October, 2006

# Assessment of Display Attributes for Displaying Metainformation on Maps

Ann M. Bisantz, *SUNY University at Buffalo*
Jonathan Pfautz, *Charles River Analytics*
Richard T. Stone, Ph.D., *SUNY University at Buffalo*
Emily M. Roth, *Roth Cognitive Engineering*
Gina Thomas-Meyers, *United States Air Force Research Laboratory*, et al.

# ASSESSMENT OF DISPLAY ATTRIBUTES FOR DISPLAYING META-INFORMATION ON MAPS

Ann M. Bisantz[1], Jonathan Pfautz[2], Richard Stone[1], Emilie M. Roth[3], Gina Thomas-Meyers[4], Adam Fouse[2]

[1] University at Buffalo, Amherst, NY; [2] Charles River Analytics, Cambridge, MA; [3] Roth Cognitive Engineering, Brookline, MA; [4] AFRL/HECS, Wright Patterson AFB, OH

In many domains, operators need to understand and act on large volumes of information from a variety of sources. Operators are particularly challenged by the need to reason about the qualifiers of that information. These qualifiers, or “meta-information”, include characteristics such as the uncertainty associated with data, the age of the data, and the source of the data. Often, these critical data qualifiers are not presented, or are not incorporated into the primary information displays used by operators. In this research, we conducted a controlled experiment to investigate the utility of four common color display attributes (hue, saturation, brightness, and transparency) for displaying meta-information under different map background, task, and meta-information-type conditions. Results indicated that participants could rank and rate display elements which varied based on saturation, transparency and brightness similarly to expected ranks and ratings. Background effects were limited; but task type and framing effects indicated that the “natural” direction for ranking may be context-dependent.

## INTRODUCTION

### Background

In many domains, operators need to understand and act on large volumes of information from a variety of sources. Operators are particularly challenged by the need to reason about the qualifiers of that information. These qualifiers, or “meta-information” (Pfautz et al., 2005), include characteristics such as the uncertainty associated with data, the age of the data, and the source of the data. For example, in military command and control tasks, commanders must reason about the location of threats. Information about those threats may come from sensors with associated uncertainties, be several hours (or days) old, and/or be derived from intelligence sources with varying degrees of trustworthiness. These critical data qualifiers are generally not presented, or are not incorporated into the primary information displays used by commanders.

While some earlier work has speculated about methods for representing meta-information (Pfautz et al., 2005), most prior research has been focused on displays of data uncertainty (see Bisantz et al., 2005, or Pfautz et al. 2006 for a review). For instance, Bisantz et al. studied the use of blurred or colored icons to display uncertainty about an object’s state; other work on scientific visualization and geographical information systems has developed techniques (albeit with limited performance testing), using a variety of graphical codes (e.g., color, texture), to indicate data uncertainty over areas on geospatial displays (Pang et al., 1997).

### Study Overview

In this research, we conducted a controlled experiment to investigate the utility of four common color display attributes (hue, saturation, brightness, and transparency) for displaying meta-information under different background, task, and meta-information-type conditions. For this study, we focused on the ability of participants to reliably rank and rate (i.e., assign a numeric value to) sets of display elements created based on the four attributes. If an attribute is useful for displaying meta-information, participants should be able to reliably rank the meta-information shown with that attribute and map a display element shown with that attribute to a specific value.

The study was designed to provide information relevant to the following research questions:

1. How does the ranking and rating of meta-information by participants vary based on the type of graphical coding used to convey the meta-information?

2. Does the utility of a display attribute for displaying meta-information vary with the background on which the information is displayed, the number of levels of meta-information being displayed, or the type of meta information (e.g., uncertainty, latency, information source) being displayed?

## METHOD

### Participants

Thirty volunteers (21 men and 9 women) were paid 15 US dollars each to participate. The volunteers ranged in age from 20 to 29 years and had normal (20/20) or corrected-to-normal vision. All of the participants had computer experience and were screened prior to the study to insure they were not colorblind.

### Independent Variables

This study included four independent variables: two within-subject and two between-subject. The within-subject variables were meta-information display attribute (variation based on hue, brightness, saturation or transparency) and background (map and grid). The between-subject variables were task framing (probability or latency/information age) and level of specificity (4, 8, or 12 levels). There were five participants in each combined between-subject condition.

### Experimental Stimuli

Participants were shown a 30.5 x 30.5 cm area displaying multiple 2.5 x 2.5 cm square colored elements. Elements were placed on randomly selected centers of a 7 x 7 grid overlaid on the map to insure consistent minimum distances between the elements (this grid was not displayed to the participants). The display area consisted of either a neutral gray color overlaid with a black grid, or a map. The map background used was sampled from a standard US military map selected to insure that the colors and texture on the map were representative of land maps. Four sets of colored elements were systematically developed to include changes in hue (multiple hues), level of brightness (for a lavender hue), level of saturation (for a magenta hue), and level of transparency (for a red hue). The hues and levels of brightness, saturation, and transparency used were chosen based on principles of color models and past research in this area (Ware, 2000). The number of levels that were displayed (one level was shown by each display element) was either 4, 8, or 12 depending on the condition. Figure 1 shows the experimental stimuli created for the condition with 8 levels of display elements.

### Experimental Tasks

Participants performed a ranking task, followed by a rating task. For both tasks, the 8 display areas (4 display attributes x 2 backgrounds) were presented to participants in random order. In the ranking task, participants were asked to rank the elements according to either latency or probability depending on the framing condition (see Table 1) by dragging numbers (i.e., the integers 1 to 12) from the side of the display onto the elements. For the rating task, a circle randomly appeared around one of the elements, and participants were asked to move a slider to a position between 1 and 100 to assign a value of latency or probability (see Table 1) to that element. Participants rated all the elements on one map before another map was shown. Participants were not provided with normative endpoints in either task (i.e., they were not told if the most saturated or brightest element corresponded to the most recent, or most certain information) to determine if there was a “natural” or stereotypical direction to the ranking.

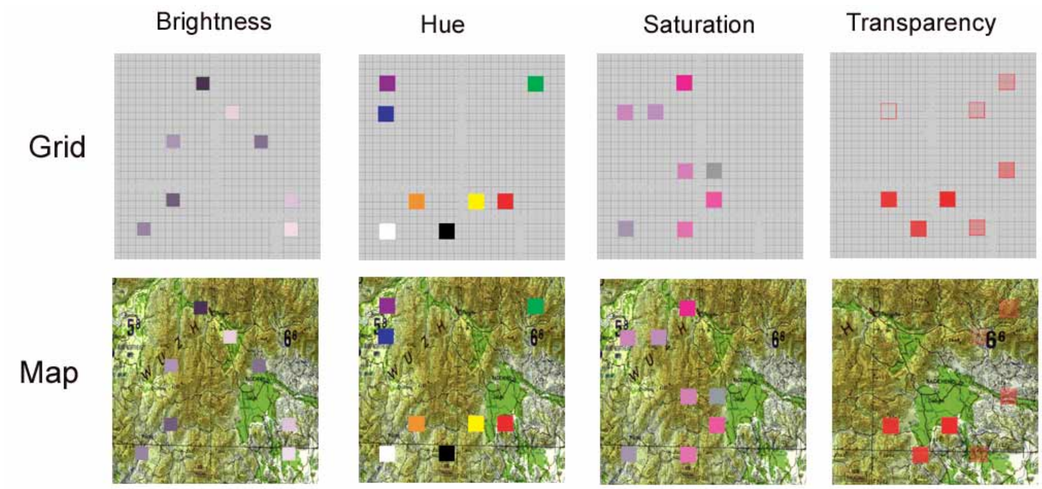


Figure 1. Experimental stimuli for the 8-level condition, for each background and display attribute condition.

**Dependent Variables**

Three dependent measures were recorded in this study: task completion time, assigned ranking (1 to 4, 8, or 12 depending on level of specificity) and assigned rating (1 to 100). Task completion time was computed from the time the first map was displayed to the time the last element on the last map was ranked or rated.

## RESULTS

The following results are discussed below: (1) the effect of time across conditions, (2) the relationship between participants' ranking and a standard or expected ranking, (3) the relationship between participants' ratings and a standard or expected rating.

**Task Time**

Task time was analyzed separately for each task and each level using a mixed-effects ANOVA (display condition x background x framing condition). There were no significant main effects of display condition, background, or framing condition on time, and only one significant interaction across all analyses run, indicating that there was no evidence to reject the null hypotheses that overall task time was not affected by the

Table 1. Instruction segments given on the task screens for the 8-level condition, for task and framing conditions. Instructions were similar for the 4 and 12 level conditions.

| Task/Condition | Task Instruction |
|---|---|
| Ranking: Probability | Please rank the regions according to the chance that a thunderstorm will occur in the region.<br>Use 1 to equal Most Likely and 8 to equal Least Likely. |
| Ranking: Latency | Please rank the regions according to how old the information about the potential thunderstorm is.<br>Use 1 to equal Most Recent and 8 to equal Oldest. |
| Rating: Probability | Please rate the circled region according to the chance that a thunderstorm will occur in the region.<br>0 = No Chance: 100 = Certain. |
| Rating: Latency | Please rate the circled region according to how old the information about the potential thunderstorm is.<br>0 = 0 hours old; 100 = 100 hours old. |

variables of interest.

## Ranking Analysis

To analyze the degree to which participants could rank display elements consistent with expectations, rank orders for each participant were correlated (using the Spearman's rank order correlation coefficient) with a normative, or expected, rank order for each display type. For hue, a standard order was generated by the experimenter to "best match" the orders provided by participants (since there is not a normative order); other orders corresponded to systematic changes in saturation, brightness, and transparency. Correlation results were summarized according to the total number of significant correlations (at $\alpha = .05$), as well as the number of significant positive correlations (to determine whether participants had a tendency to order the display elements in the same direction).

Across all conditions, absolute values of significant correlations tended to be high (from .58 to 1.0). For saturation, brightness, and transparency, participants' rank orders significantly correlated with the normative orders for 138 out of 180 (76 %) total trials. There were no apparent differences due to background (68 trials with significant correlations for map vs. 70 for grid) or framing (69 significant correlations for both probability and latency). The number of significant correlations was similar for the 8 and 12 level, and slightly less for the 4-level condition.

Additionally, for saturation and transparency, there was a directional preference with the majority (73 out of 91) of significant correlations in the same direction (with a similar pattern holding for comparisons across background and framing condition): in general, elements that were "more colored" were seen as more recent, or more likely. For brightness, however, there were differences in order preferences from the probability to the latency framing condition. Participants tended to order items so that darker squares were more certain; however, results for latency were mixed, with about half of the significant rankings associating newer information with darker squares, and half the reverse. This indicates a possible interaction between the type of meta-information and natural ordering for that display attribute. For hue, inspection of the orders provided by participants led to the creation of a post-hoc "standard" ordering, as follows: red, black, brown, purple, orange, yellow, blue, green cyan, pink, gray, and white. Inspection of these colors shows an order from darker, more colored squares to those that are lighter. However, even for this post-hoc order, there were many fewer significant correlations (only 38%) and somewhat lower correlation coefficients (min = .615, max = .786) than for the other four conditions, indicating considerable variability.

## Rating Analysis

Ratings that participants assigned to elements were analyzed in a manner similar to rankings. Expected ratings (based on the number of levels) were computed by dividing the range 0-100 by the number of intervals required, and taking the midpoint of those intervals (e.g., for the 4 level condition, there were four intervals of 25, with midpoints of 12.5, 37.5, 62.5, and 87.5). Ratings provided by participants were correlated (using Pearson's correlation coefficient) with these expected ratings for each display type.

Overall, results were similar to the ranking analysis. For saturation, brightness and transparency, there were significant correlations for 139 out of 180 (77%) trials and the absolute value of significant correlations was high (.596 to 1). The number of significant correlations was similar for the 8 and 12 level, with slightly fewer for the 4 level condition. There were no differences due to background. Like the ranking analysis, there were few differences in the number of significant ratings due to framing (67 significant correlations for probability, and 72 for latency). However, there were differences in the direction with which the elements were ordered: for saturation, brightness, and transparency, the direction of ordering was consistent for the probability condition (all correlations were in a single direction; with more colored elements seen as more certain). For latency, on the other hand, in 40 instances darker or more colored elements were seen as newer, while in 32 instances the reverse was true. For hue, similar to ranking, a post-hoc ordering was created. To best match performance on the rating task, this ordering was slightly different: black, purple, blue, red, brown, orange, green, yellow, cyan, pink, gray, and white. Like the ranking results, this order tends to move from darker, more colored squares to those that appear lighter. However, like ranking, there were many fewer significant correlations (47%) and no perfect correlations (min = .612, max = .975).

## DISCUSSION & CONCLUSIONS

Results indicate that participants could rank and rate the display elements similarly to expected levels. Performance was less consistent with hue, because (particularly for twelve levels) there is not one natural order for hue (as there is, for instance, for levels of transparency). However, there was some tendency to order hues according to a perception of darkness or intensity of color. Interestingly, background had almost no impact, indicating that at least for the map sample chosen, the map colors and annotations did not interfere with participants' abilities to interpret the colored elements (though this may differ for maps that are less homogenous and include, for example, regions of blue colored water). There was some tendency for participants to associated elements that were "more colored" (e.g., were less transparent, more saturated, and darker) with greater certainty. However, the reversals in the direction of ordering between meta-information framing conditions indicate that the "natural" direction for order may be contextually dependent. In a real world setting, training and map legends will be used to fix the ordering direction; however, it is preferable for this direction match user expectations if possible. Results from this study can be used in future research to select candidate display techniques for testing in command-and-control decision scenarios.

## REFERENCES

Bisantz, A. M., Marsiglio, S. and Munch, J. (2005). A Comparison of the Effects of Graphical, Linguistic, and Numeric Representations of Uncertainty on Decision-making. Human Factors.

Pang, A. T., Wittenbrink, C. M., & Lodha, S. K. (1997). Approaches to Uncertainty Visualization. Visual Computing, 13(8), 370-390.

Pfautz, J., Roth, E. M., Bisantz, A. M., Thomas-Myers, G., Llinas, J., Fouse, A., (2006). The Role of Meta-Information in C2 Decision-Support Systems. In Proceedings of the 2006 CCRTS Conference: The State Of The Art and The State Of The Practice.

Pfautz, J., Bisantz, A., Roth, E., Fouse, A., & Shuster, K. (2005). Meta-Information Visualization in Geographic Information Display Systems. in Society for Information Display '05 Digest.

Ware, C. (2000). Information Visualization: Perception for Design. New York: Morgan-Kauffman.