Your attention please: an evaluation of animated visual signals for ATC alarms and notifications

Dr. Sylvie Athènes, Dr. Stéphane Chatty, Alexandre Bustico
Centre d'Etudes de la Navigation Aérienne
7, avenue Edouard Belin
31055 Toulouse cedex, France
{athenes, chatty, bustico}@cenatoulouse.dgac.fr

ABSTRACT

Alarm design on air traffic control displays is a growing concern as events to be notified grow in number and diversity. Safety requires visual notifications that can be efficiently detected. It also requires that no information on the radar display is obscured by the visual notifications. We also need to design hierarchies of notifications, from most severe to begnin. Taking advantage of the current graphical capabilities of computers, we have explored various designs for transparent visual alarms. This led us to identifying dimensions in visual alarm design. We present in this paper a first experimental evaluation, in terms of detection time, of several of those dimensions: opacity, size, temporal profile of animation and signal frequency. From our results, we conclude that opacity and size are well suited to introduce some nuances in the way we convey notifications on visual displays.

INTRODUCTION

When designing user interfaces for command and control systems such as air traffic control workstations, the safety of the device itself and the adequacy and efficiency of users' actions are not the only critical issues for safety. The way information can be perceived and interpreted by users is also important to safety, especially when it comes to notifying alarms or other asynchronous events. With aircraft flying one nautical mile in ten seconds, one does not want alarms to be noticed in a matter of minutes, but rather in a matter of seconds. And in any case, one does not want alarms to go unnoticed! But whereas there are well known methods for assessing and predicting hardware safety and there is a important mass of scientific knowledge on interaction styles, there is comparatively little to be found in the literature on the efficiency of signals in modern user interfaces. This is a concern for engineers, who are more and more exposed to complex alarm design.

Over the last years, the evolution of computing systems has led to two major evolutions in the design of user interfaces for air traffic control systems: more alarms, and more sophisticated displays. On the one hand, as the underlying systems are growing more complex, new notifications have to be conveyed to users. There are new kinds of alarms associated to new types of breakdowns or problem detection, such as a ground collision risk detection. But there also are less serious events that nonetheless need to be notified, such as asynchronous communications or changes in the status of aircraft. For instance, incoming requests from aircraft through an air-ground data link have to be responded to at some point, and thus need an appropriate notification. This is new for interface designers, who so far had only to provide for a handful of severe alarms and could rely on blinking labels. A wider range of signals is now required.

On the other hand, the evolution of computing systems has provided designers with an extremely wide range of visual techniques and allowed the design of more and more visually rich displays. Between 1985 and today, air traffic control engineers have explored many ways of displaying new pieces of information on graphical radar screens: multi-windowing, colored backgrounds for geographical and meteorological information, enriched flight labels with color coding, etc. Avoiding visual clutter has become very difficult. If no real visual design is exercised, adding the necessary range of alarms (see above) to such a display can lead to disastrous results. Actually, as notifications are not considered different from other types of information in most textbooks and design guides, designers often handle them as just another piece of information to be displayed. Thus, displayed messages may obscure other pieces of information, or may go unnoticed, or both. For instance, at the CENA we have had examples of realistic simulations where controllers focusing on an area of their large screen did not notice an alarm based on color change for as long as 20 minutes!

In order to ensure alarm perception, designers tend to add sounds to notify information. Sound is indeed an efficient medium, but it cannot be used in every contexts. Air traffic controllers are usually reluctant to use sound, and prefer to keep it for emergency alarms. Therefore, other means have to be found for lower priority notifications. We believe that the graphical capabilities of computers can be used to design such solutions.

It is well known that peripheral vision is very sensitive to movement. It is also well known that painters and other visual artists know how to influence the way the eye analyses an image. Modern computers provide a significant subset of the visual techniques that are available to painters or cartoon authors. But surprisingly, very few of those capabilities have been used in air traffic control: in contrast, video games for instance make a far more extensive use of them. Among these visual techniques, most interesting is the possibility to design animations that efficiently raise attention without obscuring background information. The goal of our research is to explore the mechanism of animation efficiency, and to identify the salient design dimensions. This should allow us to provide engineers with hierarchies of visual signals to be mapped on hierarchies of alarms. In this article, we identify some dimensions, and report an experiment that provides preliminary results on their compared salience.

NOTIFYING INFORMATION

In the literature, there are classical results about visual alarm design (for a review, see[1]). However, a vast majority of these results is based on light displays and the parameters tested pertain mostly to color, shape or location. For example, the guidelines in "Human Factors in the Design and Evaluation of ATC Systems"[2] specify color (red for warning/danger, yellow for caution and green for normal/ready status), location (the central 15 degrees of the area where the controller normally looks) and blinking (no more than 2 levels, rate between 2 and 3 Hz).

More recently, researchers have stepped further away from the characteristics of visual perception *per se* and focused on the cognitive mechanisms of alarm awareness [3, 9]. Wickens [10], for example, defined a set of basic principles allowing for a better readability of the signals and their relationship in the situation at hand. Particularly efficient are the following principles: sequencing, where it should be possible to recover the order in which a series of signals progressed; grouping and prioritizing, where the spatial organization of the signals on the display should be closely related to the functional organization of the system; color, where the conventional meaning of colors should be congruent with their use (e.g. red for danger); informativeness, where each signal should be designed to provide higher-level combined information.

Lately, research work in the so-called ecological perspective has been advocating for a more integrated representation of information. In this approach, the fundamental unit of analysis is the human-environment system (or the human-machine system, as the case may be). Typically, these studies question the relevance and format of the information presented to the user [4]. Their designs attempt to exploit the large capacity of the human perceptuo-motor system as it is found in natural environments by supplying a complex, but transparent information environment. A good example is provided by head-up displays for aircraft pilots where design has focused on representation of those features of the optical array that support perception of speed, height, and flight direction [6].

DESIGNING ATTENTION GETTERS

User interface design guides tend to be conservative and address mostly the issue of color, symbol, font, location and blinking rate. We believe that we should go beyond a mere transposition of classical results. Today's graphical screens allow production of complex graphical effects that go far beyond the capacities of light signals. Their endless variations allow the design of signals for a whole range of events of different priorities. As the amount of dialogues with the machine increases, it becomes important that the user's attention is not grabbed all the time, but rather attracted discriminately. Furthermore, as the amount of information to be conveyed increases, it becomes crucial to keep fluidity of work and to avoid upstaging of the main task [5].

We are thus exploring a series of graphical attributes or effects that could support a large range of signals with distinctive attention "grabbing" capabilities. Among the possible attributes are classical parameters like the shape, size and dominant color of the graphical signal. As we are looking for signals that do not clutter the display or obscure other pieces of information, transparency as provided by modern graphical systems is also an interesting dimension. The transparency level of the visual signal can be used as well as the transparency gradient that determines its contrast with the background. Finally, animation provides a whole set of dimensions. The type of animation is probably meaningful though difficult to characterize: pulsating objects, moving objects, etc. Then for periodic animations as for any other periodic signals, classical variables are signal frequency and profile, as illustrated in figure 1.

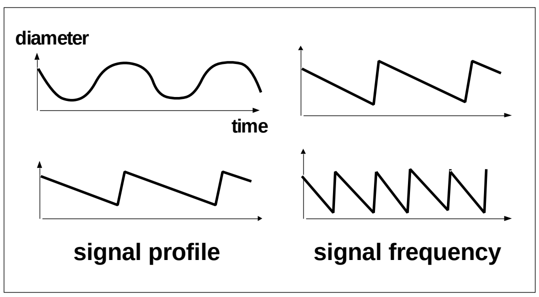

Figure 1. Variations on the way the signal evolves over time. On the left, variation on the temporal profile. On the right, variation on frequency.

The present paper reports on a subset of the above mentioned parameters. Because both shape and color have been extensively explored, we decided to keep these two parameters as neutral (or all-encompassing) as possible and to use only one shape (a circle) and one color (yellow). We chose a circle because of its relative symbolic neutrality (as opposed to the "caution" meaning of the triangle in some cultures for example) and the color yellow because it is the color with the largest perception field. We manipulated the transparency and the animation of the signals. Among the animation parameters, we chose frequency and temporal profile. We used pulsating signals, with two variations: "local" signals were small pulsating disks, whereas "non-local" or "global" signals were ripples that encompassed the whole screen, with the goal of catching the eye wherever it rests.

EXPERIMENT

We have carried out an experiment in order to assess the influence of the signals we have listed. The experiment consisted in a series of random signals that subjects had to detect and acknowledge by pointing at them on a touch screen, as illustrated on figure 2.

Design and procedure

Six right-handed subjects took part in the experiment. They were seated in front of a 21" screen equipped with a touch input layer, as shown on figure 2. The eyes-screen distance was 60 to 70cm (24 to 28 in.), depending on the length of the subject's arm.

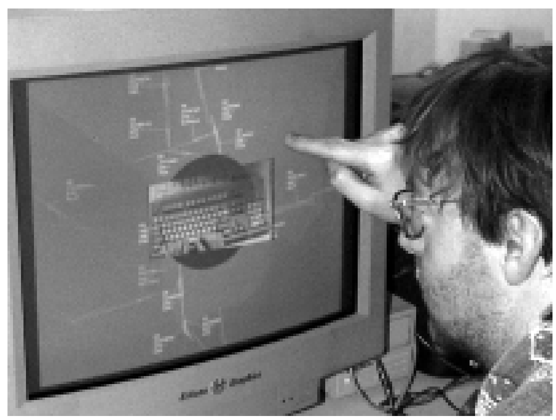

Figure 2. The experimental setup.

To evaluate the capacity of the computed signals to raise attention, we used a dual-task paradigm: signals appeared randomly at the periphery of the touch screen, while the subjects attended a Stroop color-word test displayed in the screen central part (see Figure 1). The Stroop Color-Word Test, which has now been in existence for over 60 years, yields highly reliable and stable measures of individuals' performance. For our purpose, the main interest is that reliable individual differences on each of the three scores obtained from the Stroop test maintain the same rank order of magnitude for all subjects. In other words, no matter how good or bad a subject is, his/her performance always follows the same pattern of degradation from the easiest to the most difficult Stroop task. In this experiment, the underlying assumption is that the more difficult the task, the more attention it requires. Using the Stroop test provided us with a tool to grade the amount of attention required from the subjects and a wealth of published results to check the reliability of our data (for a review, see [6, 7]).

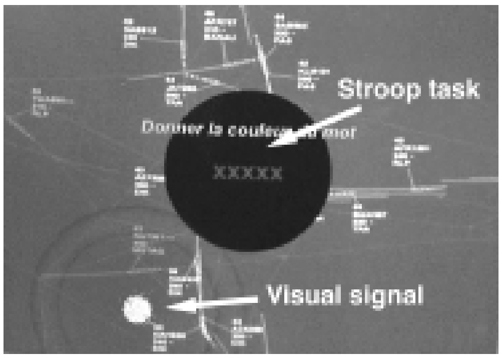

Figure 3. The experimental display: the Stroop task in the middle, and visual signals at the periphery. Here, the signal is a non-local stylised ripple converging toward a small disk.

Detection task. The signals were pulsing yellow circles whose diameters increased and decreased between 10 and 25mm. In addition, they were varied systematically along 4 dimensions: transparency (18, 30 and 50% of opacity), signal frequency (0.75, 1.50 and 2.50 Hz), temporal profile (wave or step, see Figure 1), and the local/global aspect of the signal display (the yellow circle, alone or surrounded by animated rings/waves (see Figures 3 and 4). There were 10 repetitions for each of the 36 signal configurations.

For each subject, the order of appearance of the signals was randomized over the total amount of signal configurations tested during one session. Their location was also randomized. In other words, in the course of a session, each signal could appear anywhere on the screen with the exception of the central part dedicated to the Stroop task. This allowed us to avoid biases due Fitts' law or the organization of the visual field. Subjects were instructed to point towards the signals as soon as they were detected. However, they were forbidden to visually explore the screen. Reaction time and pointing accuracy on the touch-screen were measured.

Stroop task. The Stroop color-word test consisted of words or non-words displayed in black or in color in the central part of the touch-screen. The test entailed three levels of difficulty, from the easiest to the most difficult: subjects were instructed to, either read aloud color names printed in black (e.g., the word 'red' printed in black: correct answer is red), or name the ink color of non-words (e.g.: 'xxx' printed in green: correct answer is green), or name the ink color of color names while ignoring the word (e.g., the word 'red' printed in green: correct answer is green). Consistent, significant performance degradation occurs when conflicting color words and color inks are used and subjects attempt to name the color of the ink.

Subjects were instructed to keep their attention focused at all time on the Stroop test (as opposed to visually explore the screen and look for signals) and to perform that task as quickly as possible while keeping a minimal error rate. The procedure was self-paced, a new word or non-word (depending on the condition) appearing as soon as the preceding one was read or color-named.

The experiment was divided into three one-hour sessions (one per day over three days). During each session, corresponding to one of the three conditions of the Stroop task, subjects were presented with ten repetitions of each signal configuration (i.e. 10x36), randomized over the session.

RESULTS AND DISCUSSION

We only present results pertaining to reaction time (i.e. the time needed for subjects to react to a signal while attending the Stroop task). For this variable, we performed an ANOVA with a within-subject design.

Firstly, the attentional load (as influenced by the level of difficulty of the Stroop task) does indeed have an effect on the reaction time to a signal ($p < .02$): reaction times to identical signals increase with the level of difficulty of the Stroop task. Additionally, we observe a significant main effect for three signal parameters:

- transparency ($p < .0001$): more transparent signals yield longer reaction times;
- temporal shape ($p < .02$): slow increase of the signal surface, as opposed to a step-like increase, entails a longer reaction time;
- local/global aspect ($p < .0001$): local signals yield longer reaction times.

However, some interaction effects allow for a richer description of the observed behavior.

- global signals decrease reaction time, but only for the very transparent signal ($p < .0001$).
- the level of difficulty of the Stroop task increases differentially the reaction time of the subjects ($p < .0001$). Coarsely, its effect diminishes also as the opacity of the signal increases ($p < .05$).
- more opacity decreases reaction time ; however, for some subjects this effect decreases as one gets closer to 50% opacity ($p < .05$).

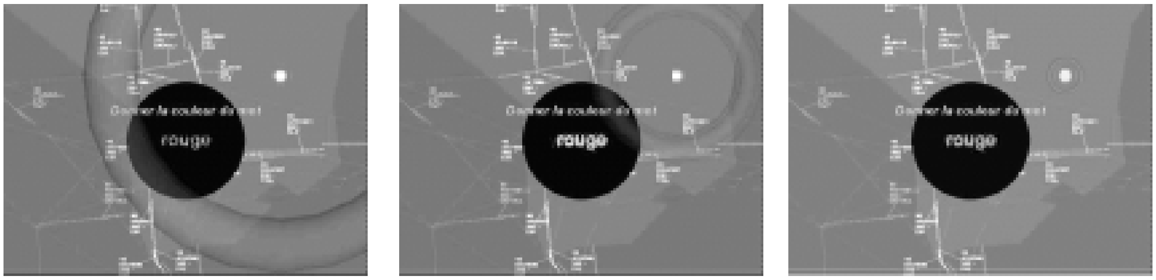

Figure 4. A non-local animation: a ripple crosses the whole screen and converges toward the point of interest.

To sum up, although not all subjects were equally sensitive to the parameters we controlled, there are three variables worth manipulating: the temporal shape of the animation, the transparency, and the local/global aspect of the signal. The differences in reaction time depend on their combination. Transparency is the trickiest, but the most promising of the parameters: although one should probably aim for opacity levels below 20% and be careful to finetune the levels appropriate to the required effects, the interaction of this parameter with the attentional load makes it a good candidate for notifications and suchlike low priority information. Further, the “enhancement effect” of the transparency parameter upon the other tested parameters hints towards specific usage where transparency could be a mode to be chosen as opposed to a set parameter. This feature could be useful when increasing workloads lead controllers to mentally postpone or tune out a subset of events. Considering that the nature of the task was to expect and react quickly to signals, differences are likely to be higher in a real life situation, therefore enabling us to design attention getters with a modulated effect. Indeed, preliminary results on the error rate indicate that the most transparent signals are more frequently missed. Therefore, such signals are appropriate for notifying events of very minor urgency, because they are to be noticed only when the user scans the screen.

CONCLUSION

This paper is a first step toward the elaboration of design principles based on the use of graphical transparency and animation. In spite of complex relationships between the parameters we presented here, this first experiment allowed us to determine which dimensions had the most important effect on visual detection. We now need to incorporate further dimensions and to explore their interactions with the parameters already tested. We will then be able to propose hierarchies of visual signals and to rate them in terms of detection time and probability. Further work also includes the introduction of sound, first as a way to shorten even further the signal detection times, and then as a mean to reinforce or enrich the effect of visual signals through the design of multimodal signals.

REFERENCES

1. Boff, K.R. & Lincoln, J.E. (Eds.): Engineering Data Compendium, Human Perception and Performance. Harry G. Armstrong Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio, 1988.
2. Cardosi, K.M. & Murphy, E.D. (Eds.): Human Factors in the Design and Evaluation of Air Traffic Control Systems. Federal Aviation Administration, Final Report, April 1995.
3. Edworthy, J. & Adams, A.: Warning Design: A Research Propective. Taylor & Francis, 1996.
4. Flach, J., Hancock, P., Caird, J. & Vicente, K.J. (Eds.): Global Perspectives on the Ecology of Human-Machine Systems. Lawrence Erlbaum Associates, Hillsdale, 1995.
5. Harrison, B.L. & Vicente, K.J.: An experimental evaluation of transparent menu usage. Proc. of CHI'96, Vancouver, Canada..
6. Jensen, A.R. & Rohwer, W.D.: The Stroop Color-Word test: a review. Acta Psychologica (25), 36-93, 1966.
7. MacLeod, C.M.: Half a century of research on the Stroop effect: an integrative review. Psychological Bulletin, Vol. 109, No.2, 163-203, 1991.
8. Rasmussen, J. & Pejtersen, A.M.:Virtual ecology of work. In “Global Perspectives on the Ecology of Human-Machine Systems”, Flach, J., Hancock, P., Caird, J. & Vicente, K. (Eds.), Lawrence Erlbaum Associates, Hillsdale, 1995.
9. Stanton, N. (Ed.): Human Factors in Alarm Design. Taylor & Francis, 1994.
10. Wickens, C.D.: Engineering Psychology and Human Performance, HarperCollins, 1992.