Visual Attention Techniques in the Graphical User Interface

Julie Wright, Shumin Zhai, Ted Selker, and Sabra-Anne Kelin

IBM Almaden Research Center
650 Harry Road, MS NWE/B2
San Jose, CA 95120 USA
{amethyst, zhai, selker}@almaden.ibm.com, sakelin@mit.edu

ABSTRACT

This paper presents an experiment in using visual techniques, called "masks", to direct a user's attention within a computer graphical user interface. Bleaching, darkening, and a solid-color pattern overlay (screening) are used to de-emphasize background material, causing the target to visually "pop-out" at the user. The tradeoff between effectively directing the user's attention and ensuring the readability of the background material is explored. Experimental results indicate that there is a wide range of darkening and screening levels that can create a pop-out effect without degrading the readability of the masked area.

Keywords: attention, pop-out effect, help systems, visual masking, user interface techniques

1. INTRODUCTION

Computers can present users with a seemingly endless variety of options and data. This complexity, which can easily overload users, has lead to the use of techniques such as 'wizards,' simplified, step-by-step graphical user interfaces that break up tasks into small, easy-to-digest parts. Consider, for instance, the task of configuring a printer. In a traditional design, a single computer dialog might be constructed that allowed specification of various printer attributes such as the printer's name, the printer driver, the output port, and so on. This would suit an experienced computer user because simple changes could be made directly, in a single step. In the case of a wizard, on the other hand, a sequence of computer dialogs would be created, one per attribute. Explanatory text is often added to the dialog to aid the user. This would more likely suit a novice user who might appreciate the extra guidance provided. The simplicity of the wizard approach, however, is often gained at the expense of the flexibility and efficiency more traditional graphical user interfaces can provide. A wizard often locks the user into a particular sequence of choices that may not be suitable for a particular user.

The tradeoff between simplicity and flexibility/efficiency suggests a third graphical user interface approach that combines both the simplicity of wizards and the flexibility and efficiency of a more traditional interface. One such an approach is illustrated by the Cognitive Adaptive Help (COACH) agent system[5], which annotates traditional user interfaces for inexperienced users (Figure 1). In COACH, dialogs are partially covered with a bitmap (mask). Only one step in the task, which should be what the user pays attention to, remains uncovered at a time, but the entire dialog is visible 'through' the mask. At each step, COACH provides help information and navigational control in an external bubble. The user steps through the task by using 'forward' and 'back' navigational controls on the external bubble. As the user gains experience with the task, the mask is removed to reveal the original dialog. The transition is eased because the user has actually been using the original dialog from the beginning.

Directing a user's attention by masking (de-emphasizing) parts of the visual display has applications beyond help or tutor systems. Many examples exist in mainstream applications, such as Lotus Word Pro and Microsoft Windows 95. When a user chooses to quit Microsoft Windows 95, the desktop is covered with a pattern of black dots, making it appear darker than the confirmation dialog. This directs attention to the confirmation dialog. In Word Pro, the brightness of colors on inactive icons or controls is seemingly reduced by an overlay pattern of the background color, causing them to look 'washed-out'.

Masked objects may retain their readability so as to provide orientation, access to obscure or advanced options, and recall of previous settings. Yet by creating perceptual distance between the highlighted (unmasked) and de-emphasized (masked) areas, it is possible to direct the user's visual attention to the emphasized area.

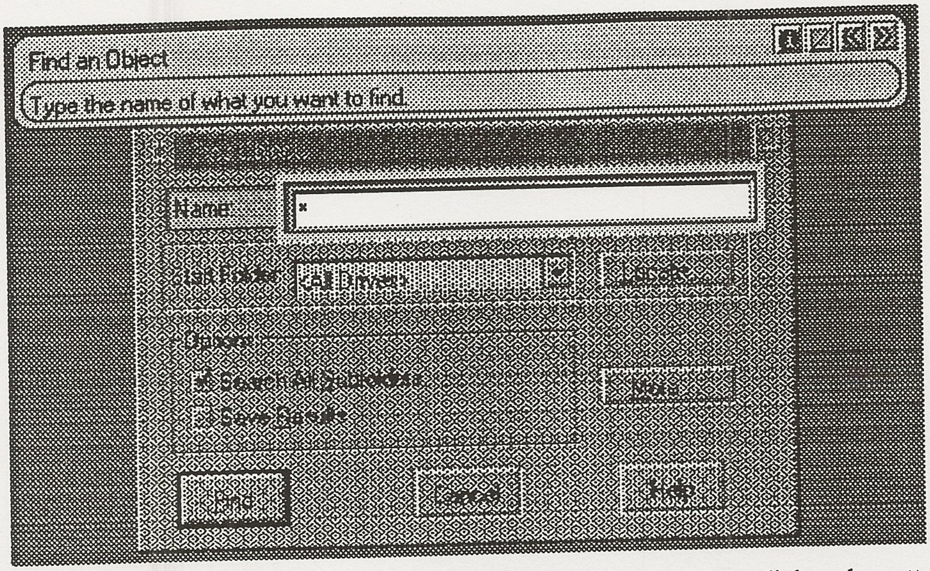

Figure 1: COACH annotation of the OS/2 Find dialog. The "help bubble" above the dialog, the pattern covering most of the dialog, and the lighter box around the name field have been added by COACH to the original dialog.

2. TECHNIQUES

An object (target) will stand out in a graphical user interface if it is sufficiently different from its surroundings (background). There are two general approaches to differentiating the target object. One approach is to examine the surrounding objects in order to identify a strategy that will work well in a given situation. For example, the objects could be examined for similarities in color, size, shape, and so on to determine which attribute is the most similar among the objects. If the objects in the background were found to be of similar size, the target object could be made to stand out by making it larger than those in the background. Or, if the background objects were found to be mostly gray or dull colors, the target object could be made a bright color, such as red, causing it to stand out because of the difference in color. There are at least two problems with this approach. First, it is not predictable to the user. Second, it might be difficult to determine a good strategy if the objects were not similar in any way.

We might think of the first approach as choosing a dialog and adapting a strategy to the dialog. A second approach is the same thing in reverse: choosing a strategy and adapting the dialog to the strategy. For this second approach, the difference between the target and background is preselected. If necessary, the dialog would be altered such that the strategy would be effective. For example, if our strategy is to make the target object appear larger than the background objects, the size of the background objects might be limited to a maximum size less than the target object's size, and the objects would be scaled accordingly. Or, if our strategy was to make the target object brighter than the background objects, we could dull the colors of all the background objects and turn the target object a bright red. Unlike the first approach, this can be predictable to the user and can work even in dialogs that contain a very diverse set of objects. This approach also avoids careful inspection of the background objects for similar features, making its application more straightforward. For these reasons, we have focused our exploration of graphical techniques on this second approach.

Many graphical techniques exist that can produce a "pop out" effect between the target and background objects. Color, luminance, contrast, area, spatial frequency, and temporal considerations are among the factors that interact to determine visual perception.[4] A color's pop-out effect is complicated. For example, an orange dot will stand out amidst a mass of yellow dots. An orange dot will also stand out amidst many red dots, but the same orange dot will not stand out among a collection of both yellow and red dots.[1] Consider also a 460-nm (wavelength) spot. Such a spot will usually appear blue. But, if that same spot is made very small, it will look white. Brightness is also a factor. The perceived brightness of a region depends heavily on the brightness of the surrounding area. A lighter background tends to darken, while a darker background tends to lighten. Most importantly, color is also a dimension that application designers rely heavily upon, resulting in wide variations in usage. Not only might the use of color to produce a pop out effect interfere with the application's use of color, but the application's use of color might also interfere with the pop out effect. For example, an object might stand out if the object is shown in color and the background is grayscale, but this hides any color cues that may otherwise be apparent. Or, the application might use red to identify completed operations. If we then try to draw the user's attention by turning the target red, we might cause confusion. Because we are interested in widely applicable solutions, we will avoid color for now.

Groupings of many small elements can form textures that also have a pop-out effect. For example, a vertical bar will stand out in a field of horizontal bars[1]. Elements of different size or complexity, such as a group of plus signs ('+') randomly oriented in a field of randomly oriented minus signs ('-') will also pop-out, but the same phenomenon will not work using a 'T' and an 'L'[3]. If texture were to be used to produce a pop out effect in a computer interface, interface features that were not the current focus would be required to be of similar shape or texture, allowing the difference in texture to stand out. This, too, would be a significant restriction in the application design.

Another common technique in graphic design is to make an object appear larger than the other objects by either scaling the object or adding a border such as a box or circle. Scaling is especially effective if the background objects tend to be of similar size. However, enlarging the target takes up more space, obscuring part of the background. While this can be minimized by making the effect as small as possible, making the effect too small may also interfere with the pop-out effect.

Several pop-out strategies were considered that did not obscure the background or change the dialog's layout. These techniques included a semitransparent image overlay, applying a gaussian blur to the background, diffusing the background to create a "crayon" effect, distorting the background to create a "wavy" look, changing the brightness and contrast to darken or lighten the background, and overlaying a periodic pattern of pixels onto the background. The pixel overlay effectively reduces the resolution of the background image by placing a "screen" of pixels on top of the background image. Each of these techniques was applied to a sample dialog for evaluation.

Three techniques were chosen for further experimentation. The techniques explored in the following experiment manipulate brightness, contrast, and resolution of the background objects. By making the background more homogeneous, these techniques are intended to produce a perceptual difference between what is masked and what is unmasked. We were interested in testing if these approaches would produce effective pop-out effect without damaging the legibility of the masked objects.

3. EXPERIMENT

3.1 Design

Four masking conditions were tested: bleaching, darkening, screening, and no-masking (see Figure 2 for the samples used). Before conducting the full experiment, several sample dialogs were created; each masked using a different technique. Preliminary experimentation indicated that the techniques of bleaching and darkening might be effective. Bleaching had the effect of washing out the colors of background objects, causing the brighter colors of the target to stand out. Darkening had the effect of placing a tinted transparent film over the background objects. The addition of the screening technique, which overlaid a uniform, periodic, yellow pattern onto the masked areas, was motivated by its low cost of implementation: only a percentage of pixels have to be replaced with yellow color.

This experiment tested five levels (from 1 to 5) of masking with for each technique were tested. The five levels of bleaching were created by reducing the contrast (-30, -45, -60, -75, -95). The five levels of darkening were created by increasing the contrast (+50) and decreasing brightness (-20, -40, -60, -80, -100). The levels of screening were created by replacing a pattern of pixels with yellow pixels (RGB 255 255 85) such that a percentage of the pixels was blocked out (12%, 25%, 37%, 62%, 75%). Note that the intervals within each technique were not necessarily equal. The values were

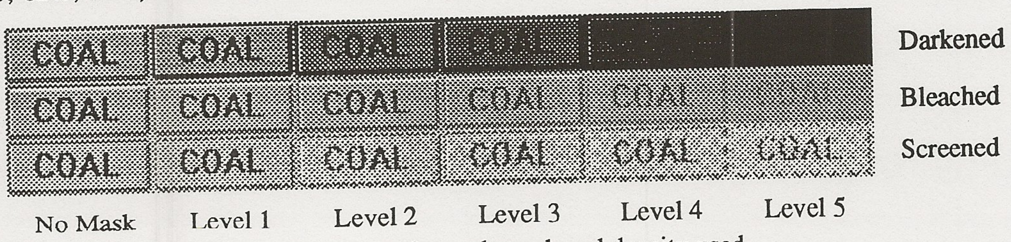

Figure 2: Sample buttons for each mask and density used.

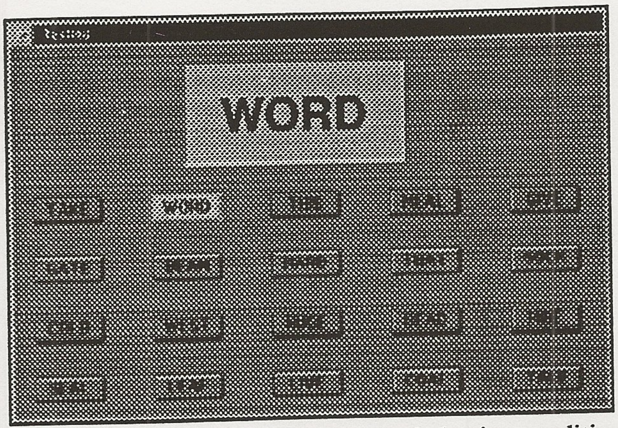

Figure 3: Test dialog for experiment (in darkening condition)

chosen and adjusted to cover as wide as possible a range of appearance (see Figure 2) according the authors' subjective experience and pilot testing. The lowest levels of masking appeared sufficiently close to no-mask condition and the highest levels of masking made the masked words nearly indistinguishable.

A target matching task was designed for this experiment. Subjects were instructed to select, as quickly as possible, the button that matches the target word at the top of the dialog as shown in Figure 3. The advantage of this experimental scenario was that all three aspects of masking could be evaluated. First, a subject's performance in locating the target word when no mask was applied served as the normal, control condition. Such a normal condition can also be considered as Level 0 masking with any of the three techniques. Second, when the target word was highlighted by various levels of masks applied to other words, the pop-out effect could be measured in comparison to the control condition. Third, when the target word was masked together with all other buttons, the readability of the masked area could be evaluated.

A total of 13 subjects performed the task under the normal control condition as well as all five levels of all three masking techniques. For each level in each masking technique, two sets of trials were run. In one set of trials, the target word was left unmasked (therefore highlighted) while the rest of the buttons were masked. In the second set, the entire dialog was masked, including the target word.

Each set of trials consisted of 5 practice trials followed by 20 test trials. Each of the test trials in a set corresponded to one of the 20 words (buttons) in the dialog (Figure 3), so all sets of trials encompassed the equal total travel distances. The order of the buttons was changed randomly between sets to avoid memorization of a button's placement.

The order of the masking conditions was determined by a Latin-square sequence of the darkening, bleaching, and screening techniques. The normal control condition was randomly inserted into the Latin-square sequence. For a given masking condition, the order in which the densities were tested was random.

During pilot testing, subjects tended to move closer to the computer screen under the more difficult conditions, presumably to compensate for increased difficulty. It seemed possible that subjects could maintain the same performance by exerting extra effort in more difficult conditions. A between-subject factor was introduced to explore this possibility. Six of the subjects were allowed to move their head position freely and seven subjects were prevented from moving closer than 1.5 feet by a string across the chest.

3.2 Results

3.2.1 Masking by Darkening

Figure 4 illustrates the performance results of the darkening technique. First, let us examine the pop-out effect when the target was the highlighted/unmasked word. The bottom curves showed that, in comparison to Level 0 (no masking), selection time was significantly reduced when the rest of window was masked (darkened) to various levels: $F_{5, 55} = 110.9$, $p < .0001$, once again showing that darkening is an effective means of creating a pop-out effect. Interestingly, even the slightest level of darkening tested appeared to make the unmasked object "pop out": mean selection time changed from 3.9 seconds

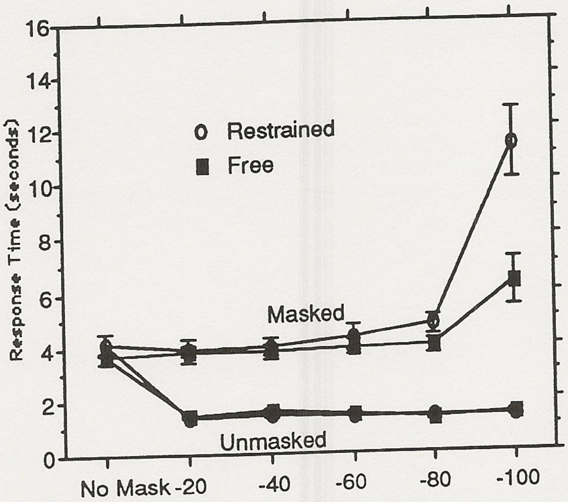

Figure 4: Mean task completion times and standard errors when the target was (top curves) masked or (bottom curves) highlighted / unmasked using the darkening technique.

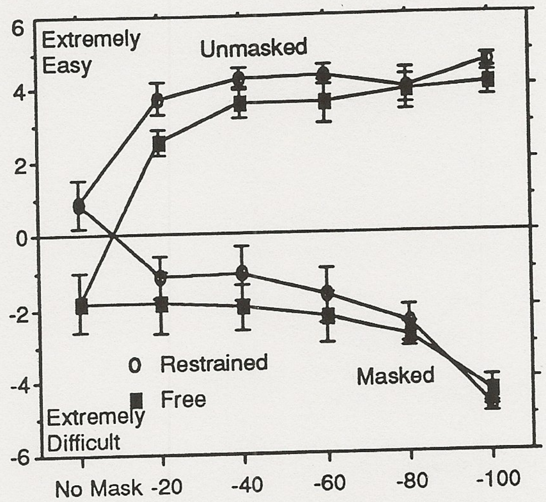

Figure 5: Subjective ratings with the darkening technique. +5 is extremely easy and -5 is extremely difficult.

at no darkening to 1.4 second at Level 1 (-20) darkening. It appeared that even the lowest level of masking enabled the subjects to be "cued" to, instead of searching for, the target. Further darkening did not make subjects perform much faster: mean selection time changed from 1.41 seconds at Level 1 to 1.36 second at Level 5. Such a difference was not statistically significant. Subjective rating data, however, appeared to be more sensitive to the darkening level increase (Figure 5): the higher levels of darkening appeared to make the unmasked target more distinguishable and were rated significantly easier than the lower levels of darkening. For example, a T-test showed $p < .0001$ significance between Level 1 (-20) and 5 (-100).

Second, as shown by the top curves in Figure 4, subjects' performance of reading the masked area was not significantly affected by darkening level except at the very extreme (Level 5). T-tests showed that selection in the masked area at the highest level (-100) was significantly slower than the other levels ($p < .0001$), but the differences among other levels tested were relatively small (from 3.9 seconds at Level 0 to 4.4 seconds at Level 4, statistically insignificant). It is remarkable that subjects maintained similar performance for such a wide range of apparent difficulty (see Figure 2). Subjective rating scores on selecting masked targets moved towards "extremely difficult" as darkening level increased (Figure 5). Not only was highest level (-100) rated significantly more difficult than other levels ($p < 0.0001$), but levels 3 and 4 were also rated more difficult than Level 0 (no darkening): $p < 0.05$.

Viewing condition (Restrained vs. Free moving) was weakly significant to selection time in masked area: $F_{1,11} = 4.4$, $p = 0.06$. By moving closer to the screen as one strategy to compensate the increasing difficulty, the free moving group had faster selection time (only) at the very dark levels. There was a strong interaction between viewing condition and darkening level ($F_{5,55} = 8.6$, $p < 0.0001$) in selecting masked targets. This was the only case (in all performance and subjective data with all masking techniques) where viewing condition made a statistically significant difference in the experiment.

To summarize, the results showed that except at the very extreme, a wide range of darkening levels can create a pop-out effect without significantly damaging the readability of the masked objects, although as the darkening level increases users may subjectively experience a slight increase in pop-out effect and a slight decrease in readability (compare Figure 4 and 5). Subjects tended to maintain the same level of performance by exerting additional effort, as indicated by the smaller time increase in selecting masked targets by the free moving group (Figure 4).

3.2.2 Masking by Bleaching

As shown in Figure 6 (bottom curves), the lowest level (-30) of bleaching tested did not appear to be effective enough to direct the subject's attention. Although it caused significantly shorter selection time than without masking (mean time 3.89 vs. 3.33, $p < .005$), stronger bleaching produced still shorter selection time: $p < .0001$. This suggests that Level 1 bleaching used in the experiment did not produce the pop-out effect needed to fully utilize the viewer's pre attentive perception. For the rest of the bleaching levels, mean selection times were (from Level 2 to Level 5) 1.65, 1.33, 1.38, 1.39 seconds, no

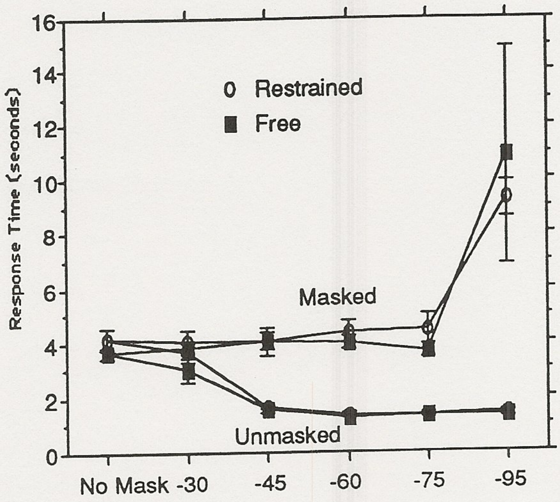

Figure 6: Task completion times when the target was (top curves) masked or (bottom curves) highlighted / unmasked using the bleaching technique.

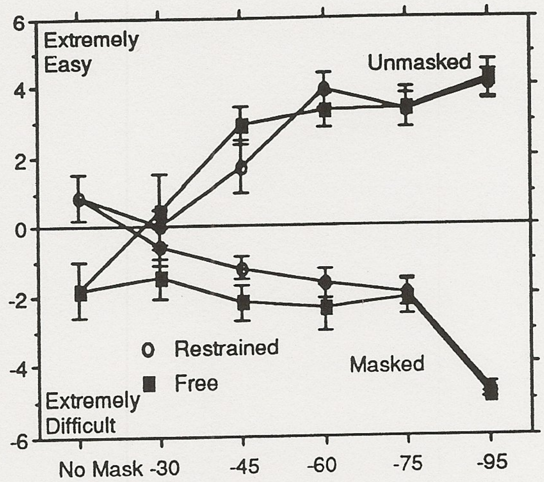

Figure 7: Subjective ratings with the bleaching technique. +5 is "extremely easy" and -5 is "extremely difficult".

significant difference was found among these means. Subjective rating indicated that as bleaching level increased, selecting the unmasked target became significantly easier (see Figure 7).

The readability of the bleached area was similar to that of the area masked by darkening: selection time at the very extreme (-95) was significantly different from the rest ($p < 0.0001$). Differences among Level 0 (no mask) to Level 4 (-75) were not significant, although subjective rating indicated increased difficulty.

Overall, in comparison to the darkening technique, a much narrower range of bleaching levels can accommodate both the pop-out effect and readability (compare Figure 6 against 4, Figure 7 against 5), although the best performances achieved by the bleaching technique were similar to those achieved by the darkening technique (about 4 seconds for selection in masked area and 1.4 seconds for selecting unmasked target).

3.2.3 Masking by screening

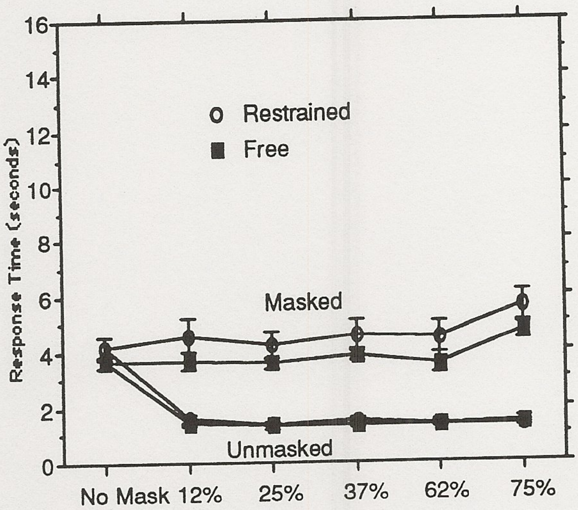

Figure 8: Task completion times when the target was (top curves) masked or (bottom curves) highlighted / unmasked using the screening technique.

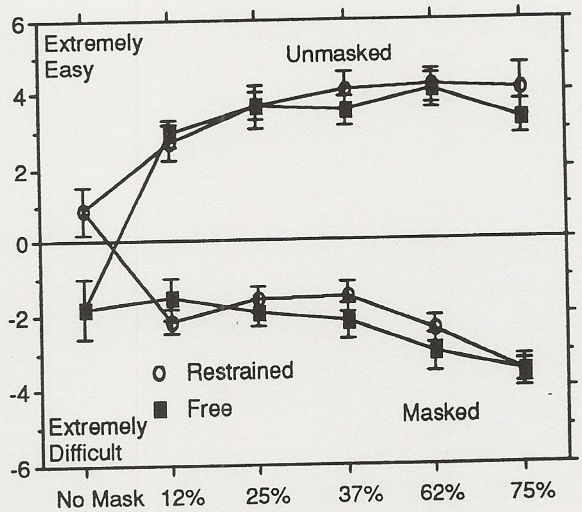

Figure 9: Subjective ratings with the screening technique. Horizontal axis is the degree of screening. +5 is "extremely easy" and -5 is "extremely difficult".

As shown in Figure 8, even the slightest level of screening tested (Level 1, 12% pixels screened out) appeared to provide a strong cue: mean selection time of unmasked targets changed from 3.9 seconds at Level 0 screening to 1.5 second at Level 1 screening ($p < 0.0001$). Further screening did not appear to be much more effective; mean selection time changed from 1.5 seconds at Level 1 to 1.38 seconds at Level 5 (not statistically significant). Subjectively, even the lowest level of screening (level 1) made the task significantly easier ($p < 0.0001$). Further increase in screening level made the task seem even easier ($p < 0.05$) (Figure 9).

As the percentage of pixels masked out increased, the readability of the masked area was expected to decrease. Indeed, T-tests showed significantly different ratings from lower levels to higher levels (e.g. $p < 0.05$ from Level 2 to Level 4) (see Figure 9). However, the time of selecting the target word from the masked area remained flat except at the extreme Level 5 (Figure 8). From Level 0 to Level 4, mean selection time changed from 3.9 to 4.1 seconds (not statistically significant). Only Level 5 was significantly different from the rest ($p < 0.0001$). This once again showed the user's ability to overcome increased difficulty without significantly reducing time performance.

Overall, a wide range of screening, from 12% of pixels to 62% pixels being masked out, produced the optimal performance (again 4 seconds for selection in masked area and 1.4 seconds for selecting unmasked target).

4. DISCUSSION

All three techniques used in the experiment were effective when adequate levels of masking were selected. The highlighted target selection time at such levels was around 1.4 seconds for all three techniques, much faster than with no masking (around 4 seconds). Bear in mind that motor control (moving cursor to target) was involved in target selection. The visual difference between finding a pop-out target and searching for the same target without masking cues was much greater than 4.0:1.4, the ratio of response times for the two conditions. Furthermore, for several levels of each mask, the time to find an item in the masked area was similar to Level 0 (no mask), suggesting that the readability of the masked areas had been well preserved.

The range of such adequate masking levels that can provide strong cues to direct a user's attention without readability loss in the masked area differed among the three techniques. The flat area of the performance curve where performance did not differ significantly in this experiment. It was -20 to -80 with the darkening technique, -45 to -75 with the bleaching technique, and 12% to 62 % for the screening condition. Bleaching provided a narrower optimal range than the other two techniques provided. Darkening and screening both worked well, although they offer very different masking appearances and require different implementation mechanisms for interface designers to choose.

Although subjective data tend to vary from individual to individual, the rating data collected in the experiments did offer information not revealed by the performance data. The rated difficulty changed faster than the performance measurement as masking level moved to the two extrema. It is interesting that this increase in subjective difficulty did not correspond to immediate performance loss. Subjects tended to compensate for the increased difficulty, perhaps by exerting extra effort. See a textbook such as one by Wickens[6] for related theories of performance and resources.

Note also that the experimental scenario is simpler than real world applications where both the visual features and the user's cognitive activities are much more complex. The target and background objects in the experiment were practically identical in every way except that the masking effect was not applied to the target object. Preliminary testing involving a more complex dialog consisting of several different types of dialog controls (bitmap buttons, entry fields, etc.) found no significant effect of object type. However, the greater diversity of a more realistic dialog might reduce the effectiveness of the mask, requiring a stronger masking level than would be needed for a more homogeneous dialog. A narrowed range of effective masking level is expected in the more complex applications.

5. CONCLUSION

Focusing a user's attention in a computer's graphical user interface can be used to provide automated assistance to the user or to convey the urgency of a given task. An experiment was conducted to evaluate means of drawing the user's attention to various parts of a computer dialog. Objects were de-emphasized by applying a visual effect, called a "mask", causing the unmasked objects to stand out. Of the three masking techniques studied, two of the techniques, darkening and screening out the background objects, can provide a wide range of masking levels that could produce strong, possibly pre-attentive, visual cues to pop-out the targets and still allow the user to read the masked background area. Masks with as little as 12 % pixels being screened out could produce a strong pop-out effect without sacrificing the readability of the background. Bleaching out the background objects was not as widely effective.

The studies presented here indicate that important information can be made to pop-out using darkening or screening masks. Darkening or screening the less important parts of the user interface provides strong attention cues to the important information. Yet the non-important information can still be accessed. Such a way of directing a user's attention by masking is a graceful human machine communication method: tell the user what the user interface agent "thinks" is important, while allow the user act according to his own will.

ACKNOWLEDGMENTS

We would like to thank Ron Barber, Paul Maglio and other members of the USER (User System Ergonomics Research) group at the IBM Almaden Research Center for their contributions to the development of this research.

REFERENCES

1. D'Zmura, M. Color in Visual Search. *Vision Research, 31, 6* (1991), 951-966.
2. Harrison, B., and Vicente, K. An Experimental Evaluation of Transparent Menu Usage. *Proceedings of CHI 96: Human Factors in Computing Systems* (1996), 391-398.
3. Julesz, B., and Bergen, J. R. Textons, the Fundamental Elements in Preattentive Vision and Perception of Textures. *The Bell System Technical Journal, 62, 6* (July-August, 1983), 1619-1645.
4. Rogowitz, B. E. The Human Visual System: A Guide for the Display Technologist. *Proceedings of the SID, 24, 3* (1983), 235-252.
5. Selker, T. COACH: A teaching agent that learns. *Communications of the ACM 37, 7* (1994), 92-99.
6. Wickens, C. D. *Engineering Psychology and Human Performance, Second Edition* (1992). HarperCollins Publishers, Inc.