Wolfe, J. M. (2005).Society for Information Display, International Symposium Digest of Technical Papers, 2, 1395-1397.

41.1: *Invited Paper:* How Might the Rules that Govern Visual Search Constrain the Design of Visual Displays?

Jeremy M. Wolfe

Visual Attention Lab, Brigham and Women's Hospital & Department of Opthalmology, Harvard Medical School, Cambridge, MA, USA

Abstract

In visual search attention can be guided toward a target item among distracting items by a limited set of basic features. This talk will summarize the rules of effective guidance (e.g. you can efficiently guide attention toward a red and vertical item but not a red and green item).

Humans are good at searching for things in a crowded visual world. Once, we searched the savannah for predators and prey. Today, we search the cabinet for a mug, the countertop for the coffeepot, and the refrigerator for the cream. Moreover, we create stimuli for others to search through - from web pages to medical images to panels of instrumentation. The rules, forged over evolutionary time, apply to these searches of our own time. The purpose of this paper is to review some of those rules in the hope that knowledge of the limits on human search will help in display design.

1. Basic Visual Search

To begin, if an observer is asked to find a target (e.g. an "F" in Figure 1a) among a number of similar distractors (e.g. "5"s), that observer will behave as though she is processing those items at a rate of one every 25-50 msec [1]. This assumes that the observer does not need to move her eyes to each item in turn because eye movements proceed at a much slower rate of 3-4 sec. It also assumes that each item can be recognized rapidly once attention is directed to it. This is "face-in-the-crowd" search where one face is much like all others. Why don't we have to search in this laborious fashion most of the time? The core of the answer is that we can use simple information to *guide* our attention to more complex targets.

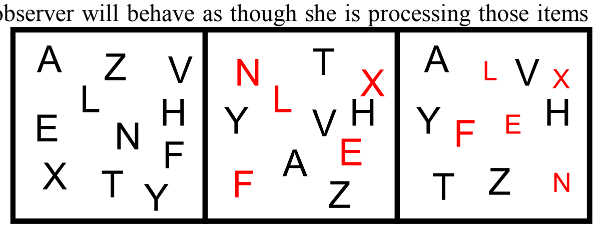

Figure 1: Find the "F" in each panel. In 1b, it is red. In 1c, it is big and red.

2. Guided Search

If basic feature information differentiates the target from, at least, some of the distractors, search can be speeded. Thus, in 1b, observers searching for a "red F" can guide attention to the red items, avoiding the black [2]. If there is more information, guidance can be more effective. In Fig 1c, the target is a "big, red F". Attention can be guided toward the intersection of the set of big items and red items - an excellent place to look for big red items [3].

3. Basic Features

Only a limited number of stimulus properties can guide attention. The exact list is a matter of debate but there are probably not fewer than a dozen guiding features and not more than two dozen [for a recent list, see 4]. The most efficient searches are those with a target item defined by a single basic feature (Fig. 2a). Some features are fairly obvious candidates like color and orientation. Others, like line termination [5] are less obvious and many, seemingly reasonable, candidates like intersection type do not work [Fig. 2c, ref 6].

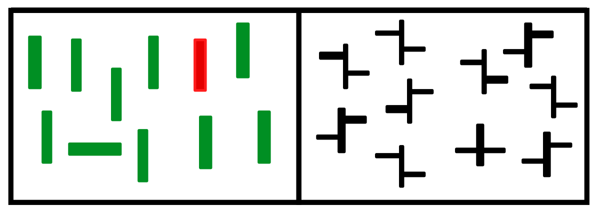

Figure 2: Color and orientation "pop-out" Intersection type (X vs T) does not.

4. Top-down vs. Bottom-up guidance

Basic features in a display can guide attention in either of two modes. Top-down guidance is based on the observer's attentional set (e.g. "look for red"). However, an item might attract attention without the observer needing to know anything about that item in advance. This is <u>bottom-up</u> guidance and is based on the local differences between an item and its neighborhood. A red item among green will tend to attract attention even if the observer wasn't looking for anything. The appearance of a new object, marked by a luminance onset, is perhaps the gold standard for a bottom-up guiding stimulus [7]. Strong bottom-up guidance is sometimes referred to as "attentional capture" [8]. However, you cannot assume that your flashing ad will grab eyeballs on a webpage. Observers can usually learn to avoid capture by new stimuli if they form a top-down attentional set to do so [9].

5. Similarity

If you are trying to attract attention, it is important to keep in mind that bottom-up guidance is based on differences. First, the larger the target-distractor differences, the more salient the target. That point is illustrated in the top panels of Fig. 3. Second, the larger the differences among distractors, the *less* salient the target will appear. The distance, in color space from targets to distractors is the same in the upper right and lower left of Fig. 3, but the target is hard to find in the lower left because of distractor heterogeneity [10]. Distractor differences are particularly problematic when, as in the lower left, the target lies between the distractors in color space[11, 12]. In the lower right, everything is salient and nothing attracts attention.

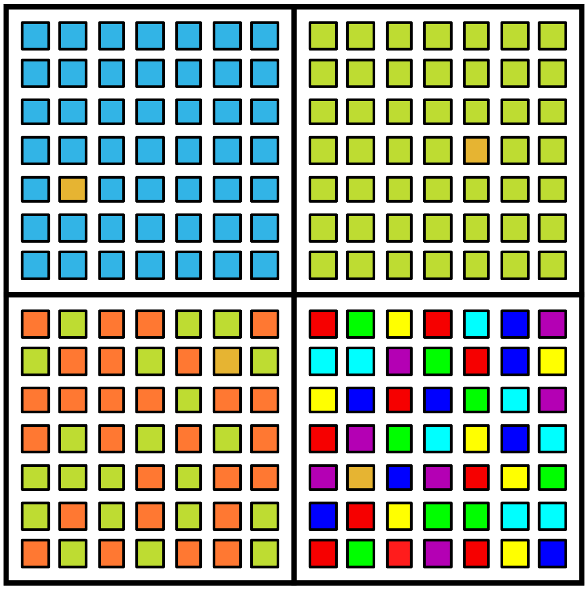

Figure Three: Why is the pale orange square easier to find in some quadrants than in others?

6. Categorical processing

The rules of bottom-up stimulus salience are fairly intuitive. Less obvious is the fact that top-down guidance has a very limited "vocabulary". Consider orientation. While it is easy to tell if a line is vertical, it is not easy to find a vertical line among lines tilted 20 deg to the left and right (4a). It is easier to ask that attention be guided to items that are defined *categorically*. In orientation, the categories are steep and shallow, left and right. Thus, in 4b, search for the 10 deg tilted item is fairly easy since it is the only "steep" item. In 4c, search for the 10 deg tilted item is harder because it is neither uniquely "steep" nor uniquely "left-tilted". Note that all the relationships between items are the same in 4b and 4c [13]. The items are simply rotated 20 deg counter-clockwise in 4c (and, of course, the target is repositioned). In size, the guiding vocabulary seems to be just "big" and "small". In 4d, it is hard to find the medium-sized target.

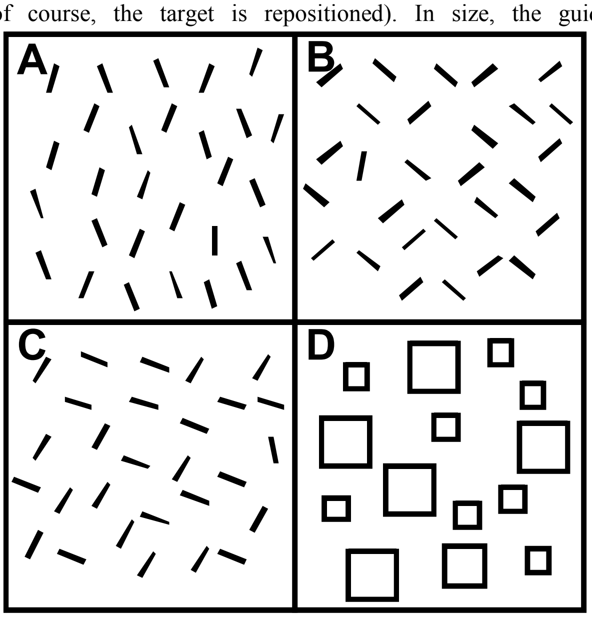

Figure Four: Targets are vertical (4a), 10 deg from vertical (4b&c) and medium-sized (4d).

Of course, this does not mean that multiple orientations, sizes, colors, etc. cannot be used in a display. However, the choice of features is constrained when it is important that users be guided to one item without being distracted by others.

7. Crowding

Just as the deployment of attention is coarsely coded to a feature attribute like orientation, it is also coarsely coded in space, This can be illustrated by the phenomenon of "crowding". Look at the center of Figure Five where it says "Look Here". If you now attend to the pattern at the top of the figure, you will be able to tell that there is a set of discrete lines. However, you will probably find yourself unable to count the lines. Somehow, you cannot get your attention to limit itself to just one line [14]. If you keep your eyes in place and shift attention to the letters at the bottom of the figure, you will find yourself unable to read the letters in the middle of the array. As with the lines, you are unable to restrict your attention to a single letter when it is crowded by others. This is not a limit in your acuity because you will notice that you can read the isolated "S" and, indeed, you may be able to read the letters at the corners of the clump of eight letters [15].

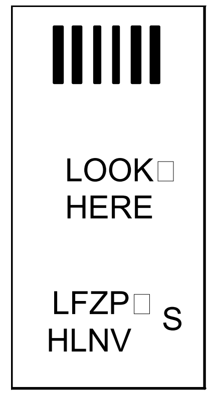

Figure Five: Crowding

8. Combining features

Though the deployment of attention is guided by fairly coarse mechanisms in both space and feature domains, the news is not all grim. As illustrated in Fig 1c, we can guide attention based on the combination of multiple sources of information. Under most real-world situations, knowing that a target is a moving, big, red, shiny object is likely to guide attention to only a few items even if the specifications of "moving", "red" etc. are quite crude. Under most circumstances, the crude descriptors will limit the set of possible targets to a small subset of all items. The exceptions are the "needle in a haystack" searches where all items have much the same set of features.

Guidance to conjunctions is also subject to some fairly severe limitations. Most notably, we do not appear to be able to guide attention to the conjunction of two values drawn from the same feature space. Thus, on the left of Figure 6, it is hard to find the target defined by a conjunction of two colors, red & blue [16]. Interestingly, it is easier if one color is the color of the whole item

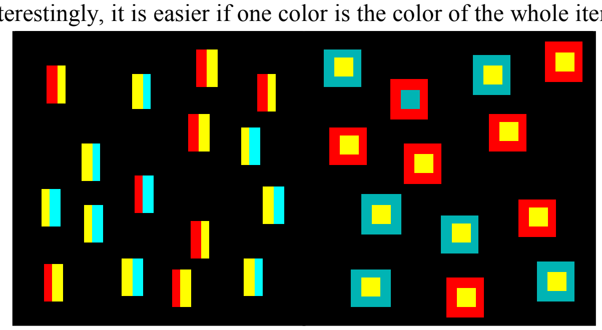

Figure 6: It is hard to guide attention to the item that is red and blue but easier to find the red item with a blue center.

and the other color is the color of a part (The red item with the blue center in Figure 6)[17].

9. Objects

This sensitivity to object structure raises the last constraint for discussion here. All else being equal, attention is guided to objects (though the precise definition of objects is not clear) and that attention spreads throughout an object [18]. Moreover, it proves to be very difficult to guide attention to a specified part of an object. On the left of Figure 7, there is a green horizontal element. It is hard to find because all of the plus-shaped objects contain green and horizontal. The same pixels are present on the right of the figure but now guided search is easier because only one object has the attributes "green" and "vertical" [19].

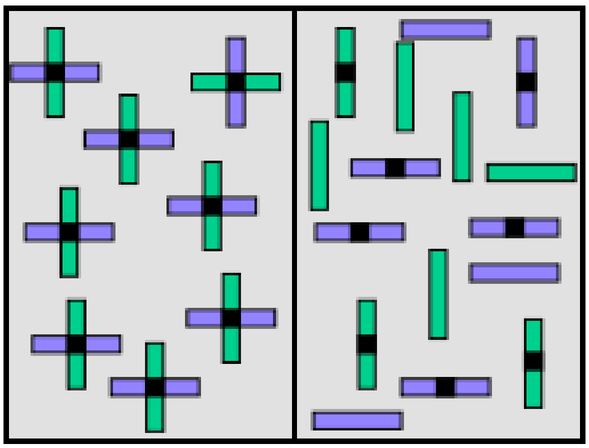

Figure 7: It is hard to find a green horizontal region when it is embedded in objects that all contain green & horizontal features (left). It is easier if there is only one object contains both green & horizontal (right).

10. References

[1] J. M. Wolfe, "What do 1,000,000 trials tell us about visual search?," *Psychological Science*, vol. 9, pp. 33-39, 1998.

[2] H. E. Egeth, R. A. Virzi, and H. Garbart, "Searching for conjunctively defined targets.," *J. Exp. Psychol: Human Perception and Performance*, vol. 10, pp. 32-39, 1984.

[3] J. M. Wolfe, K. R. Cave, and S. L. Franzel, "Guided Search: An alternative to the Feature Integration model for visual search," *J. Exp. Psychol. - Human Perception and Perf.*, vol. 15, pp. 419-433, 1989.

[4] J. M. Wolfe and T. S. Horowitz, "What attributes guide the deployment of visual attention and how do they do it?," *Nature Reviews Neuroscience*, vol. 5, pp. 495-501, 2004.

[5] B. Julesz and J. R. Bergen, "Textons, the fundamental elements in preattentive vision and perceptions of textures.," *Bell Sys. Tech. J.*, vol. 62, pp. 1619- 1646, 1983.

[6] J. M. Wolfe and J. S. DiMase, "Do intersections serve as basic features in visual search?," *Perception*, vol. 32, pp. 645-656, 2003.

[7] S. Yantis, "Stimulus-driven attentional capture," *Current Directions in Psychological Science*, vol. 2, pp. 156-161, 1993.

[8] R. W. Remington, J. C. Johnston, and S. Yantis, "Involuntary attentional capture by abrupt onsets.," *Perception and Psychophysics*, vol. 51, pp. 279-290, 1992.

[9] C. L. Folk, Remington, R W, Johnston J C, "Involuntary covert orienting is contingent on attentional control settings," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 18, pp. 1030-1044, 1992.

[10] J. Duncan and G. W. Humphreys, "Visual search and stimulus similarity," *Psychological Review*, vol. 96, pp. 433-458, 1989.

[11] M. D'Zmura, "Color in visual search," *Vision Research*, vol. 31, pp. 951-966, 1991.

[12] B. Bauer, P. Jolicoeur, and W. B. Cowan, "Distractor heterogeneity versus linear separability in colour visual search," *Perception*, vol. 25, pp. 1281-1294, 1996.

[13] J. M. Wolfe, S. R. Friedman-Hill, M. I. Stewart, and K. M. O'Connell, "The role of categorization in visual search for orientation.," *J. Exp. Psychol: Human Perception and Performance*, vol. 18, pp. 34-49, 1992.

[14] J. Intriligator and P. Cavanagh, "The spatial resolution of visual attention," *Cognit Psychol*, vol. 43, pp. 171-216., 2001.

[15] A. Toet and D. M. Levi, "The two-dimensional shape of spatial interaction zones in the parafovea.," *Vision Research*, vol. 32, pp. 1349-1357, 1992.

[16] J. M. Wolfe, K. P. Yu, M. I. Stewart, A. D. Shorter, S. R. Friedman-Hill, and K. R. Cave, "Limitations on the parallel guidance of visual search: Color X color and orientation X orientation conjunctions.," *J. Exp. Psychol: Human Perception and Performance*, vol. 16, pp. 879-892, 1990.

[17] J. M. Wolfe, S. R. Friedman-Hill, and A. B. Bilsky, "Parallel processing of part/whole information in visual search tasks," *Perception and Psychophysics*, vol. 55, pp. 537-550, 1994.

[18] R. Egly, J. Driver, and R. D. Rafal, "Shifting attention between objects and loctions: Evidence from normal and parietal lesion subjects.," *J. Experimental Psychology: General*, vol. 123, pp. 161-177, 1994.

[19] J. M. Wolfe and S. C. Bennett, "Preattentive Object Files: Shapeless bundles of basic features," *Vision Research*, vol. 37, pp. 25-43, 1997.