2007

Preliminary Studies in Complex Military Iconic Symbology

D. W. Repperger

G. Thomas

J. Miller

D. Aleva

S. Fullenkamp

PRELIMINARY STUDIES IN COMPLEX MILITARY ICONIC SYMBOLOGY

D. W. Repperger[1]
[1]Air Force Research Laboratory, AFRL/HECP
Wright-Patterson AFB, Dayton, Ohio 45433

G. Thomas[2], J. Miller[2], D. Aleva[2], S. Fullenkamp[3]
[2] Air Force Research Laboratory, AFRL, HEC
Wright-Patterson AFB, Dayton, Ohio 45433
[3] General Dynamics, Inc.
Dayton, Ohio 45433

Two studies are reported on human performance using military visual iconic objects that increased in complexity by augmenting the number of dimensions. The first experiment demonstrates certain perceptual human biases that appear when attempting to correctly identify the state of a standard military icon feature. The second experiment, based on the results of the first experiment, revealed the expected information overload conditions that occur which give rise to exponential laws. The results of the second experiment provide preliminary guidelines on proper synthesis of new iconic objects in terms of their information complexity as well as how they are perceived by operators.

Introduction

A challenging problem in aviation studies involves how to best design key attributes of visual displays to best match the needs of operators as well as to enhance their performance. Investigation of pilot error is fundamental to the improvement of visualization systems. Hooey & Foyle (2006) developed a taxonomy regarding navigation errors during taxi operations. They defined three classes of errors (planning, decision, and execution errors). It was found that the heads-up-display best mitigated execution errors which helped disambiguate the environment. In the investigation of flare accident rates and probable causes for improper flares, Benbassat & Abramson (2002) found that particularly high flare accident rates occurred. An extenuating factor to reduce the problem was pilot experience and improved instruction sets. The method at which display context affects the attention capture is demonstrated on a study of the effectiveness of abrupt onset signals (Nikolic, Orr & Sarter (2004)). The rationale was that aviation has data-rich, event-driven domains. A challenging problem in this regard is to better understand why pilots miss changes in the status and behavior of their automated systems. In the arena of accident investigation analysis, an error analysis and prevention approach within an organization can be useful in conjunction with an objective criteria for exploring human error frameworks (Wiegmann & Shappell, 2001).

There are many stressors that influence performance of humans with aviations systems. The effects of sleep loss, time pressure, and workload level on performance can influence accomplishment during a dynamic task such as air traffic control (Lichacz (2005)). For performance of a mid air traffic detection task, a stressor, clutter, inhibited detection of command changes and traffic (Ververs & Wickens (1998)). Using head-up and head-down displays, the effects of display location, intensity and degree of clutter were investigated. Mid air traffic detection was superior with a head-up display reflecting an attentional trade-off. For air traffic control systems, Metzger & Parasuraman (2006) studied an automated decision aide and data link in an effort to enhance the efficiency and capability of the air traffic control system. The concern was that additional demands are required for the limited visual attentional resources of the controllers. With the use of automated multimodal conflict cures, perhaps mitigation of these demands could occur, especially under high-traffic loads. For the task of correct assessment of in flight weather conditions, Wiggins & O'Hare (2003) investigated decision making. It was found that experts and novices used in-flight cues differently, indicating a need to change instructional sets to accommodate the experience of the users. Another study to support these results was conducted by Endsley & Kris (1994) in the design of expert systems to reduce workload and thus improve upon interface guidelines. Strong differences were found between experts and novices. Recommendations were made on how best to present information so that it matches the reasoning processes of expert systems in future cockpits.

The complexity of displays is an important consideration on how to influence pilot communications and workload. One case in point was reported by Prinzo (2003) for call sign procedures. The goal was to use cockpit displays of traffic information to enable pilots to acquire, identify, verify, and maintain predefined spacing intervals between aircraft. The suggestion was to modify call sign procedures in present aviation systems. Earlier, a list of sixteen factors that contribute to complexity in airspace tasks was developed by Mogford, Murphy & Guttman (1994). Complexity factors were identified both as direct (verbal reports) and indirect (multidimensional scaling) to help discover potential factors. In the investigation of cognitive complexity, Cummings & Tsonis (2006) define three sources: environmental, organizational, and display. Their efforts explored how these components of complexity could be effectively partitioned, measured and compared. They found that color in a display may actually contribute to the cognitive complexity. For complex tasks, from the perspective of a pilot that has to perform conflict detection and resolution tasks, it was found that display dimensionality, conflict geometry, and time pressure of in-flight disagreement produced performance decrements, especially for nonlevel conflict geometries and increased time pressures.

Studies on displays are also very pertinent. In comparing two dimensional to three dimensional displays, to study how visual renderings modulate performance in free-flight environments, the independent variables include maneuver choice, flight safety and mental workload. It was shown by Alexander, Wickens & Merwin (2005) that tradeoffs occurred for different display types. Similarly, Thomas & Wickens (2006) showed 3-D displays eliminated ambiguity costs but increased vulnerability to increased workload. Putting additional information on traffic displays, for the tasks of collision detection and avoidance (Prichett, (2000)), can modify behavior. The performance, however, may not significantly suffer by the changes in the participant's behavior. The tasks studied included collision detection and avoidance during parallel approaches. For cockpit informational displays, Zhang (1997) was concerned with reducing complexity by distributed representation. The goal was to make the distribution across the external information displays in the cockpit in concurrence to the internal minds of the pilots. Again with limited-size displays, Prinzel, Comstock, Glaab, Kramer, & Arthur (2004) examined the efficacy of synthetic visual displays in aircraft that have limited-size display spaces. Issues such as field of view and display size were considered. The general conclusion was that all sizes and texturing methods were viable candidates for synthetic vision displays.

Objective

A special form of a visual display in military applications includes iconic objects, which portray meta information to an operator or commander. Using standard military iconic objects, the fundamental objective will be to see at what level of complexity the visual rendering (complex iconic object) will fail to benefit the operator from a performance point of view. A number of issues require investigation, including how to develop a taxonomy on synthesizing the dimensions of the complex object, and how to discern favored dimensions among those less favored dimensions of an iconic object.

Experiment 1

The first experiment examines the relative saliency of the alternative dimensions that are displayed. The term "saliency" will refer to the ability of a particular dimension (attribute) of an iconic object to elicit accurate responses from the user.

Method

Experimental Design. The first experiment was constructed as a full factorial design in which each subject was exposed to all levels of the independent variable of interest (icon complexity). The number of dimensions (complexity) was varied from 2-10 in a random manner.

Subjects. Six adult subjects participated in this experiment. They ranged in age from 31 to 63 years. They were not directly compensated for participating in this experiment. However, they were compensated for their involvement via salary. Each session lasted approximately 1 hour in duration.

Apparatus. Figure 1 shows the equipment and a typical display image used to collect data. The LCD display (17" x 24") had a 1920 x 1280 pixel resolution. The software was developed using Microsoft Visual Basic to display the requisite imagery and to query the observer regarding the state of an individual dimension of a displayed icon. In the first experiment, the complex icon was portrayed for a fixed duration (t_d) of 1000ms (1 second). Lighting conditions were adjusted so the subjects felt comfortable that the display was prominent enough to discern the icon representation.

Stimuli. Figure 2 shows an example of a ten-dimension icon (Mil-Std 2525B), which is a standard military icon for the US Armed Services. In this visual rendering, the ten dimensions of the icon could have features in either one of two binary states. For the example in Figure 2, the ten dimensions are related to features and states as follows:

(1) The icon is "surface", rather than "air." This is because the shape is round.
(2) The icon is "hostile", rather than "friendly." The color is red.
(3) The icon is "planned." This is indicated by the dashed line around the boundary.
(4) The icon is "reduced." The "-" sign appears in the lower right area.
(5) The direction vector is North since the arrow is more "north" than "south."
(6) The direction vector is East since the arrow is more "east" than "west."
(7) The velocity of the direction vector is "red" indicated by the color of the arrow.
(8) The task force is the rectangular box in the upper left part of the icon. The "no" condition has the box rotated up. The "yes" condition is a flat orientation.
(9) The quantity is "9" as indicated by the number in the lower left area of the icon. The alternative choice would be "6".
(10) The nuclear state is "yes" as indicated by the mushroom cloud object in the upper center of the icon. The alternative state is a "chemical" object which is produced by rotating the mushroom cloud object 180 degrees so that it looks like a chemical flask. This is consistent with the methodology in MIL- Std 2525B.

Training Procedure. Subjects were trained for two days. The training sessions were 1 hour in length and allowed the participants to familiarize themselves with the experimental procedure, the ten icon dimensions and the two possible states (or absence of a state) of each of the icon dimensions that were portrayed and how they would be queried on the state of the iconic dimension. The icons were displayed for 1 second during the training sessions. The data collection phase consisted of 2 sessions of 470 trials for each session. In each experimental trial an icon containing 2 – 10 of the dimensions described above was displayed for t_d = 1 second before the icon disappeared and the question appeared regarding the state of one of the dimensions of the icon. Subjects initiated a trial by clicking a start button, observed the subsequent icon display and then answered a "yes" or "no" or "not applicable" query about a specific state of a dimension. The subjects had as much time as they needed to answer the question and they could take a break if they desired at any time. No feedback was given on whether the response was correct or incorrect. There were two orders of the presentations of the icons that were randomized to eliminate memorization by the subjects.

Experiment 2

The second experiment reordered how the complexity of the iconic object was synthesized based on the results of the first experiment. Both complexity level and time duration (length of the stimulus presentation) were independent variables. The levels of complexity varied from 2 to 10 dimensions. The four levels of presentation time of the stimulus were either 1, 0.5, 0.25 or 0.125 seconds.

Method

Experimental Design and Procedure. Using the same apparatus and similar subject pool as in the first experiment, the following changes were made in the experimental design and procedure:

The assignment of the attributes or features of the complex icon was predicated on the results of Experiment 1.

Results and Discusion

Experiment 1 Results

The purpose of the first investigation was to quantify performance as a function of the complexity (number of dimensions displayed) of the icon as well as rank order each iconic dimension in terms of saliency. In the first study, however, the taxonomy for defining the icon dimension number was randomly assigned. To evaluate performance, a signal detection theory approach was employed to analyze the data. Overall accuracy of detection was the dependent performance measure of interest. Figure 3 displays a truth table that characterizes the responses made by the subjects in this experiment. The sensitivity and specificity plots shown in Figure 4 were determined from the displayed icon dimensions in experiment 1 which had one of two states: "on" or "off." Using methods from signal detection theory, it is possible to define performance accuracy of correct identification as follows (for equal states (balanced design) of the 470 trials):

$$n_1 + n_4 = 235 \text{ trials} \quad (1)$$

$$n_2 + n_3 = 235 \text{ trials} \quad (2)$$

This leads to the following definitions of sensitivity and specificity for this experiment:

Sensitivity. The ability to correctly identify the state of the dimension being in the "on" state.

Specificity. The ability to correctly identify the state of the dimension being in the "off" state.
From Figure 3, it can be shown that:

$$\text{Sensitivity} = n_1/(n_1+n_4) = n_1/(235) \quad (3)$$

$$\text{Specificity} = n_3 / (n_2 + n_3) = n_3/(235) \quad (4)$$

One goal of the first experiment was to investigate how sensitivity and specificity vary as the number of dimensions contained in the icon increases. Figure 4 shows the resulting sensitivity and specificity plots as the icon dimensions were randomly assigned. It is obvious in Figure 4 that the 6-dimension condition had reduced saliency. This is seen by the sudden drop in specificity as the icon complexity increases. The accuracy also showed a reduction in the 6-dimension condition in Figure 5. This result indicated that the experiment had to be carefully redesigned. The assumption that all features were equally salient was incorrect. To overcome these human perceptual biases that occurred, we redesigned the icons and queries to account for this effect encountered.

Experiment 2 Results

From Experiment 1 and Figure 5, the dimensions of the icon were first rank ordered in terms of their performance accuracy. Then a new set of iconic objects were designed for testing in Experiment 2 where we were careful to make sure that the categories of multidimensional icons (3 dimensions, 4 dimensions… 10 dimensions) including iconic features and queries were more equally represented in each category to eliminate the bias encountered in Experiment 1. Figure 6 now displays accuracy of correct detection as the dependent variable and dimension of the iconic symbol as the independent variable. As expected for experiment 2, there was now a monotonic relationship between increasing icon complexity and decreasing accuracy. In an effort to understand the effect of time pressure (icon display duration) on performance, averages were computed across all features, dimensional conditions and subjects for the four levels of stimulus presentation times selected. In Figure 7, it is seen that as the time duration decreases, the accuracy is affected in a degraded sense. Figure 8 plots the accuracy performance versus the log of the stimulus presentation time when averaged across all features, dimensional conditions and subjects. The straight line has a regression coefficient of $R^2 = 0.944+$ providing strong support to the concept that increases in time pressure produces an exponential characterization of the detrimental effects on the overall accuracy of the iconic identification.

Summary and Conclusions

Two experiments are reported on the ability of operators to correctly glean the states of a complex military iconic object. In the first experiment, it was demonstrated that humans have certain perceptual biases to features presented which impacts how iconic objects should be designed. In the second experiment, as time pressures are increased in the study, there was an exponential loss of performance due to the limited capacity of humans to glean information from complex visual renderings, such as military iconic objects.

Acknowledgments

This work was supported, in part, by the AFOSR grant: 313HC54, "Complexity of Glyphs and Systems" and the USAF CPE (Commanders Predictive Environment) Program.

References

Alexander, A. L., Wickens, C. D. & Merwin, D. H. (2005). Perspective and coplanar cockpit displays of traffic information: implications for maneuver choice, flight safety, and mental workload. *The International Journal of Aviation Psychology*, 15(1), 1-21.

Benbassat, D. & Abramson, C. I. (2002). Landing flare accident reports and pilot perception analysis. *The International Journal of Aviation Psychology*, 12(2), 137-152.

Cummings, M. L. & Tsonis, C. G. (2006). Partitioning complexity in air traffic management tasks. *The International Journal of Aviation Psychology*, 16(3), 277-295.

Endsley, M. R. & Kris, E. O. (1994). Information presentation for expert systems in future fighters aircraft. *The International Journal of Aviation Psychology*, 4(4), 333-348.

Hooey, B. L. & Foyle, D. C. (2006). Pilot navigation errors on the airport surface: Identifying contributing factors and mitigating solutions. *The International Journal of Aviation Psychology*, 16(1), 51-76.

Lichacz, F. M. J. (2005). Examing the effects of combined stressors on dynamic task performance. *The International Journal of Aviation Psychology*, 15(1), 45-66.

Metzger, U. & Parasuraman, R. (2006). Effects of automated conflict cuing and traffic density on air traffic controller performance and visual attention in a datalink environment. *The International Journal of Aviation Psychology*, 16(4), 343-362.

Mogford, R. H. Murphy, E. D. & Guttman, J. A.

(1994). Using knowledge exploration tools to study airspace complexity in air traffic control. *The Inter. Journal of Aviation Psychology*, 4(1), 29-45.

Nikolic, M. I., Orr, J. M. & Sarter, N. B. (2004). Why pilots miss the green box: how display context undermines attention capture. *The International Journal of Aviation Psychology*, 14(1), 39-52.

Prichett, A. R. (2000). Display effects on shaping apparent strategy: a case studying collision detection and avoidance. *The International Journal of Aviation Psychology*, 10(1), 59-83.

Prinzel, L. J., Comstock, J. R., Glaab, L. J., Kramer, L. J. & Arthur, J. J. (2004). The efficacy of head-down and head-up synthetic vision display concepts for retro-and forward-fit of commercial aircraft. *The Inter.l Journal of Aviation Psychology*, 14(1), 53-77.

Prinzo, O. V. (2003). Pilot's visual acquisition of traffic: Operational communication from an in-flight evaluation of a cockpit display of traffic information. *The International Journal of Aviation Psychology*, 13(3), 211-231.

Thomas, L. C. & Wickens, C. D. (2006). Display dimensionality, conflict geometry and time pressure effects on conflict detection and resolution performance using cockpit displays of traffic information. *The International Journal of Aviation Psychology*, 16(3), 321-342.

Ververs, P. M. & Wickens, C. D. (1998). Head-up displays: Effect of clutter, display intensity and display location on pilot performance. *The Inter. Journal of Aviation Psychology*, 8(4), 377-403.

Wiegmann, D. A. & Shappell, S. A. (2001). Human error perspectives in aviation. *The International Journal of Aviation Psychology*, 11(4), 341-357.

Wiggins, M. W. & O'Hare, D. O. (2003). Expert and novice pilot perceptions of static in-flight images of weather. *The International Journal of Aviation Psychology*, 13(2), 173-187.

Zhang, J. (1997). Distributed representation as a principle for the analysis of cockpit information displays. *The International Journal of Aviation Psychology*, 7(2), 105-121.

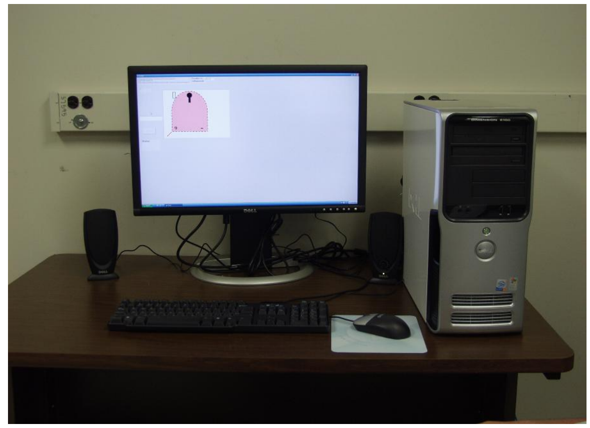

Figure 1. The experimental setup to run subjects

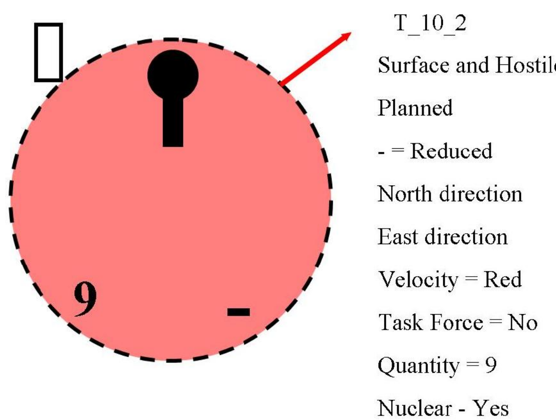

Figure 2. A ten dimensional military icon

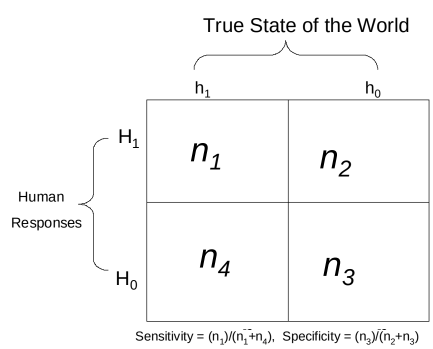

Figure 3. Truth Table for Experiment 1

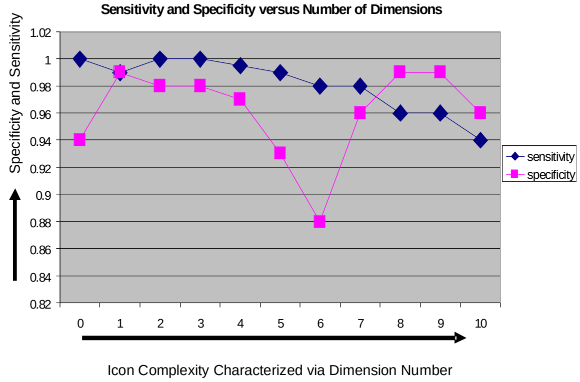

Figure 4. Sensitivity – Specificity for Experiment 1

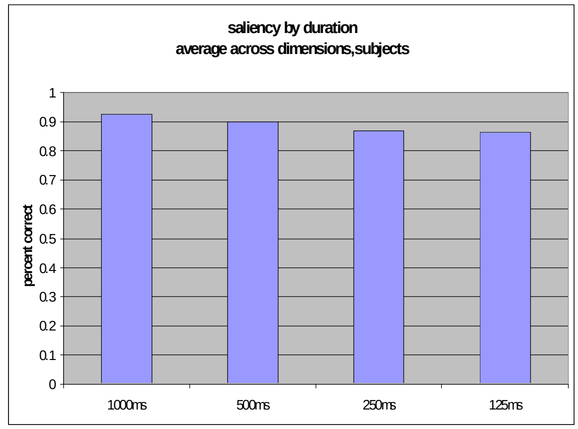

Figure 7. Averaged Accuracy vs. time duration

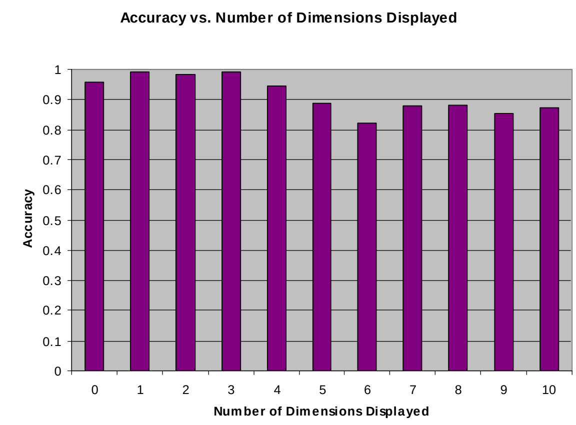

Figure 5. Accuracy vs. Dimension Number – Exp. 1

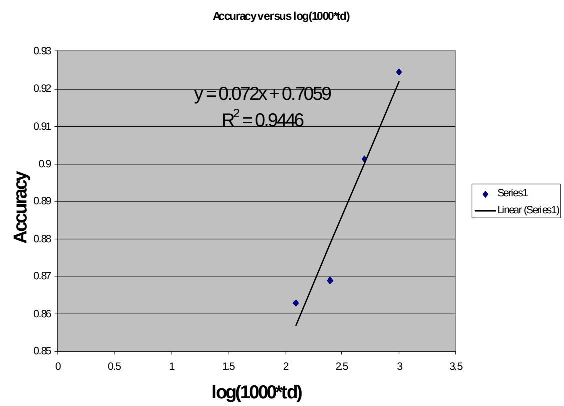

Figure 8. Experiment 2 – Accuracy – vs. Log(time)

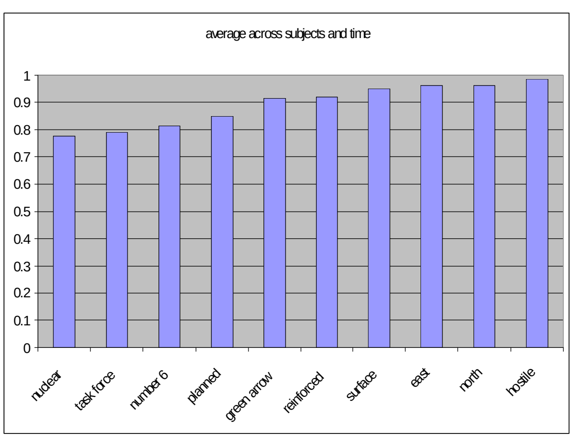

Figure 6. Accuracy versus complexity – Exp 2.