle among display conditions, we examined the display condition selected when entities were engaged (for all times entities were engaged; see Figure 6). A two-way $\chi^2$ test indicated that there was a significant impact of probability visibility condition on display selection, $\chi^2(2) = 392, p \geq .00$. Participants for whom numeric probabilities were visible were more likely to engage entities when they were displayed as missiles, whereas those without probabilities visible engaged entities in the most-likely-transparent condition.

**Discussion**

In this research, we extended findings from previous studies by using transparency to represent uncertainty in object classification and by exploring methods for representing classification uncertainty across more than two categories. Additionally, we investigated the effect of including numeric annotation along with a graphical representation of uncertainty and the impact of allowing participants to change the display.
Analysis of score across display conditions showed that when participants could toggle between display conditions, they performed best. Performance in this condition was significantly better than in the condition in which probabilities were not represented graphically (i.e., the most-likely-solid condition). When participants could toggle between conditions, they engaged fewer entities when viewing the most-likely-solid representation, indicating a preference for more uncertainty information.

Addition of numeric probability information did not improve overall task score performance. This finding is consistent with results of Finger and Bisantz (2002) and Bisantz et al. (2005), who showed that participants who used graphical representations generally performed at least as well those who saw numeric representations, and Bisantz et al. (2009), who showed that participants could successfully rank order and rate regions according to levels of uncertainty (and other forms of meta-information), which were represented by color saturation, brightness, or transparency.

In this study, however, there was evidence that participants’ strategies with respect to firing on entities were influenced by the form of representation. Specifically, participants who saw graphical information regarding probabilities (or who could toggle to do so) fired on entities with less evidence (lower probability) that they were missiles. Because the probabilities associated with entities tended, with some randomness, to approach the actual identity of an entity as the track moved across the screen, this finding suggests that in the most-likely-solid condition (the condition with no graphical representation of uncertainty), participants waited to fire—perhaps to see whether the icon would shift from displaying one type of entity to another (e.g., shift from a solid missile to solid plane icon). In contrast, when they had some information regarding the probability associated with an entity, they fired sooner (i.e., at a lower probability).

Because all entities moved from the edge to the center of the screen (the city), we would expect results regarding the distance from the city to be consistent with the level of probability results. For the visible numeric probability condition, this was the case. When the icon was solid, participants waited longer (until the object was closer to the city) than when the icon was transparent. However, without numeric annotation, participants fired when entities were closer to the city than with numeric annotation, across all display conditions. This difference (compared with the probability measure) may be attributable to the fact that distance changed continuously, whereas probabilities changed in discrete steps. When we considered both probability and distance to the city, it appears that participants were less conservative (fired sooner) when they had more information about the probabilities.

One explanation is that with more information, participants could detect more precisely the point at which the probability an entity’s missile classification reached a particular threshold. Note that depending on the task context, the “correct” choice to take action with more- or less-certain information could vary. That is, in some circumstances, it may make sense to act sooner (e.g., if there is
Uncertainty Visualization for Dynamic Decision Making

Results indicated that even when they had access to numeric probabilities, participants observing icons that did not graphically indicate the level of uncertainty had a lower score, engaged entities with a higher probability of being a missile, and engaged entities when they had moved closer to the city than when icon transparency indicated the uncertainty level (or when they could toggle to a condition that did). Also, overall workload was higher (although not significantly) for the most-likely-solid condition compared with either condition in which uncertainty information was presented by transparency, even when numeric probabilities were present. In the toggle condition, participants still appeared to prefer a condition in which uncertainty information was represented graphically, even when numeric probabilities were present. These results, combined with those from past studies, indicate that graphical representations of uncertainty are useful and may provide information beyond that provided by numeric annotations.

Researchers have studied graphical representations of uncertainty in classification when an item could belong to two categories. We extended that work by creating novel display methods for uncertainty classification when more than two categories were present. Participants were shown either the probability that an entity belonged to one category or the most likely category of an entity. This method could be extended to support additional categories beyond the three tested here. The condition in which participants could toggle among conditions supported the best task performance, suggesting that providing such flexibility to users would be beneficial.

It is reasonable to consider, for instance, situations in which users at one moment may be interested in the probability that all entities were threats (therefore highlighting the most or least threatening objects, depending on the mapping of salience to uncertainty; Bisantz et al., 2009) whereas at another moment may be interested in the most likely categorization for the objects being viewed. Although participants in the toggle condition rated workload as somewhat higher, this effect was not significant and did not adversely affect performance compared with the other conditions.

Because the toggle condition was always performed last, we further investigated score to see whether the advantage for this condition could be attributable to practice. Mean scores across the four trial blocks (in order) were 0.41, 0.40, 0.42, and 0.45. A regression analysis of score by trial across all eight trials was significant, $F(1, 191) = 6.822, p = .010$, but was not significant when the two toggle trials were removed, $F(1, 143) = 1.47, p = .227$. Thus, participants performed similarly for the first three blocks and then improved on the fourth (toggle) block. This evidence does not indicate a consistent effect of practice across all four blocks, although it is possible that there was some cumulative effect of three blocks of practice that resulted in better performance on the fourth block.
There was some indication that the second trial in each block was better than the first for all four blocks, $F(1, 191) = 3.73, p = .055$, but this did not translate to improved performance when switching display conditions. Also, because trials were constructed in pairs (the number of birds, planes, and missiles were held constant across the pair of trials), the nature of the trial may have contributed to this effect.

We did not include a condition in which participants could see only numeric indicators (i.e., either without icons or associated with a consistent icon regardless of likely object type) or could toggle numeric probability labels on or off. These configurations could be tested in further research. On the basis of these results, however, it seems likely that including numeric labels can provide value when combined with graphical uncertainty representations. If screen clutter or information density is a concern, providing users with the ability to turn labels on and off may be a solution if uncertainty information is also shown graphically. We also did not vary the number of uncertainty levels, or granularity, with which the information was presented to participants. It is possible that preference for different representations (i.e., as expressed through the toggling condition or choice of numeric annotation) may depend on how many categories of probability information are possible and thus must be remembered and distinguished.

We also did not evaluate how representational choice may interact with the nature of the probabilities displayed. Testing more specifically how different representations affect decision making when the underlying uncertainty has different characteristics (e.g., category probabilities are very high or very low) would be an interesting extension to this research. For instance, research has indicated that in some conditions, people tend to overweight small probabilities (Kahneman & Tversky, 1984). Graphical representations could be developed to explicitly counteract such biases, for instance, by changing the degree of transparency so that the degree of change in transparency is increased for very low probabilities.

Finally, although some aspects of the current task were based on realistic operational constraints (such as time pressure, the fact that defense actions may not always be successful, and classification probabilities that change because sensors gain more reliable information over time), we did not attempt to mimic an actual missile defense situation or environment. In such settings, for instance, object speed would be a useful indicator of identity; automation would be used to pre-filter unimportant objects, and the costs and benefits of correctly or incorrectly engaging objects would be different, and more extreme, than the scores provided in the game. Additionally, we did not test displays with domain experts. Future work should address the implementation and testing of these display concepts in studies that are closer to operational environments.

Conclusions

This study extended past research by investigating the use of transparency representations on decision making in a dynamic task and by comparing different schemes for representing uncertainty in object classification. Results provided
continued support for the use of graphical uncertainty representations, even when combined with numeric representations, and also for systems that allow users to select among display methods. Future work should extend these results from controlled tasks to more operational settings and experienced users.

Acknowledgments

This work was sponsored by the Missile Defense Agency and was performed under Government Contract No. W9113M-07-C-0172.

References


**Ann M. Bisantz** is a professor of industrial and systems engineering at the University at Buffalo, State University of New York. She conducts research in areas of uncertainty representation, decision making, interface design, and cognitive engineering methods in a variety of domains, including military command and control, health care, and emergency management. She received her PhD from the Georgia Institute of Technology.

**Dapeng Cao** is a PhD candidate in the department of industrial and systems engineering at the University at Buffalo, State University of New York. He conducts research in knowledge representation, decision support, simulation and optimization in the domains of health care, and command and control. He received his MS in Human Factors from the University at Buffalo.

**Michael Jenkins** is a research assistant and PhD candidate at the University at Buffalo, State University of New York, pursuing his Ph.D. in Human Factors Engineering. His research interests include complex systems design, human-machine trust, decision making, and ecological interface design. He received his M.S. in Human Factors in Information Design and MBA from Bentley University.
Priyadarshini R. Pennathur is currently a post-doctoral research fellow in the department of anesthesiology and critical care medicine at the Johns Hopkins University School of Medicine. Her research interests are in cognitive engineering, human factors in healthcare and patient safety. She received her PhD in Industrial and Systems Engineering from the University at Buffalo.

Michael Farry is a senior scientist at Charles River Analytics Inc. There he manages and leads project teams that blend well-grounded cognitive engineering techniques, innovative design ideas, state-of-the-art software engineering practices, and thorough evaluation methods to deliver technology that is effectively and soundly tailored to user requirements. He received MEng and SB degrees from the Massachusetts Institute of Technology, where his research focused on the construction of social networks from sensor network data.

Emilie Roth is the principal scientist of Roth Cognitive Engineering. She has been involved in cognitive systems analysis and design in a variety of domains, including nuclear power plant operations, railroad operation, military command and control, medical operating rooms, and intelligence analysis. She is a Fellow of HFES, serves on the editorial board of the journal Human Factors, and is associate editor of the Journal of Cognitive Engineering and Decision Making. She received her PhD from the University of Illinois.

Scott S. Potter is a senior systems engineer at General Dynamics C4 Systems | Viz. Since joining Viz in 2010, he has been co-principal investigator on the DARPA PAL Program, including leading a series of Double Helix exercises to identify directions for increased decision support. He has over 20 years of experience in the design of decision support tools for complex, dynamic, high-risk environments for a variety of domains including military command-and-control work, and intelligence analysis. He received his PhD from Ohio State University.

Jonathan Pfautz is a principal scientist and vice president, Cognitive Systems at Charles River Analytics Inc. in Cambridge, Massachusetts. His major research interests include cognitive and perception-based visualization and user interface design, cognitive work analysis, social network analysis, human behavior modeling, and the applications of cognitive engineering to software system development across a variety of military and nonmilitary domains. He received his PhD from Cambridge University.