15 Interactive data visualization for mitigation planning: Comparing and contrasting options[1]

Beth Yost, The MITRE Corporation

1. Introduction

Interactive data visualization can greatly assist decision makers in comparing and contrasting their options as they choose a course of action for mitigation. Admittedly, ineffective visualizations can lead to poor decisions (consider, for example, the many descriptions of unintelligible charts used in the process of deciding whether or not to launch the space shuttle *Challenger* (Tufte, 1997)). By contrast, effective visualizations can lead to a deeper understanding of the robustness of various courses of action, resulting in better outcomes for a greater proportion of decisions. In this chapter, we cover the underpinnings of effective visualizations that help decision makers better comprehend their options. Because using interactive visualization to compare and contrast courses of action represents a nascent area of study, we also discuss the research to date in this area and describe promising directions for future investigation.

1.1. Definition

In essence, visualization for mitigation means supporting option awareness. In this chapter, we use the definition of option awareness provided by Drury, Pfaff, Klein, and Liu (2013, p. 658): "… individual and team decision makers develop *option awareness* (OA): the perception and comprehension of the relative desirability of the available options, as well as the underlying factors, trade-offs, and tipping-points that explain that desirability…" Option awareness assists users in selecting a good course of action by creating an understanding of how effective that action will be under various conditions. This requires forecasting of plausible futures, starting with the particulars of a specific current situation. Ultimately, visualizations that promote option awareness should present the landscape of plausible outcomes for each option in a way that enables the user to easily compare those landscapes and choose the most robust option.

1.1.1. Types of decisions

Most visualization tools today simply present situation awareness data; the developers presume that the user will extrapolate from the details of a situation to decide on a course of action. This leaves all analysis of options and forecasting of outcomes to human cognition, with none of the burden carried by machines capable of analyzing many options quickly. Computer-supported option awareness alleviates this overload by presenting visualizations to augment cognition and increase understanding.

[1] Approved for Public Release; Distribution Unlimited. Case Number 14-2487

As with situation awareness, individuals can achieve progressively higher levels of option awareness, and new interactive visualization tools can give decision makers much broader and deeper views of their options. Figure 1 shows the three levels of option awareness, with each level building on the insight and understanding gained from the previous level. A visualization that enables option awareness level 1 enables users to see and understand the comparative robustness of their options more rapidly. If the visualization enables option awareness level 2, users understand the key trade-offs between those options and the factors that make one option more robust than another. Option awareness level 3, which builds on levels 1 and 2, enables users to propose more effective options based on a deep understanding of relative robustness and the factors driving that robustness.

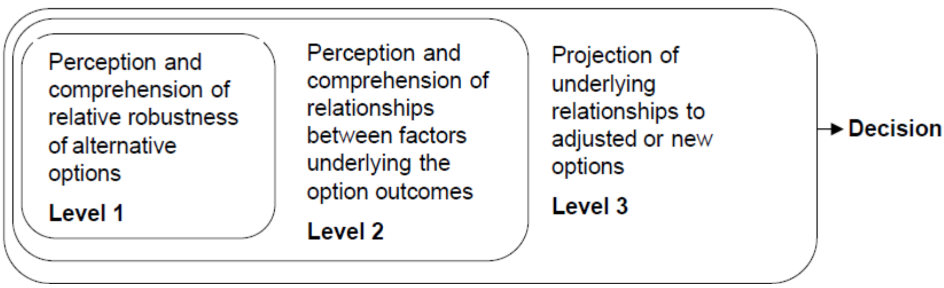


Figure 1. Levels of option awareness. Adapted with permission from Klein, Drury, Pfaff, & More (2010).

1.1.2. Types of data

A key driver for visualization in general and option awareness visualization in particular is the type of data to be visualized. Fundamentally, the option awareness data includes:

- A set of alternative options
- A set of conditions that can vary or are unknown (uncertainty)
- Criteria for evaluating the options
- Evaluation of the options

The evaluations can be generated by computer simulation or by a group of subject matter experts, or can be found from historical data. They result from examining each option across a range of possible conditions that could vary or may be unknown and measuring the outcome in terms of each relevant criterion (e.g., cost or lives lost). The complexity in option awareness data stems from:

- Options that are multi-faceted (e.g., the option has two aspects that may interact with each other in some way)
- Varying or unknown conditions, including aspects of the current situation (e.g., a flu outbreak with unknown lethality) or uncertainty about option effectiveness (e.g., what percentage of people will comply with a request to stay home from work)

Table 1 shows this breakdown for the simple notional dataset of the virus outbreak that we use as an illustration throughout this chapter. In this example, the decision maker must determine how to respond to the outbreak. The option has two components: (a) whether to implement social distancing (i.e., ask schools and workplaces to close) and (b) what percentage of the population to vaccinate (10% vs. 50%). These factors interact with each other because social distancing would result in fewer people becoming exposed to the virus, and thus vaccination would have less impact on the outcome. The uncertainty in this case comes from not knowing the deadliness of the new virus. Here we consider only three possibilities: 30%, 60%, or 90% of people exposed and not vaccinated will die from the virus. A real scenario would probably involve multiple unknown variables (e.g., unknown lethality + unknown percentage of the public that would comply with social distancing) and would require evaluation of tens to hundreds of possibilities (not just three). This would further complicate the data, resulting in an even greater need for effective visualization to aid the decision maker in understanding the robustness of each option and the underlying factors. In the illustrative example discussed in this chapter, the criteria we use to assess the goodness of each option are cost and lives lost.

Table 1. *Notional option awareness data: Most aggregation (top left) to most detailed (bottom right)*

		Criteria **Option**	Single Weighted Score (i.e., Score)	Each Criterion Separately (i.e., Cost, Lives)
Uncertainty	Combined to Show Robustness (i.e., criteria includes set of values, one from each possible Deadliness condition)	Single Name (i.e., Option)	**Option Deadliness Score** OptionN10 <not shown {1, 33, 65} OptionN50 <not shown> {35, 53, 71} OptionY10 <not shown> {81, 85, 88} OptionY50 <not shown> {77, 79, 81}	**Option Deadliness Cost Lives** OptionN10 <not shown> {$34, $58, $82} {243, 486, 729} OptionN50 <not shown> {$63, $77, $90} {135, 270, 405} OptionY10 <not shown> {$37, $40, $43} {27, 54, 81} OptionY50 <not shown> {$76, $78, $79} {15, 30, 45}
		Each Aspect Separately (i.e., Dist, Vacc)	**Dist Vacc Deadliness Score** No 10 <not shown> {1, 33, 65} No 50 <not shown> {35, 53, 71} Yes 10 <not shown> {81, 85, 88} Yes 50 <not shown> {77, 79, 81}	**Dist Vacc Deadliness Cost Lives** No 10 <not shown> {$34, $58, $82} {243, 486, 729} No 50 <not shown> {$63, $77, $90} {135, 270, 405} Yes 10 <not shown> {$37, $40, $43} {27, 54, 81} Yes 50 <not shown> {$76, $78, $79} {15, 30, 45}
	Each Possibility Separately (i.e., Deadliness = 30, 60, or 90)	Single Name (i.e., Option)	**Option Deadliness Score** OptionN10 30 65 OptionN10 60 33 OptionN10 90 1 OptionN50 30 71 OptionN50 60 53 OptionN50 90 35 OptionY10 30 88 OptionY10 60 85 OptionY10 90 81 OptionY50 30 81 OptionY50 60 79 OptionY50 90 77	**Option Deadliness Cost Lives** OptionN10 30 $34k 243 OptionN10 60 $58k 486 OptionN10 90 $82k 729 OptionN50 30 $63k 135 OptionN50 60 $77k 270 OptionN50 90 $90k 405 OptionY10 30 $37k 27 OptionY10 60 $40k 54 OptionY10 90 $43k 81 OptionY50 30 $76k 15 OptionY50 60 $78k 30 OptionY50 90 $79k 45
		Each Aspect Separately (i.e., Dist, Vacc)	**Dist Vacc Deadliness Score** No 10 30 65 No 10 60 33 No 10 90 1 No 50 30 71 No 50 60 53 No 50 90 35 Yes 10 30 88 Yes 10 60 85 Yes 10 90 81 Yes 50 30 81 Yes 50 60 79 Yes 50 90 77	**Dist Vacc Deadliness Cost Lives** No 10 30 $34k 243 No 10 60 $58k 486 No 10 90 $82k 729 No 50 30 $63k 135 No 50 60 $77k 270 No 50 90 $90k 405 Yes 10 30 $37k 27 Yes 10 60 $40k 54 Yes 10 90 $43k 81 Yes 50 30 $76k 15 Yes 50 60 $78k 30 Yes 50 90 $79k 45

We follow visualization expert Ben Shneiderman's visual information seeking mantra of "overview first, zoom and filter, then details on demand" (Shneiderman, 1996, p.336) to make visualizations of the data more manageable for a decision maker. At the highest level, the data presented for each option includes the name of the option and a set of scores for each option. The top left cell in Table 1 shows the highest level of aggregation of the data: an overview. The two aspects of the option (dist and vacc) are combined into a name that describes the complete option (e.g., OptionN10). The two criteria (lives and cost) are combined into a single weighted score. The specific conditions under which each option was tested (i.e., the deadliness percentages of 30, 60, and 90) have been grouped by option, resulting in a set of values that represent the plausible outcomes for that option under uncertainty (the uncertainty about the deadliness of the virus). At the other end of the spectrum, as decision makers drill down to obtain more details, they would move toward the bottom right cell in the table, which shows the highest level of detail for this simple example dataset.

The characteristics of the data constrain the space of relevant visualization techniques. Users can attain option awareness level 1—understanding the robustness of each course of action—even from the most aggregated form of the data, the overview, if visualized effectively. However, displaying only the aggregate data hides the underlying factors that drive robustness. To achieve level 2 option awareness users require a capability to interact with the data, filter the data to show only the most worthwhile options, and drill down into the data to find more details so they can begin to understand the factors driving the outcomes. This helps users to understand the trade-offs: for example, two options may have a relatively similar average outcome, but one has a much better best case and much worse worst case than the other.

1.2. Why Visualization Matters

The visual interface constitutes the last gateway between data about the robustness of the options and the decision maker's mind. A clear presentation of the option awareness data reduces the cognitive burden. However, if the chosen visualization technique obfuscates options and outcomes, decision makers will find it difficult to fully comprehend the resulting analysis of their options.

Designers must create cognitively and perceptually effective visualizations that allow the user to interact with the visualization in appropriate and meaningful ways. Something as simple as the terminology used to convey the data can affect risk perception and the resulting decisions (e.g., due to the framing effect) (Ancker, Senathirajah, Kukafka, & Starren, 2006). Further, if users do not understand the inner workings of a tool, particularly a modeling and simulation tool, they may blindly trust the tool too much or lose trust in the tool altogether (Sheridan, 2002). To realize the powerful benefits of using interaction visualizations to analyze options, new visualizations and visual analytics tools must effectively address these types of cognitive issues.

1.3. What Makes Visualization Difficult

Option awareness visualization comes with its own unique set of challenges that future research must address. At a high level, some of the challenges in this space include:

- Complexity of the data

  - Achieving greater levels of option awareness will require visualizations that provide deeper understanding of the underlying, inherently complex data. How do we effectively portray the robustness of each option? How do we create visualizations that clearly show users how the different factors interact?
- Diverse users without statistics backgrounds
  - As this chapter will show, statistical analysis and concepts such as mean and standard deviation are often used to summarize the set of scores that represent the robustness of an option (e.g., in the example above, OptionN10 has a set of scores {1, 33, 65}). Yet many decision makers do not have a statistics background. How do we create interactive visual analytics tools useful to groups of stakeholders who have diverse backgrounds and experiences? Which information could and should the visualization abstract away?
- Need for effective visualization design principles to enable trustworthy, defensible decision support
  - Much has been written about creating effective situation awareness displays, but little about creating effective option awareness displays. Rigorous design principles will lead to visualizations that effectively support option awareness. Further, in the context of selecting between options, good design principles are important not just for cognitive and perceptual reasons, but also for:
    - Ensuring appropriate levels of trust – While decision makers will not rely exclusively on a visual analytics tool for information, that tool will be one of several sources. Decision makers need a way to determine the reliability of the results and the level of trust to put in what they are seeing.
    - Ensuring ethical designs – Designers must create visualizations that show the data in a way that does not unduly influence the user’s decisions.
    - Ensuring defensible results – Option awareness level 2 gives decision makers the understanding they need to explain or defend the reasons why they chose one option over another.

The remainder of this chapter presents more details on these challenges.

1.5. How Visualization Fits into the Bigger Picture

This chapter centers on visualization for comparing options and gaining understanding of the trade-offs, but comparing alternative options represents only one step in the mitigation process. In reality, visualization can help in other areas related to mitigation; for instance, in generating alternative options, assisting those tasked with executing the selected option, or monitoring the effectiveness of the selected option.

First, a subject matter expert needs to identify and generate alternative options – whether proven or novel – before they can be modeled and compared. Most real-world cases fall into the middle of this continuum. Well-known options typically involve checklists and well-documented processes. Visualization can allow users to compare all of the options and can create the basis for generating new variations. By observing which factors lead to good or bad outcomes, a user acquires the basis for creating new, even more effective options. Visualizations that provide option awareness levels 2 and 3 assist with this process (see sections 3.1.2 and 3.1.3 in this chapter).

Second, once decision makers select a particular action they must prescribe how to execute the action. The individuals responsible for carrying out the action will need detailed instructions. The details must include what to do and how to do it, and may require discussion and revision (e.g., the Postal Service may be asked to deliver vaccines to homes). Those responsible for carrying out the action must also know what resources will be available to support the action (e.g., how much vaccine to deliver) and the scope of the action (e.g., to whom to deliver the vaccines). Visualization could help decision makers communicate this type of information to the action takers.

Third, as the actions take place and the situation evolves, decision makers must assess the results and adapt the response based on that assessment. For example, if people are asked to stay home from work, someone must determine what percentage of the public is complying. The greater the percentage of people who comply with what is asked, the more effective the option. This calls for practices and tools that enable the initial data and the reasons for selecting a particular option to be available when assessing the effectiveness of the actions being taken. This would contribute to effective mitigation because decision makers could better understand how well the situation on the ground matched the predictions of the original models used to choose among options. All of these areas call for research, including work to build a system that provides early indicators of how well the mitigation actions are working.

We limit the scope of this chapter to visualization for comparing the relative robustness of a set of options, understanding the trade-offs between those options, and gaining a deep enough understanding of those factors to generate novel options. The assessment phase represents a full circle back to using visualization for understanding and detection. In the remainder of this chapter, we first briefly describe how decision makers use option analysis today, which leads into the discussion of the state of the art. We then discuss areas that require more research to further advance the state of interactive visualization tools for option awareness.

2. State of the Practice

This section summarizes the primary techniques used for visualization today, highlights some findings, and suggests shortfalls in current capabilities.

2.1. Manual Techniques

Today, most analysts compare different options by hand, perhaps with the assistance of a spreadsheet to capture the data, and then visualize the results in a simple static diagram that can be placed on a presentation slide. For example, in its Analysis of Alternatives (AoA) Handbook (AFMC, 2008) the Air Force includes a section on presenting the results of the AoA, as illustrated in Figure 2. This type of chart is sometimes referred to as a horse blanket because of its colorful design.

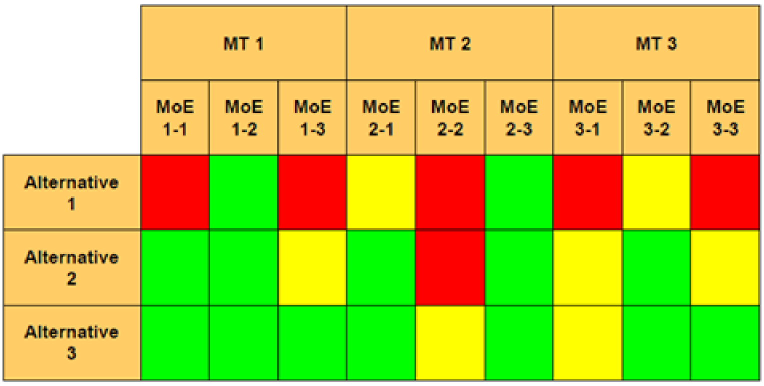


Figure 2. Horse blanket – static visualization for comparing options (AFMC, 2008).

The AoA Handbook uses this horse blanket (or heatmap) visualization approach to compare options ("Alternative 1," etc.) across a number of different conditions ("MT 1" is Mission Task 1), for a number of different criteria ("MoE 1-1" stands for Measure of Effectiveness, a measure relevant to the first mission task). The handbook stresses that designers must apply a method for mapping colors based on the score relative to a meaningful threshold and discourages use of weighted averages across criteria, as these can be misleading. If we apply this static visualization to our example dataset shown in Table 1 we obtain the visualization in Figure 3. Each cell shows the average value.

Figure 3. Horse blanket (heat map) applied to notional virus outbreak dataset.

The simple example in Figure 3 reveals the drawbacks of this visualization method. By showing only the average values, it does not indicate robustness across plausible deadliness percentages (e.g., best and worst case values for each option). This means it does not provide option awareness level 1. Since it does not show the robustness, it also cannot show the factors underlying that robustness. Additionally, users have no easy way to interact with, filter, or drill down into the data to explore the relationships among factors and thus to achieve option awareness level 2. Section 2.2 briefly discusses a few emerging visualization techniques used to achieve option awareness.

2.2. Interactive Techniques

In addition to the static visualizations found in presentations, there are commercial software products that provide interactive visualizations for comparing options. These interactive tools tend to depend heavily on different analytic techniques. Some of these tools provide some support for option awareness level 2, in which a person understands the factors underlying the outcomes. Sensitivity analysis and classification trees are two examples of techniques found in these products.

Sensitivity analysis allows users to determine how much of the variance in the outcome is due to different sources of uncertainty. For example, based on the data in Table 1, the variance in the set of scores associated with each option results from the choice of whether to implement social distancing, the choice of how much of the population to vaccinate, and the deadliness of the virus.

The simplest approach to sensitivity analysis involves varying a single condition at a time and examining the impact on the outcome. Tornado charts show the results of this approach. For example, in the tornado chart in Figure 4, the occupancy rate of schedule flights (i.e., capacity of scheduled flights) was varied from 40% (worst case) to 60% (best case). The chart shows that if occupancy is at only 40% the airline will lose money; however, if occupancy hits 60% the airline will profit by approximately $35,000 annually. The sort order places the variable that causes the most variance in the outcome on top. In this example, the occupancy of scheduled flights has the greatest influence on the annual profit while the cost of insurance has the least.

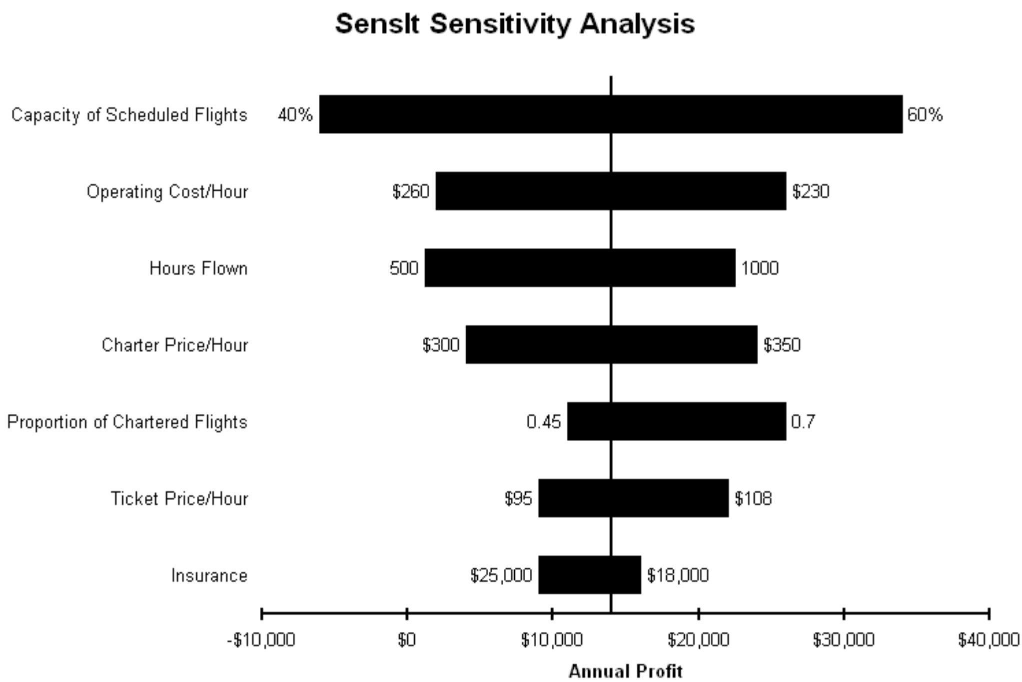


Figure 4. Tornado chart. Adapted with permission from Sensit (2013).

Sensitivity analysis shown in tornado charts falls short on option awareness level 1. Since each factor is varied independently, the analysis involves a single notional scenario rather than values from multiple varying or uncertain conditions. However, such charts do start to provide option awareness level 2 by showing the factors that have the greatest impact on the outcome. For example, if we consider our notional example from Table 1, our tornado chart would look like Figure 5 for a worst case scenario. The effect of each factor (Dist, Vacc, Deadliness) on the total number of lives lost clearly identifies social distancing as the most important factor because it changes the percentage of the population exposed to the virus. It does not show interaction effects (i.e., that if social distancing is implemented the percentage of the population vaccinated matters less), but does provide a simple method of ranking the importance of the underlying factors. The need remains for specific visualization approaches that could begin to explain conditions under which a particular option will produce a better or worse outcome.

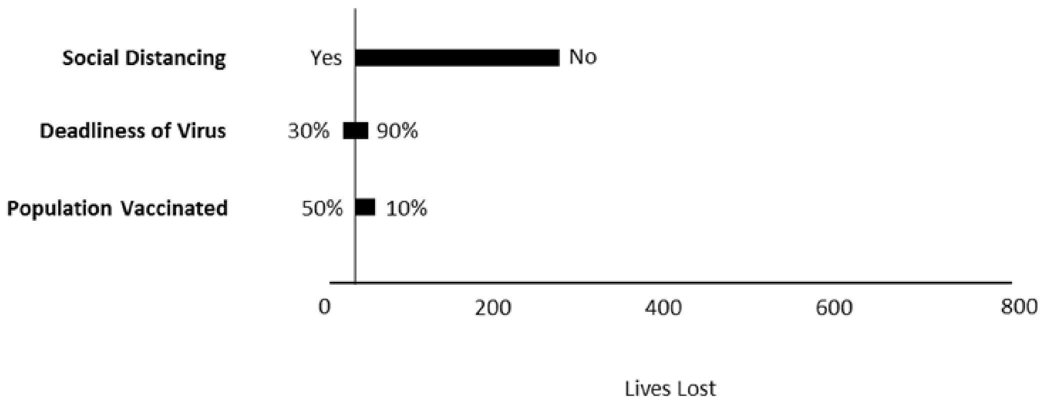


Figure 5. Tornado chart showing notional virus outbreak dataset.

A full summary of all the commercial solutions is beyond the scope of this chapter. The example in Figure 4 came from SensIt, an add-on for Excel. Other companies, such as Vanguard Software, have similar software solutions for generating non-interactive visualizations.[2] We mention them purely as examples of existing commercial software that represents the types of visual analytics tools that will eventually support users in achieving progressively higher levels of option awareness.

Another interactive visualization technique that aids in reaching option awareness level 2 is a decision tree. The data mining algorithms that produce classification trees can generate decision trees that bin the outcomes into categorical groups (e.g., good, moderate, and poor outcomes). The leaf nodes in a tree visualization show the number of good, moderate, or poor outcomes that resulted from following a particular path down the tree. Figure 6 gives an example for the notional virus dataset, showing that the decision to implement social distancing was the most crucial factor in ensuring good outcomes. Social distancing led to the fewest number of lives lost, regardless of the percentage of the population vaccinated or the deadliness of the virus.

[2] See http://www.vanguardsw.com/products/vanguard-system/risk-analysis.htm

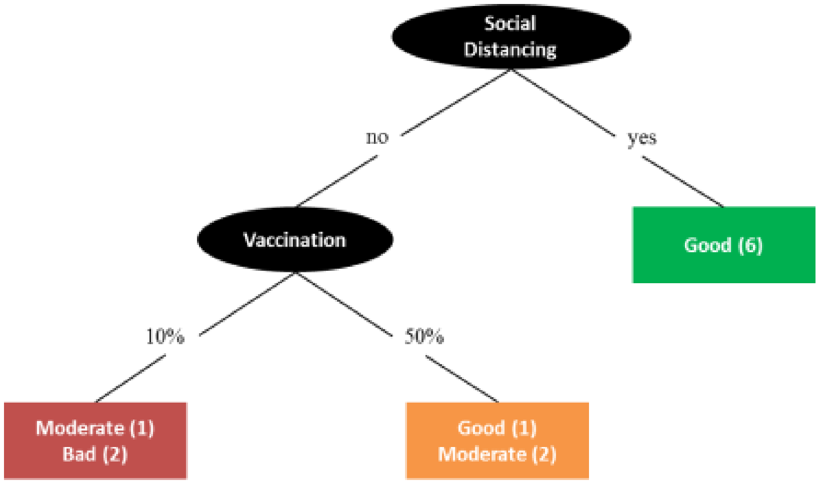


Figure 6. Decision tree showing option awareness level 2 for notional dataset's number of lives lost.

The factors that have the most influence on the outcome appear at the top of the tree. This approach helps explain the factors behind a particular outcome. As in the case with the manual techniques, much work remains to support option awareness effectively.

2.3. Lessons Learned from Practice

Readers must keep in mind that while this chapter uses a small, simple, notional dataset for illustrative purposes, in practice the source dataset can be large and complex. The scalability of the visualization techniques and interaction with the visualizations must enable display of large complex datasets in ways that domain experts can quickly understand. Often the complexity involved means that people resort to very simplistic stoplight chart indicators (as in the horse blanket), consider only a single scenario (as in the tornado chart), or show only the best and worst cases for a given option.

2.4. Gaps

Better interactive visualization tools, especially those that appeal to domain experts, are needed to bring these capabilities to an audience without a strong background in statistics. Users need tools that can enable all three option awareness levels and provide a seamless transition between levels. These tools must be accessible for end users, with appropriate training on basic statistics and advanced visualization techniques where necessary. Research is just beginning to address some of these issues.

3. Visualization Techniques for Option Awareness

While existing visualization techniques can show the necessary types of data, using these techniques to visualize option awareness data represents a new direction. In this section, we describe a few techniques that developers could incorporate in future tools to promote option awareness levels 1 and 2. How well visualization can enable and support option awareness level 3 remains an open question.

3.1. Option Awareness Levels

3.1.1. Option awareness level 1 – Showing distribution

Option awareness level 1 requires tools that show the relative robustness of each option. This means displaying the set of plausible outcomes for each option. The challenge for option awareness level 1 lies in showing the distributions of outcomes. Designers and developers must determine how much statistical analysis is needed and appropriate for the audience and what can be simplified.

Visualization Approach

Box plots, also known as box-and-whisker plots, are a statistical visualization technique created by Tukey (1977). There are many variations of the box plot; a basic plot is shown in Figure 7.

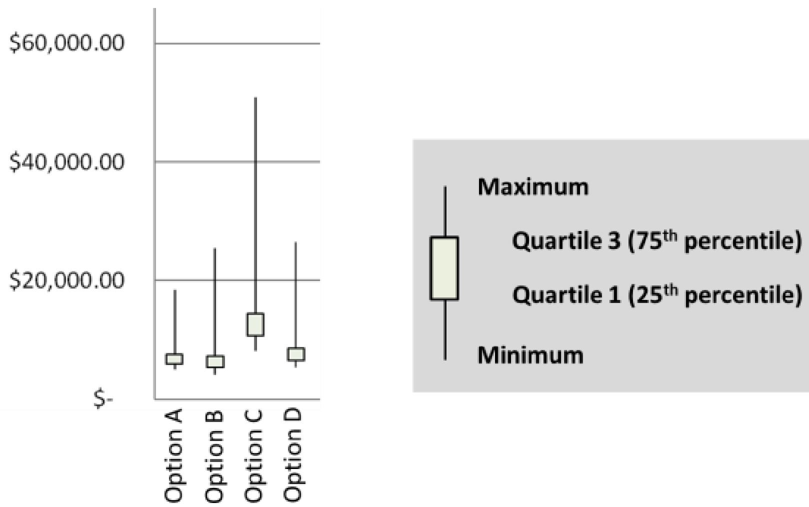


Figure 7. Box plot showing distribution of plausible futures (option awareness level 1) for four options.

Data

Each of the four notional options (Options A, B, C, and D) in the dataset had 35–40 possible outcomes in the dataset represented in Figure 7. The algorithms calculated the minimum cost possibility, maximum cost possibility, first quartile, and third quartile.

Encoding

The vertical line represents the range from the best case to worst case. The box in the middle shows the middle 50% of all plausible outcomes for that option (i.e., the 25th–75th percentile).

Interaction

Users may wish to sort data based on the compactness of the range of plausible outcomes (length of the line), the best outcome (bottom of the line), the worst outcome (top of the line), or the characteristics of the middle 50% of plausible outcomes (size of the box, top or bottom of the box).

Advantages and Disadvantages

The box plot provides an excellent statistical summary of the data, allowing users to compare best, worst, or most likely outcomes under uncertainty. However, users must understand the concepts of the 25^{th} and 75^{th} percentiles and basic statistics to interpret the plot.

Alternative Visualizations

There are many alternative visualization techniques that show the actual values rather than a statistical summary of distribution, as well as many variations on the basic box plot. The most basic alternative technique is the histogram. Figure 8 shows the landscape of plausible outcomes and provides option awareness level 1. The challenge in developing such a histogram is to appropriately bin the data into groups (e.g., $4,000–$4,999 is a bin). A wise selection of bins can create an effective way to show option awareness level 1 using the actual data points rather than statistical summary data.

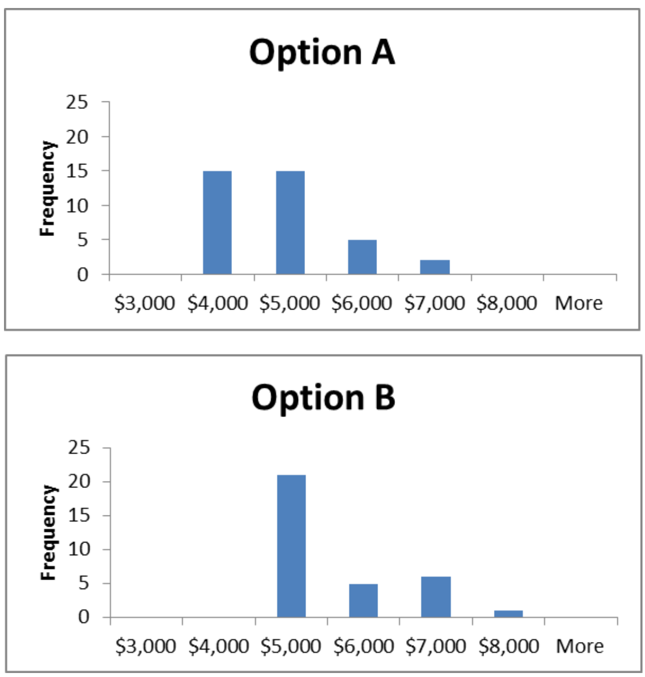


Figure 8. Histogram showing landscape of plausible futures.

Variations of the box plot show values in addition to the statistical summary data. These variations include density plots, density plots with a rug, bean plots, and violin plots; Yau (2013) gives a brief tutorial on these plots and how to implement them using the programming language R. Bean plots show all of the specific data points while addressing the issue of overlapping points. Other proposed techniques include showing all outcomes as lines, but making each line semi-transparent to reveal where many points are superimposed on each other (Barnes, 2013).

Some past work has compared these visualization designs and discussed the trade-offs among different visualization techniques for showing the distribution of values (Ibrekk & Morgan, 1987). The main takeaways were that simple explanations of the visualizations aided understanding, but only minimally, and basic visualization techniques like box plots or the semi-transparent approach offered particular promise. Visualization techniques for showing distribution of values have seen little adoption, perhaps because of the need for users to understand basic statistics. Future research should address the comparative effectiveness of these techniques to support option awareness and gain adoption.

While all the examples provided may lead to option awareness level 1 because they show the data distribution, they do not provide insight into the underlying factors that caused the difference in outcomes. That insight requires effective multidimensional visualization techniques that show the relative importance of factors that result in good or bad outcomes for a particular option.

3.1.2. Option awareness level 2 – Showing multidimensional data

Option awareness level 2 requires a way to display the underlying factors driving the outcomes and the interactions among those factors. In our simple notional dataset, this means we need a way to effectively understand the impact that social distancing, vaccinations, and deadliness of the virus have on the outcomes. We also need to know how those variables interact. Additional benefits of visualizing the factors include that showing users what factors affect better vs. worse outcomes provides a foundation for developing new courses of action (see section 3.1.3). Furthermore, users can immediately see the sources of uncertainty and can then question the underlying assumptions made in the models. Moreover, seeing the factors can help users compare the relative importance of each source of uncertainty in a given situation. The challenge for option awareness level 2 lies in showing the most critical factors and the relationships among them. In many datasets this challenge is further compounded by a large number of variables.

Visualization Approach

A key visualization technique in this area is the decision tree, classification tree, or variant of it (see section 2.2 and Figure 6).

Data

The data includes all of the plausible future paths for all of the options and the corresponding outcomes.

Encoding

Some of the nodes in the tree represent a decision point (e.g., whether to implement social distancing) or variable (e.g., deadliness of the virus). The decision points or variables with the greatest impact on the outcome appear higher in the tree. The leaf nodes at the lowest level of the tree structure represent the outcomes. Nodes can be combined and aggregated in cases where a variable has no influence on the outcome. For example, if social distancing is implemented then the outcome will always be good, regardless of vaccinations or deadliness, so the nodes under yes to social distancing can be collapsed into a single node that represents good outcomes.

Interaction

Users may wish to filter the data so that only paths that lead to good outcomes are shown in the tree. They also may wish to prune the tree by selecting a value for a particular option (e.g., no to social distancing).

Advantages and Disadvantages

While decision trees can prove highly effective in aiding option awareness level 2, they can become complex when a situation involves many variables. Users may require training to become familiar with the underlying analysis and understand what the tree shows. To attain option awareness level 2, users need new, easy-to-grasp visualizations that show the interactions among factors and the impact they have on the outcomes (i.e., that display the data from all plausible futures, not just one).

Alternative Visualizations

There are alternative visualization techniques, techniques other than decision trees, which can assist with option awareness level 2. These techniques for visualizing multidimensional data include a scatterplot matrix and parallel coordinates. A full discussion of approaches for multidimensional data visualization, which would help in establishing option awareness level 2, is outside the scope of this chapter. We cover this topic briefly only to help the reader understand what research has already accomplished and where open research challenges remain.

3.1.3. Option awareness level 3 – Identifying new options

Once users have option awareness levels 1 and 2 they will understand the factors underlying the outcomes for a particular option(s) and will need a way to create a new option and confirm the predicted robustness of that new option. Creating a new option calls for a capability to specify options not previously considered (e.g., apply quality assurance to a vaccination program to ensure a minimum vaccination level). Interaction controls could help achieve option awareness level 3 provided they feed new option information back to level 1 or 2 visualizations.

3.2. Gaps

While visualization and interaction techniques can move us one step beyond the current state of the practice, the biggest gains would come from (a) new visualization techniques that provide option awareness level 3, (b) visualizations for option awareness that are easy for non-statisticians to understand and use, (c) tests of the effectiveness of these new techniques for supporting selection of options in the real world (e.g., which multidimensional visualization techniques best support selection of an option and aid understanding of factors underlying the option’s robustness), and (d) an integrated capability for achieving progressively higher option awareness levels (e.g., bringing together option awareness levels 1, 2, and 3). Further research is needed to understand how well existing techniques work when used in real-world situations where the uncertainty would decrease over time and the option awareness data could be refined in real time. This calls for visualization methods that clearly show users where, how, and why the robustness of options has changed.

4. State of the Art

In this section we focus on the most recent advances and highlight cutting-edge research on visualization for option awareness. Because these techniques are new, much remains to be done to confirm their effectiveness. In the coming years, however, these capabilities should make their way into the operational community.

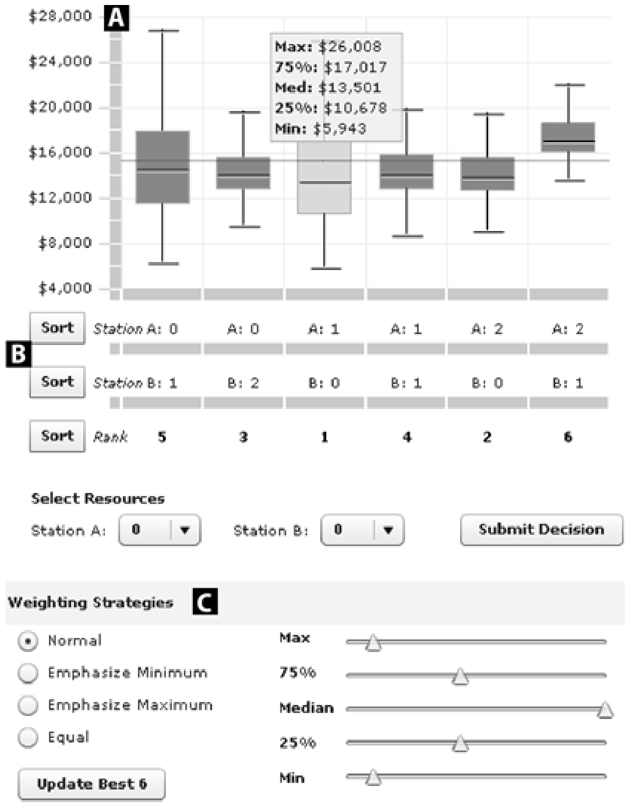
4.1. Areas of Active Research

While researchers have produced many systems over the years to increase situation awareness, the distinction between visualization needed for situation awareness versus option awareness is still emerging. Significantly less effort has centered on visualizing the robustness of different options and interacting with the option awareness data in a meaningful and comprehensible way. The research literature in this area is very recent, and scanty.

We first discuss design principles proposed for effective option awareness and some new interaction and visualization techniques. We then present research on bringing option awareness visualizations to a broader audience (to users with strong domain knowledge who may not be statisticians).

4.1.1. Design principles for effective option awareness visualizations

Researchers have developed many design principles to undergird situation awareness visualizations (Endsley, Bolte, & Jones, 2003). However, Drury, Pfaff, Klein, and Liu first proposed a set of visualization design principles to support option awareness (Drury, Pfaff, Klein, & Liu, 2013). These principles emerged from multiple experiments. Figure 9 gives an example of an interactive box plot visualization used in some of these experiments.



User interface example showing (A) frequency distributions of outcomes for top six options, (B) sorting by rank or quantity of resources, and (C) controls to customize the weighting strategy used to rank options, including four presets.

Figure 9. Option awareness visualization used in experiments informing design principles. Adapted with permission from Drury, Pfaff, Klein, & Liu (2013).

Drury et al.'s design principles are both interaction oriented (i.e., allow users to specify the values for the input parameters, to add their own weights to each criterion, and to filter data and sort options) and visualization oriented (i.e., show only the top few options and show the outcome under a variety of conditions rather than a single set of most likely conditions). Their contributions and insights cover much of option awareness level 1 and just begin to address option awareness level 2. These represent the first visualization design principles specifically designed for option awareness, and will surely evolve as option awareness concepts become more widely adopted and researchers learn more from real-world use of their models and visualizations.

4.1.2. New interaction and visualization techniques

As described in section 3, many new visualization capabilities have begun to emerge in the field of visual analytics. These capabilities align nicely with some of the underlying needs in the area of option awareness, as visualization of uncertainty is an essential part of option awareness. These interactive visualizations provide new mechanisms for achieving option awareness. However, much remains to be done to effectively visualize option awareness data.

In the remainder of this section we describe some visualization techniques that reflect the latest thinking about visualizing uncertainty: a capability essential to achieving option awareness. Heinzl (2012) gives a sampling of the latest visualization and interaction techniques, some of which could be used to obtain option awareness level 1 because they relate to visualizing the distribution of values. Others provide an analysis of the parameter space – a type of visualization that closely aligns with option awareness level 2 and with understanding the factors underlying the robustness of each option.

Most of the latest techniques related to option awareness level 1 extend or combine previously known techniques. Pang's (2012) perspective on the field of uncertainty visualization covered visualization techniques such as histograms, box plots, uncertainty glyphs, embellishments, pseudo-coloring, transparency, and fuzziness and dust clouds. Potter (2012) described visualization techniques that directly relate to option awareness level 1 and visualization of the distribution of values, such as error bars, spreadsheets, and many variations on box plots (including additions for aesthetics, value prevalence, sample size, confidence levels, moments, and modality). The newest visualization technique she discussed was the summary plot (Figure 10) which combines four related plots into a single plot that presents more information than any of the visualizations would alone. Potter, Kniss, Riesenfeld, and Johnson (2010) describe the summary plot as an advanced visualization technique that experts could use to more fully understand the robustness of each option. Such plots yield even more information on the comparative robustness of different options to viewers with a strong statistics background. However, as Potter points out, users need new visual metaphors to help visualize uncertainty.

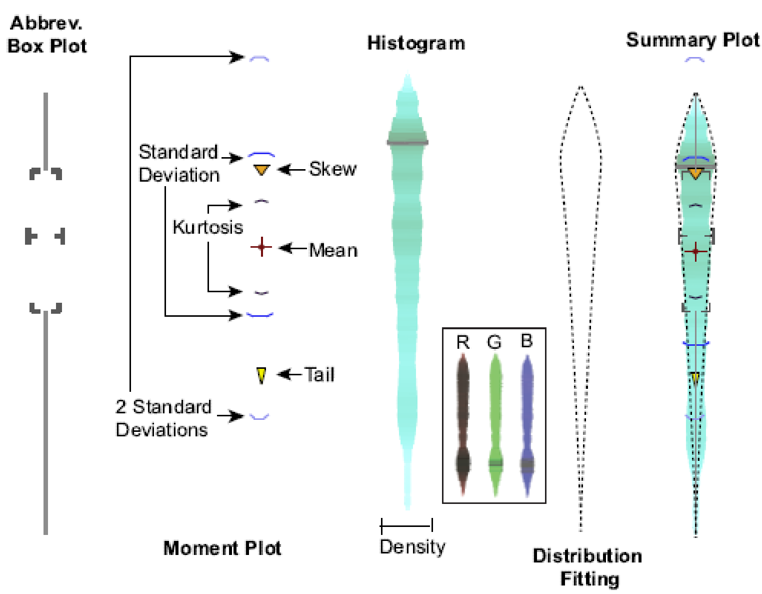


Figure 10. Summary plot. Adapted with permission from Potter, Kniss, Riesenfeld, & Johnson (2010).

The latest visualization techniques related to option awareness level 2 tend to center on the interactions among multiple linked views of the data. A good example of this additional interaction is Heinzl's (2012) visualization, which assists car designers in such tasks as selecting the fuel injection timing and engine speed in order to meet their targets with respect to torque and fuel consumption. The visualizations used include scatterplots and parallel coordinate plots with box plots overlaid, and incorporate multiple linked views. Figure 11 shows examples of these visualizations. Note that the visualizations in this example show multiple aspects of the option to be selected and two criteria for evaluating the goodness.

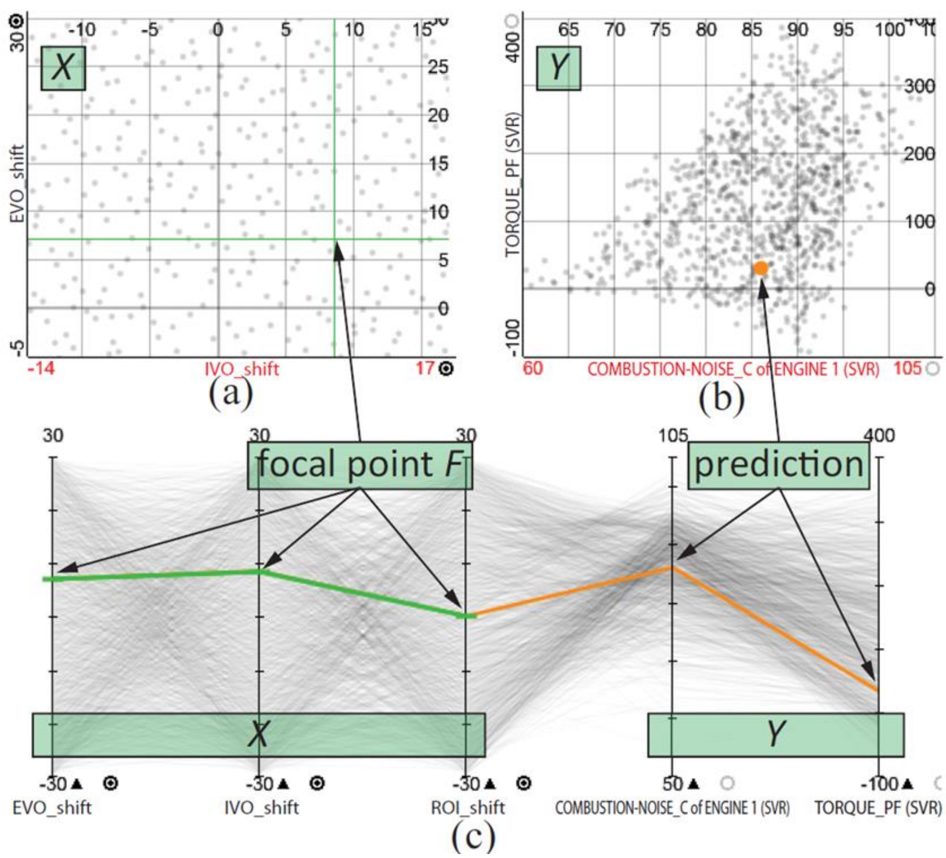


Figure 11. Scatter plots and parallel coordinate plots used for car design. Adapted with permission from Berger, Piringer, Filzmoser, & Groller (2011).

Some researchers have begun to address ways of promoting better understanding of the space of possible futures as a scenario plays out over time. One tool, World Lines, does this by using a horizontal tree visualization below a timeline (Waser, Fuchs, Ribicic, Schindler, Bloschl, & Groller, 2010). As the user selects a new action to try at a particular point in time, a new branch grows out of that tree. The user can see multiple scenarios (branches in the tree) in a single view and compare how they develop over time. The authors use the example of a flooding scenario in which users must decide how to place sandbags in a way that minimizes the number of flooded buildings and uses as few sandbags as possible. World Lines does not necessarily provide option awareness level 1 or 2, since it does not show the robustness or factors underlying the goodness of the chosen options. However, it does allow users to steer the modeling by specifying input parameters at particular points in time and to try new scenarios. This may begin to provide support for option awareness level 3. The tool also represents a step toward an ability to provide more information on the temporal dimension of option awareness data. Future research should explore this area further.

Most of the tools described in the research literature are best suited for experts with statistical knowledge. Engineers can use these interactive visualization tools and approaches, but the domain expert would require training on a number of different statistical concepts to understand the resulting visualizations. In the next section we examine how innovative tools might reduce the necessary training and provide more intuitive visualizations that would help make option awareness more accessible to domain experts who do not have a background in statistics.

4.1.3. Bringing capabilities to a wider audience

Visualization could greatly assist decision makers who could benefit from understanding the robustness of their options but aren't statisticians. The visualization community has recently recognized the importance of filling this gap because more and more mainstream audiences now use interactive visualizations. In fact, a recent master's thesis (Danziger, 2008) described creating "Information Visualization for the People" to address these opportunities.

Media organizations provide a source of inspiration. News media must rapidly build visualizations that a diverse population can easily interact with and understand. For example, the *New York Times* and *Washington Post* now have interactive visualization departments. The *Washington Post* recently published one such visualization (Andrews, Cameron, & Keating, 2013). This interactive visualization could inspire future option awareness level 1 visualizations because the data has the characteristics needed: a distribution of values (for a given medical procedure). The equivalent data in our example relates to the distribution of outcomes for a given option. In the *Washington Post* visualization, they chose to show the lowest value, highest value, and average value. By selecting a particular state from the drop-down menu at the top, users can see where a particular state appears in the distribution, revealing how that state compares to the others in terms of being closer to the minimum or maximum or closer to the average.

A second example, this one from the *New York Times*, provides a visualization that could serve as an inspiration for providing option awareness level 2. The *New York Times* created this visualization (Bostock & Carter, 2012) in the Fall of 2012 during the U.S. presidential race. The example shows the rough equivalent of an interactive decision tree. A user can drill down into the data by mousing over paths in the tree to obtain more detail, and can filter the data by selecting a particular path or by selecting the outcome in a particular state shown at the top. The filtering prunes the tree to show the remaining paths to victory for each presidential candidate. Similar visualizations and interactions could provide level 2 option awareness to users if populated with different data.

These visualizations, particularly in their interactive form, increase the understandability of complex data, thus reducing the human cognitive bottleneck. The data may be complicated, but the visualizations need not be. Researchers must apply this approach of simplifying complex data as much as possible, but no more, to option awareness so that the visualizations can be used by a wider audience. News media organizations such as those mentioned here probably have especially valuable information regarding how much to simplify an analysis to make the process clear to the average person who does not have a strong statistics background.

However, research must still determine the effectiveness of such visualization techniques in the context of supporting robust decision making. To generate option awareness, decision support scenarios may call for more complex information than that used in the examples above. Nonetheless, these visualizations provide important examples of techniques that do not require explanation and instead enable users to focus only on understanding the data. The key is to find an appropriate balance.

4.2. Gaps

All of the option awareness concepts, new visualization design ideas, the ability to simplify those designs, and the underlying design principles are evolving and would benefit from greater insight. New, more effective visualization techniques are needed for all levels of option awareness, particularly for level 3. The design principles must evolve together with the concepts and practice, based on use in the real world. Research should also seek to develop methods for bringing all three levels of option awareness together into a single integrated tool: one that allows users to progress easily from one level of option awareness to the next. As the basics evolve, these visualizations must eventually be adapted to extend beyond the desktop to visualizations appropriate for decision makers who use tablets or mobile devices to display data; in other words, to move the capabilities out of command center-like environments and into the hands of users in the field.

5. Call to Action

To bring the new technologies and methods into operational use all of these techniques must coalesce and the emerging set of design principles must evolve. Widespread use of visualization tools depends on advanced design principles based on observations made in the field rather than only in a laboratory.

In addition to systems based on these advanced visualization design principles, users will need to understand the underlying analysis that generated the data shown in the interactive visualizations, and will need to be trained on how to apply these visualizations for decision making. User training and experience, combined with better design principles and tools, could avoid many potential pitfalls in the use of visualizations. Further, as visualizations evolve, researchers must determine how much to simplify the complexity of the underlying analysis and visualization techniques. Effective visualizations must achieve a balance between supplying enough detail to support good decisions while not displaying so much detail that users dismiss the tools as too complex.

In the meantime, developers should seek to smooth the transitions from situation awareness displays into option awareness displays. Tools are also needed that allow users to progress from one level of option awareness to the next. As a situation plays out some of the uncertainty will disappear, which calls for rapid feedback into the tools to update underlying assumptions and uncertainties. Option awareness visualizations must also update option comparisons as uncertainty decreases to help users know when to adapt their mitigating actions as a situation evolves.

The future of option awareness visualization is wide open, with research only starting to focus on option awareness as a distinct concept. Realizing these capabilities in future visual analytics tools offers great potential benefits to decision makers faced with uncertainty.

References

AFMC. (2008). *Analysis of alternatives (AoA) handbook: A practical guide to analyses of alternatives*. Kirtland Air Force Base, NM: Air Force Materiel Command, Office of Aerospace Studies.

Ancker, J. S., Senathirajah, Y., Kukafka, R., & Starren, J. (2006). Design features of graphs in health risk communication: A systematic review. *Journal of the American Medical Informatics Association (JAMIA), 13*(6), 608–618.

Andrews, W., Cameron, D., & Keating, D. (2013, May 8). *Disparity in medical billing*. Retrieved from http://www.washingtonpost.com/wp-srv/special/national/actual-cost-of-medical-care/

Barnes, R. C. (2013, June 2). *Visualizing value distributions and event time clusters in one dimension*. Retrieved from https://www.hci-matters.com/blog/2011/10/12/archives/289/

Berger, W., Piringer, H., Filzmoser, P., & Groller, E. (2011, May). Uncertainty-aware exploration of continuous parameter spaces using multivariate prediction. *Eurographics /IEEE Symposium on Visualization 2011 (EuroVis 2011), 30*(3).

Bostock, M., & Carter, S. (2012, November 2). *512 Paths to the White House*. Retrieved from http://www.nytimes.com/interactive/2012/11/02/us/politics/paths-to-the-white-house.html

Danziger, M. (2008). *Information visualization for the people.* (Master's thesis). Cambridge, MA: Massachusetts Institute of Technology. Retrieved from http://cms.mit.edu/research/theses/MichaelDanziger2008.pdf

Drury, J., Pfaff, M., Klein, G., & Liu, Y. (2013, July). Decision space visualization: Lessons learned and design principles. Proceedings of the *15th International Conference* on *Human Computer Interaction (HCI) International 2013, vol. 8007,* 658-667. Las Vegas, NV: Springer.

Endsley, M. R., Bolte, B., & Jones, D. G. (2003). *Designing for situation awareness: An approach to user-centered design.* New York, NY: Taylor & Francis Inc.

Heinzl, C. (2012, October). *Industrial applications of uncertainty and parameter space analysis*. IEEE VisWeek 2012 Tutorial on Uncertainty and Parameter Space Analysis in Visualization. Seattle, WA. [PowerPoint Slides.] Retrieved from http://www.cg.tuwien.ac.at/research/publications/2012/VisWeek-Tutorial-2012-Uncertainty/S6b_Vis2012_Tutorial_CH_Handouts.pdf

Heinzl, C., Bruckner, S., Groller, M. E., Pang, A., Hege, H.-C., Potter, K., . . . Moller, T. (2012, October). *IEEE VisWeek 2012 Tutorial on Uncertainty and parameter space analysis in visualization*. Seattle, WA. [Abstract.] Retrieved from http://www.cg.tuwien.ac.at/research/publications/2012/VisWeek-Tutorial-2012-Uncertainty/

Ibrekk, H., & Morgan, M. (1987). Graphic communication of uncertain quantities to nontechnical people. *Risk Analysis, 7*(4), 519–529.

Klein, G. L., Drury, J. L., Pfaff, M., & More, L. (2010, June). COAction: Enabling collaborative option awareness. In *Proceedings of the International Command and Control Research and Technology Symposium (15th ICCRTS).* Santa Monica, CA.

Pang, A. (2012, October). *Uncertainty visualization*. IEEE VisWeek 2012 Tutorial on Uncertainty and Parameter Space Analysis in Visualization. Seattle, WA.

Potter, K. (2012, October). *Statistical uncertainty: From quantification to visualization*. IEEE VisWeek 2012 Tutorial on Uncertainty and Parameter Space Analysis in Visualization. Seattle, WA.

Potter, K., Kniss, J., Riesenfeld, R., & Johnson, C. R. (2010). Visualizing summary statistics and uncertainty. In *Computer Graphics Forum*, *Proceedings of Eurovis 2010*, *29*(3), 823–832. Retrieved from http://www.sci.utah.edu/publications/potter10/Potter_EuroVis10.pdf

SensIT (2013). Tornado chart. Retrieved from http://www.treeplan.com/sensit-for-sensitivity-analysis.htm

Sheridan, T. B. (2002). Humans and automation: System design and research issues. Santa Monica, CA: Wiley-Interscience.

Shneiderman, B. (1996, September). The eyes have it: A task by data type taxonomy for information visualization. In *Proceedings of the IEEE Symposium on Visual Languages*. p. 336–343. Washington, DC: IEEE Computer Society Press.

Tufte, E. T. (1997). *Visual explanations.* Cheshire, CT: Graphics Press.

Tukey, J. W. (1977). *Exploratory data analysis.* Reading, MA: Addison-Wesley.

Waser, J., Fuchs, R., Ribicic, H., Schindler, B., Bloschl, G., & Groller, M. E. (2010). World lines. *IEEE Transactions on Visualization and Computer Graphics, 16*(6), 1458–1467.
Yau, N. (2013, June 2). *How to visualize and compare distributions*. Retrieved from Flowing Data: http://flowingdata.com/2012/05/15/how-to-visualize-and-compare-distributions