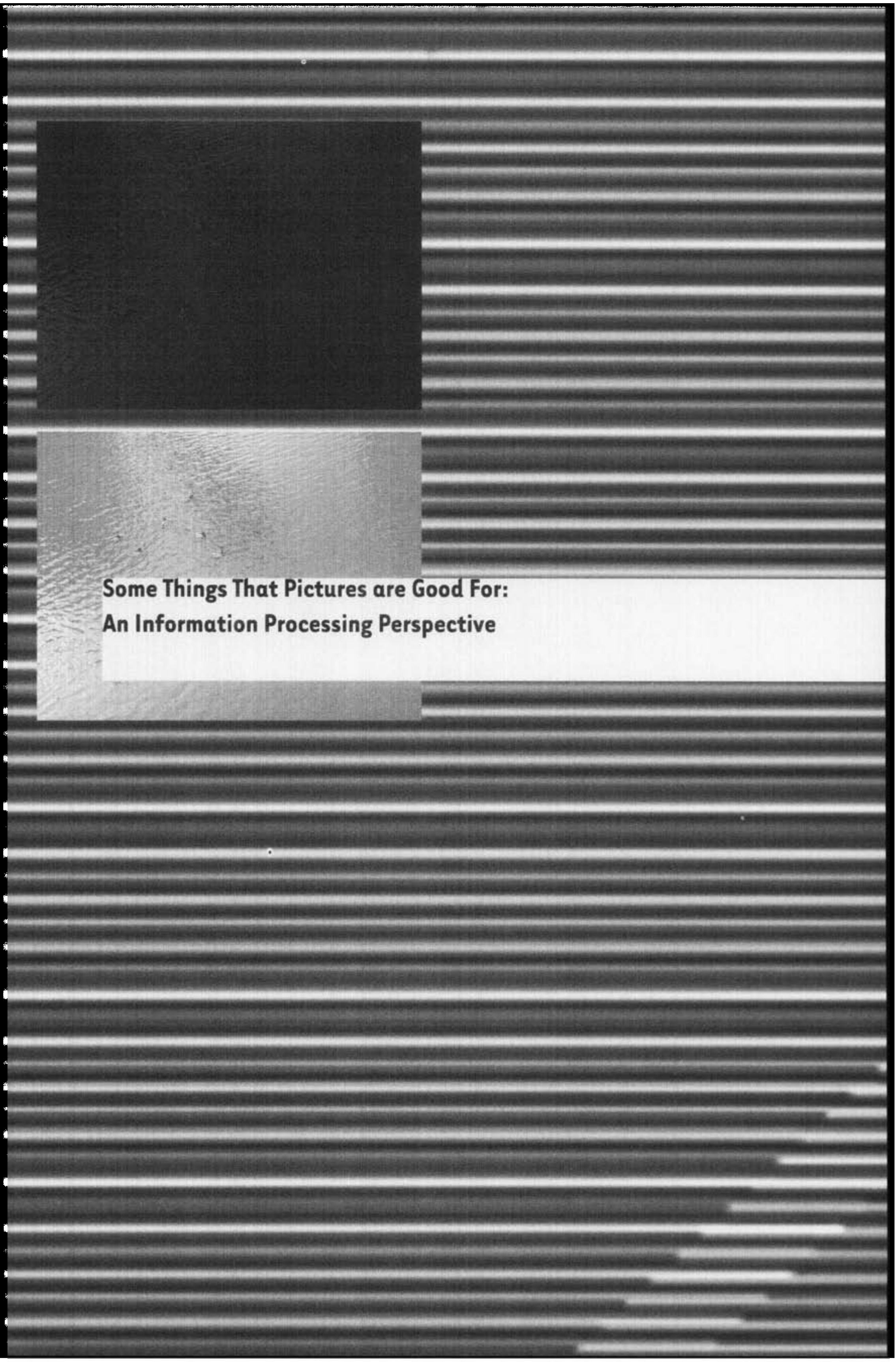
Some things that pictures are good for: An information processing perspective

Lester Loschky

Visible Language; 2001; 35, 3; Research Library
pg. 244



LESTER LOSCHKY

Biography

Lester Loschky is a doctoral candidate in Psychology, studying visual perception and cognition, at the University of Illinois at Urbana-Champaign, USA (e-mail: loschky@uiuc.edu). His current research deals with the interrelationships between visual resolution, eye movements, visual attention and working memory. More generally, his interests are in visual perception, both visual and linguistic compr>ension and the visual arts. This is reflected in the three fields he has studied: perceptual and cognitive psychology (currently at the University of Illinois), second language acquisition and instruction (MA in English as a Second Language, University of Hawaii at Manoa) and the visual fine arts (BFA in drawing and sculpture, Columbia College).

Abstract

Our visual experience of the world is extremely limited in scope both spatially and temporally. This is due to extreme restrictions on our visual attention, our region of high resolution within the field of view and our visual short-term memory, as shown by research on visual perception and memory. However, we have developed very efficient ways of dealing with these limitations. One biologically based scheme is to make rapid eye movements around our visual environment several times per second. This allows us to attend to items in our visual environment serially that we could not attend to simultaneously, and allows us to refresh our leaky visual short-term memories at the same time. A second entirely human invention is to make and view pictures. Pictures have a great capacity for allowing us to direct a person's attention to things they might not have noticed. Pictures also allow us the time to carefully explore visual information by attending to details that otherwise might have disappeared in our ever-changing world. Likewise, because pictures can hold information in a stable form, we don't have to use our limited visual short-term memories to hold onto their contents. Instead, we have the potential to repeatedly look back at any detail whenever the need arises in order to more deeply process its contents without loss of information due to the image changing. In this way, pictures facilitate our contemplation of visual information. Of course, pictures do not remove the inherent limitations on our visual attention, resolution and short-term memory, as clearly shown in the pictorial demonstrations contained in this article. However, pictures do extend our abilities to deal with these limitations in ways that greatly enrich our visual experience.

Human Perceptual and Cognitive Limitations

Those of us blessed with sight live in a rich visual world full of shapes, colors and lines, objects, people and events. Our eyes provide us with new visual information on an almost continuous basis.[1] And, though we rarely stop to think about it, most of us, if asked, would probably say that we are well aware of what is in our immediate surroundings — our visual environment — at any given moment. But psychological research on human perception over the last few decades has come to show that our visual experience of the world is in fact quite limited.

If the last statement above strikes you as obscure, it is quite understandable because most of the time we are completely unaware of how little visual information we are able to take in or hold onto. Nevertheless, a wealth of research suggests that our visual experience has the following characteristics:

1) *Visual attention is extremely limited.*

We pay attention to (and thus consciously experience) only a very limited number of things in our visual environment at any given moment. Those things which we do not attend to can go completely unnoticed even when they are literally right in front of our eyes. This phenomenon has been shown experimentally and is known as "inattentional blindness."[2] In a number of such experiments, a simple visual stimulus (e.g., a small black square, or a white circle, roughly the width of a pencil at arm's length) could be briefly flashed on a computer screen (for about 1/5 of a second) exactly at the center of vision (i.e., directly where participants were looking) and never be noticed by the majority of viewers, if they were doing a task that required them to pay attention elsewhere. Later, after the same people had been tested once in this way and were now more wary, when such a stimulus was flashed again, virtually all of them could correctly identify its shape and location.[3] While such a result may come as a shock at first, it should really come as no surprise based on personal experience. Most of us can probably remember looking in vain for something, perhaps our keys or wallet, only to find after near exasperation that what we were looking for was in plain sight the whole time ("It was right in front of my nose")! Our failure to "see" something that is literally right in

T	W	F	Z	G	X
+	+	+	+	+	+
25	20	15	10	5	0

FIGURE 1.
A Test of Visual Acuity as a Function of Distance From the Center of Vision.
If you hold the test figure at a distance of approximately 20 cm (or 8 inches) from your left eye, and you stare at the plus on the far right (above the number 0) with your left eye, you should have no problem reading the letter above it. However, as you try to read the letters above further points along the line, all the while holding your gaze on the far right plus, you will find it becomes increasingly difficult. Numbers below the pluses are the approximate degrees in visual angle that each plus is from the center of vision. Most people cannot read many letters beyond 10 degrees from the center of vision.

front of our eyes for an extended period of time can be explained by the fact that we can only pay attention to a limited region of our visual field at any given time, often only a single object. If we don't pay attention to something, even though we are looking at it, we will not consciously "see" it.[4]

2) *Visual resolution has severe spatial limitations.*

We can visually perceive only a small region of our visual field in high resolution. The rest of our visual field (peripheral vision) is degraded.[5] This seems in many ways similar to what I just said above, but it is actually quite different.

FIGURES 3 a–i

A Test of Visual Short-term Memory (or "Visuospatial Working Memory")

The test shows a series of everyday objects on a normal background (in this case, tools on a workbench). There are a series of images, each on a separate page. Look for just a moment at each of the images, lettered a–i without looking back at any previous images. When you get to image h, it will ask you which object was at the location marked by a square. Try to remember which object appeared there, and then quickly look at the next image i. That image will show four choices and will ask you to pick the one that you think was at the cued location. Afterwards, selecting your answer, you can check to see if you were right by looking back and finding the object that was indeed at that location. Frequently, viewers report only remembering a single feature of the object. (The scene backgrounds and objects were created by Gregory J. Zelinsky, State University of New York, Stony Brook.)

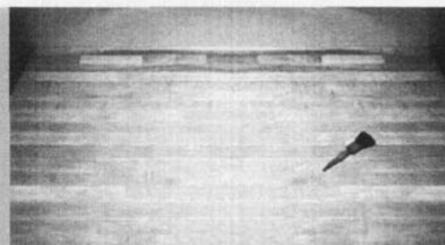


FIGURE 3a

[1] In fact our vision is interrupted briefly on average three to four times a second when our eyes move and during those times we are nearly functionally blind. Matin, E. 1974. Saccadic Suppression: A Review and Analysis. *Psychology Bulletin*, 81:12, 899–917. However, these 'blank periods' on average last only 0.05 of a second and only make up about twelve percent of the time we spend looking. Hallet, P.E. 1986. Eye Movements. In Boff, K. L. Kaufman and J. Thomas, editors. *Handbook of Perception and Performance*, Volume 1. New York: Wiley and Sons.

[2] The first published article looking at this phenomenon was Rock, I., C. Lennett, P. Grant and A. Mack. 1992. Perception Without Attention: Results of a New Method. *Cognitive Psychology*, 24, 502–534. A recent book on the topic is Mack, A. and I. Rock. 1998. *Inattentional Blindness*. Cambridge: MIT Press.

[3] Rock et al. 1992 and Mack and Rock, 1998.

[4] See also Simons, D.J. and C.F. Chabris. 1999. Gorillas in Our Midst: Sustained Inattentional Blindness for Dynamic Events. *Perception*, 28:9, 1059–1074.

[5] For a good introduction to this topic and an explanation based on the structure of the retina, see Thibos, L.N. 1998. Acuity Perimetry and the Sampling Theory of Visual Resolution. *Optometry & Vision Science*, 75:6, 399–406.

FIGURE 2a.
A High-resolution Picture.

FIGURE 2b.
A Multi-resolutional Picture

Version 2b of the picture has been degraded in the visual periphery so that it matches the fall-off of visual resolution of the human visual system. The most degradation is on the left side of the image. If you hold the image at a distance of 32 cm (or 12.5 inches) from your left eye, and look at the cross, you should be unable to detect the degraded portion of the image on the left. (Image produced using an algorithm developed by Dr. Jian Yang of Imaging Research and Advanced Ddevelopment, Eastman Kodak Company, Rochester, NY, USA.)

To get a better idea of precisely what I mean, try the peripheral vision test in figure 1. If you hold the test sheet at the listed distance and keep your eyes fixated at the O mark, you will find that it becomes difficult if not impossible to read the letters that are much further than the 10° distance. Figure 2b shows an application of this fall-off of visual resolution with distance from the center of vision.[6] The image in Figure 2b should look identical to that in figure 2a if you hold page at the required distance and keep your eyes **fixated on the cross. But** once you move **your eyes away from the** cross in figure **2b, you can easily see the** image degradat**ion in the further areas of** the image. This **shows that visual resolu**tion is separabl**e from paying attention,** because you shou**ld not be able to read** the letters in **figure 1, or detect the image** degradation in **figure 2b even when you** know where to pay attention and are trying.[7] The fact that we are only able to see very clearly at the center of vision is one of the chief reasons we constantly move our eyes (more on this later).

3) *Visual short term memory is severely limited.*

We are able to remember only a very limited number of details of what we have just looked at (including the identities of objects and their locations) at any moment.[8] Furthermore, as we look from one object to another in a scene, our short-term visual memory is always best for those things we have most recently

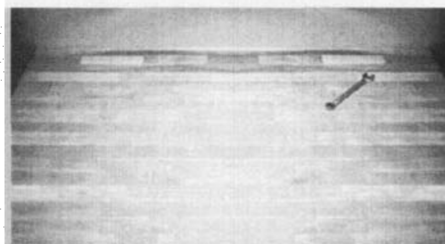


FIGURE 3b

[6] The program that produced the multi-resolutional image in Figure 2b was created by Dr. Jian Yang of Imaging Research and Advanced Development, Eastman Kodak Company. For truly imperceptible peripheral image degradation, it is necessary to take into account the exact luminance and contrast levels of the image, as well as the precise viewing distance and the center of gaze. Thus, the printed image here is only a very rough approximation of what can be achieved with a properly calibrated computer monitor and eye tracking system.

[7] Interestingly, however, it has recently been shown that paying attention to something increases our visual resolution for it, even when it is far from the center of vision. Yeshurun, Y. and M. Carrasco. 1999. Spatial Attention Improves Performance in Spatial Resolution Tasks. *Vision Research*, 39:2, 293-306. Of course, paying attention was not able to completely overcome the fall-off of resolution with distance from the center of vision, rather, it improved resolution for attended objects in comparison to unattended ones.

[8] See, for example, Irwin, D.E. 1992. Memory for Position and Identity Across Eye Movements. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 18:2, 307-317.

looked at, and worse for those things we looked at only a few seconds earlier.[3] You can experience this for yourself by taking the memory test in figure 3. If you are like most people, you will have a hard time with the test, even if you get the correct answer. Indeed, most people are quite surprised to find out how difficult such a seemingly simple task is, and they frequently report only remembering a single feature of the object (e.g., its color, shape, or orientation). Such tests point out how limited our visual short-term memories are; our surprise shows how unaware of this fact most of us are.

Together, the above three characteristics of vision put extreme limits on our moment-to-moment visual experience. Yet, surprisingly, we are almost never aware of these extreme restrictions. It is only when we are tested, either by our circumstances, or, much less commonly, by experiments such as those illustrated above, that we realize how little visual information we are able to take in and retain for any short period of time. Nevertheless, taken together, the above statements suggest that we miss much (or most) of the detail (and potential meanings) in our visual surroundings, and most of those things that we actually pay attention to we quickly forget (except for the general substance of what they represent to us, if anything). Indeed, one might wonder how we manage to get along in life if we do so poorly at such simple visual tasks. Apparently, for our survival, we do quite well enough. We pay attention to those things we deem important, ignoring much of the rest of the scene, abstract the information important to our needs, and forget the rest

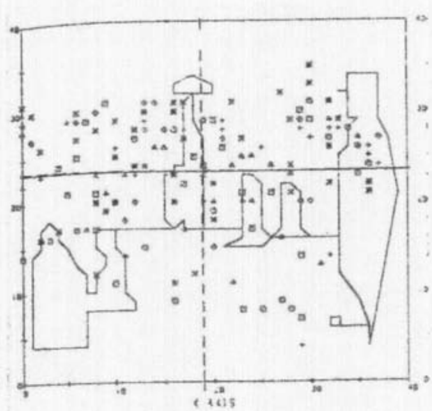


FIGURE 4

Eye Fixation Plot on a Schematic Representation of Georges Seurat's "Sunday Afternoon on the Island of La Grande Jatte."
The symbols represent each of the eight viewers whose eye fixation data are plotted on the graph. The broken lines (both horizontal and vertical) represent the average location of all viewers' eye fixations. The solid line (which almost entirely overlays the horizontal broken line) is the best fitting line through all of the eye fixation points. Both measures show that the fixations center on the central figure of a woman with a parasol. (Courtesy of Calvin F. Nodine and James J. McGinnis. 1983. Artistic style, compositional design, and visual scanning. Visual Arts Research, 12, 3).

of the details. Whenever we need more information or have forgotten something, we simply investigate the environment by moving our eyes to our areas of interest and repeat the above process. Our visual system, though seemingly impoverished, is actually quite economical — indeed, the credo of the visual system should probably be "waste not, want not"![10]

It turns out, however, that we have developed ways of making up for our visual limitations. One that I just mentioned above is biological and behavioral: we move our eyes constantly (normally about three to four times a second) in order to bring those objects that draw our attention into the high resolution area of our visual field. Moving our eyes back to things we looked at before is also useful for refreshing our memory. Since our short-term memory is so leaky, we often need to look again at what we looked at before in order to refresh our memory for what it was we just saw only a moment before.

A second way we have of coping with visual limitations is a purely human invention: making and using pictures.[11] Pictures help us to deal with our limited visual attention and memory spans and, in doing so, pictures greatly enrich our visual experience.

Attentional Assistance

Pictures help us cope with our limited visual attention spans in two important ways: by directing our attention, and by allowing us time to attend at leisure in order to explore visual information. In fact, as we shall see, these are interrelated themes.

FIGURE 3c

[9] For a classic paper on this topic, see the following. Phillips, W.A. 1983. Short-term Visual Memory. *Philosophical Transactions of the Royal Society of London*, B302, 295-309. For more recent work, see Zelinsky, G. and L. Loschky. 2001. Forgetting What We Have Just Seen: Recency Effects for Objects in Scenes Revealed by Eye Movements. (Manuscript submitted for publication.)

[10] For an early discussion of this idea, see O'Regan, K. 1992. Solving the "Real" Mysteries of Visual Perception: The World as an Outside Memory. *Canadian Journal of Psychology*, 46:3, 461-488.

[11] In this paper, I generally use the term "picture" to describe what James Elkins refers to as hypographemic images, i.e., something "close to the ideal of a purely visual image," (90) such as a photograph. However, it seems that the arguments I am making about the value of pictures may more broadly extend to the most general domain of images, referred to by Elkins as "gramma," the Greek term for a "picture, written letter, or piece of writing" (83). Elkins, James. 2000. *The Domain of Images*. Ithaca, New York: Cornell University Press.

Imagine several photographers all gathered in one room taking pictures. Looking at their photographs later, each picture will focus our attention on different things from that same room. This highlights one of the primary reasons we make pictures: to direct people's attention to things we want them to see. Pictures can direct our attention in many different ways including their compositional structure (i.e., the arrangement of objects across the image), by using "pop-out" feature contrasts (i.e., contrast within perceptual dimensions such as color, size, orientation, etc.) and by motion cues in animated images.

Let us first consider compositional devices for guiding a viewer's attention. The most basic compositional device is to put the point of interest in the center of an image. Studies have shown that people looking at pictures spend most of their time looking in and around the center of the images.[12] Of course, this begs the question of why people look at the center of images. Is it because of learned habit, or because high resolution vision is at the center of vision and this somehow maps onto the center of a picture, or because the center is the geometric mean of the image — i.e., we end up visiting the center as our attention travels across the image from place to place, or simply because the most salient objects usually happen to be in the center? In fact, this question remains unanswered, but it is possible that several of the above-mentioned factors interact to give us this "center bias" in pictures.[13]

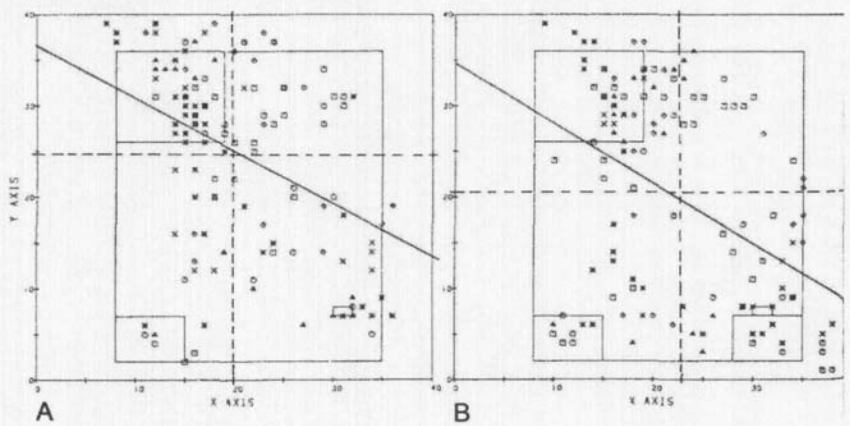


FIGURE 5

A) Eye Fixation Plot on a Schematic Representation of Piet Mondrian's "Opposition of Lines, Red and Yellow."

The symbols represent each of the eight viewers whose eye fixation data are ploted on the graph. The broken lines (both horizontal and vertical) represent the average location of all viewers' eye fixations. The solid line is the best fitting line through all of the eye fixation points, and indicates a diagonally oriented balance of attention in the picture.

B) A Modified Version of The Mondrian Composition.

This version was modified by the authors by adding a rectangle to the bottom right corner. This resulted in changed fixation patterns and thereby lowered both measures. (Courtesy of Calvin F. Nodine and James J. McGinnis. 1983. Artistic Style, Compositional Design and Visual Scanning. *Visual Arts and Research* 12, 4.)

A more sophisticated use of picture composition to guide the viewer's attention in a picture is to attempt to 'balance' the image. From an aesthetic standpoint, a balanced picture is one in which the opposing forces of the picture achieve an equilibrium; from an attentional perspective, a balanced composition is one in which the viewer's attention is more or less evenly distributed throughout the picture, though there will likely still be a bias towards the center and away from the edges.[14] By careful arrangement of the elements in a picture, it is possible to affect the overall balance of where viewers place attention. This was nicely shown in a study of viewer's looking patterns on a number of art images.[15] The study showed that, when eight different viewers looked at a reproduction of Georges Seurat's "Sunday Afternoon on the Island of La Grande Jatte" both the average eye position and the balance line (the best fitting line through all the viewers' eye fixation points) crossed at the central figure (a woman holding a parasol), as shown in figure 4. But, a well-balanced composition will not necessarily focus viewers' attention at the center of the picture. For example, figure 5a from the same study shows that when the same viewers looked at a reproduction of a painting by Mondrian, the eye fixation points were aligned more diagonally. Of particular interest, we see that in figure 5b, when the authors made a slight change to Mondrian's composition (by adding a small rectangle in its bottom right corner), it dramatically changed the viewers' eye fixation center point. Thus, this study clearly supports what artists have argued for centuries, namely that the way an image is composed can alter where viewers look in a picture, i.e., what they pay attention to in it.

Nevertheless, research on the effects of compositional balance on eye move-

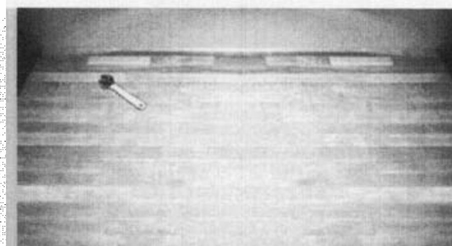


FIGURE 3d

[12] Buswell, G.T. 1935. *How People Look at Pictures: A Study of the Psychology and Perception in Art*. Chicago: University of Chicago Press. See also Mannan, S.K., K.H. Ruddock and D.S. Wooding. 1997. Fixation Patterns Made During Brief Examination of Two-dimensional Images. *Perception*. 26:8. 1059-1072.

[13] Locher, P. 1996. The Contribution of Eye-movement Research to an Understanding of the Nature of Pictorial Balance Perception: A Review of the Literature. *Empirical Studies of the Arts*. 14:2. 143-163.

[14] Locher. 1996.

[15] Nodine, C. and J. McGinnis. 1983. Artistic Style, Compositional Design and Visual Scanning. *Visual Arts Research*, 12. 1-9.

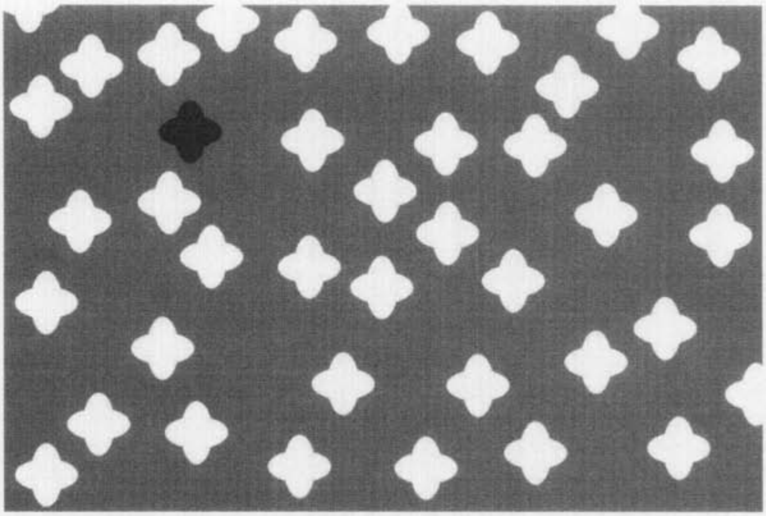


FIGURE 6a

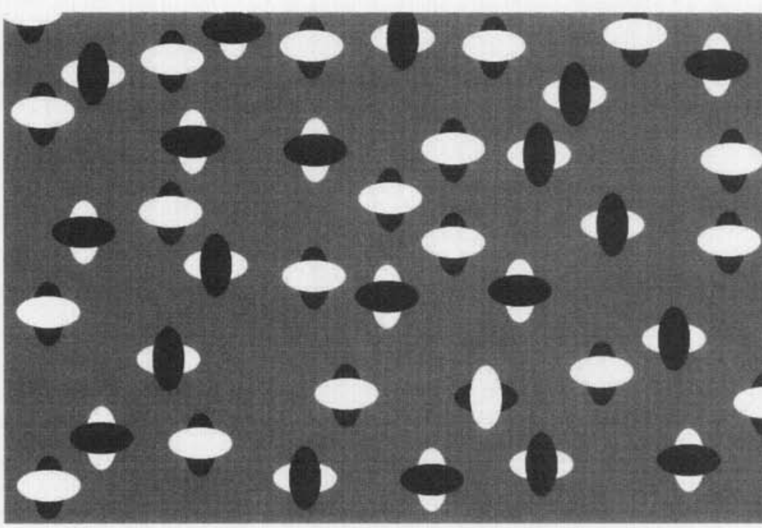


FIGURE 6b

FIGURE 6a

An Image Illustrating the Phenomenon of Feature-based Attentional "Pop-out."

The unique item in the image "pops out" (i.e., immediately captures attention) by virtue of differing in terms of a single feature dimension (in this case, luminance) from all other items in the image.

FIGURE 6b

An image Illustrating the Phenomenon of Feature Conjunction Search.

The unique item in this image is defined by a particular conjunction of features (orientation, luminance, and overlap) that are shared in different combinations by all other items in the display. Thus, the unique item can only be found by a time-consuming attentional search through the image.

ments also indicates that the effects of balance are subtle, and they are stronger for viewers with more experience in viewing art than for naïve viewers.[16] Thus, other, less subtle means may be more effective in directing viewers' attention in pictures. One such unsubtle means of directing attention is to create attentional "pop-out" by contrasting one element in the picture with all other elements along a single perceptual feature dimension. As can be easily seen in figures 6a and 6b, the unique item immediately pops out in the former, but not in the latter. This is because in figure 6a, the unique item differs from all other items on a single feature dimension, brightness. In contrast, in figure 6b, the unique item differs from all other items in terms of a unique conjunction of features, in this case brightness, orientation and overlap.[17] A striking example of using feature pop-out in a more complex image can be found in a long tracking shot in the film, "Schindler's List." It is an aerial shot of the Warsaw ghetto, and we see thousands of people walking in all directions. A single figure of a small girl, however, is colored, while the rest of the figures are monochrome. This instantly draws the viewer's attention to the girl from among the thousands of other figures in the shot. There appear to be a number of feature dimensions within which pop-out can occur, including size, orientation, color, brightness and perceived depth in the plane.[18]

Finally, a particularly powerful method of attracting attention in a picture is through motion, which explains why waving at someone is often the best way to get their attention. We can explain this in terms of the connections between low-level motion detection mechanisms in the brain and a phylogenetically primitive brain area, the superior colliculus, which guides attention and eye move-

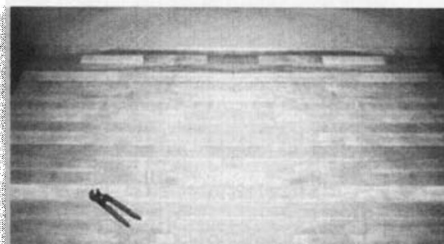


FIGURE 3e

[16] Locher, 1996.

[17] Two important papers on this topic follow. Treisman, A.M. and G. Gelade. 1980. A Feature-integration Theory of Attention. *Cognitive Psychology*, 12:1, 97-136. Wolfe, J.M. 1994. Guided Search 2.0: A Revised Model of Visual Search. *Psychonomic Bulletin & Review*, 1:2, 202-238.

[18] Wolfe, 1994.

FIGURE 7

A "Spot the Difference" Picture Pair.

The difference between the two image versions is much larger than most viewers realize at first. In fact, roughly twenty-five percent of the image area of each version of the picture differs from the other version. There are two probable reasons why it is so difficult to perceive all the differences between the images: 1) the limited extent of visual attention at any given point in time allows only a few objects to be recognized and 2) the fact that the differences do not change the 'gist' (i.e., overall meaning) of the scene means they do not stand out. (Images courtesy of Gregory J. Zelinsky, State University of New York, Stony Brook.)

ments.[25] Indeed, this connection is particularly strong in lower animals such as frogs, for whom visual motion is the preeminent cue for orienting attention.[26] Thus, in animated/moving pictures, the viewer's attention can be most effectively guided to an area of interest by having it move. A special case of this is time-lapse photography. In this case, we use a picture sequence to create a perception of motion, which is otherwise imperceptible (e.g., "watching the grass grow") because the rate of change is normally too slow for our motion detectors to respond to.

In sum, pictures can direct our attention to an intended point of interest, which is much more difficult to achieve in our normal visual environment — usually the best we can do is to point at something and say "Look at that!" But pictures allow us far greater subtlety and precision in guiding a person's attention, and they can do so across vast distances of time and space. Therefore, we can consider the ability to direct visual attention to be a special feature of pictures, which from a communication standpoint, represents expanded possibilities for the sender of information.

Conversely, a second attentional benefit of pictures is related more to expanding the possibilities for the receiver of information. Specifically, because pictures can freeze details in time and space, this allows us to leisurely explore a scene and take in all its particulars. Compare this situation to what we see while walking down a bustling city street. We are met with a million details many of which are changing on a moment-by-moment basis. Our attentional limitations are far exceeded and we are able to take in only a very few details at a given place and time. Much of what is going on around us is completely missed and becomes "background." It is only by taking or making a picture of the scene that we can freeze the details in it. This

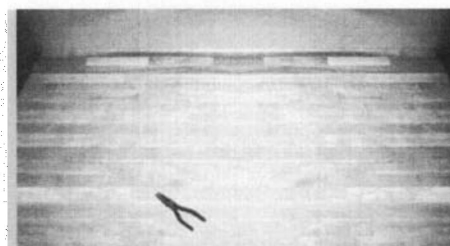


FIGURE 3F

[25] Newsome, W. T. 1997. Deciding about Motion — Linking perception to action. *Journal of Comparative Physiology A-Sensory Neural & B>avioral Physiology*, 181:1, 5-12.

[26] Ewert, J.P. 1974. The Neural Basis of Visually Guided B>avior. *Scientific American*, 230:3, 34-42.

extends our ability to explore the image in greater depth, giving us time to send our eyes to more points of interest, which in turn enables us to view them in high resolution at the center of vision.

This benefit of pictures can be explained in terms of two basic findings from research on visual attention. First, our visual attention is spatially and temporally limited. It takes time to scan a scene and attend to all of its details, which are spread over an extended region of space. In this context, it is worth considering figures 6a and 6b again. Unless an item pops out and grabs our attention almost instantaneously as in figure 6a, we may only notice it by scanning through the image until we finally hit upon it as in figure 6b, and this takes time and effort.[21] This explains the idea b>ind the expression "like trying to find a needle in a haystack." A needle looks a lot like a piece of hay (both are long, thin, and straight), and thus finding it takes a long time.

Second, paying attention to something is necessary in order to recognize what and where it is.[22] Most people are unaware of this fact, because we perceive the 'gist' of a scene, for example whether it is an office or a forest, almost 'instantaneously,' i.e., within a single eye fixation (or about 100-200 ms).[23] Thus, we have the impression of seeing everything in a scene at any given moment in time. In fact, however, recent research indicates that what we are perceiving is limited to 1) the specific item we are paying attention to at that moment and 2) the general spatial layout and category of the scene.[24]

Taken together, these two aspects of visual attention result in our generally being unaware of the specific identities and locations of items in a scene until we take the time to eventually come across them. For example, take a look again at the picture shown in figure 2a. One can very quickly recognize that it shows a street market scene. But even though I've looked at this picture numerous times, only recently did I first notice the third of three small children standing between the two central women. Likewise, few people seem to notice the crow in the picture. Thus, it is through freezing the details in a picture in the temporal and spatial dimensions that we are afforded the opportunity to leisurely attend to the various parts of an image.

This, in large part, explains the difficulty of detecting a change between two otherwise identical images that are separated in time and space. This is best exemplified by "spot the difference" picture pairs, as shown in figure 7. When we look at either picture in a spot the difference pair, we see what looks like the same picture because both pictures have the same 'gist' (or overall meaning). In this case, both versions show an office. If only the identity and/or location of individual items differ between the two versions, without changing the gist, these differences will not be readily apparent until we carefully pay attention to and compare each particular item in each picture.[25]

Notice, of course, that this kind of picture change violates a real world constraint — namely that changes to a real world scene are generally accompanied by perceived motion, with the exception of grass growing, etc. If the changing pictures were shown one after the other, as in an animated image sequence, we would perceive motion due to the change and this would immediately draw our attention to the change. Thus, "spot the difference" puzzles can be considered to be 'defective picture pairs' because they inhibit our ability to detect a change, by blocking the attentional guidance property of perceived motion. Nevertheless, because their details are frozen in time and space, we have the opportunity to detect the change if we take the time to look carefully. Thus, to the extent that the attentional guidance aspect of pictures is decreased, we must depend on the expanded possibilities for attentional exploration in pictures (because details are frozen) to compensate for it.

Memory Management

Because pictures can freeze details, they also expand the potential for the person receiving visual information to more deeply contemplate it. This is a by-product of the fact that through storing visual information in an unchanging form, pictures help us overcome our visual memory limitations, both long term and short term. The use of pictures as long-term storage of visual images is probably their best-known function. People constantly take photographs as records of their experience. Without photographs, we often have difficulty remembering even those images that should be the most memorable, for example what our parents looked like when we were children, or what our children looked like when they were young. This explains the expressions of surprise we often make when we look at our old photographs — "I can't believe I looked like that!"

Most of us, however, are less aware of the way pictures serve as aids for our short-term memory as well. Earlier I suggested that you try the visual short-term memory test in figure 3 as a way of illustrating how little we remember of what we see around us on a moment-by-moment basis. But the importance of short-term memory and its varying uses can also be surprising to discover. For example, language compre-

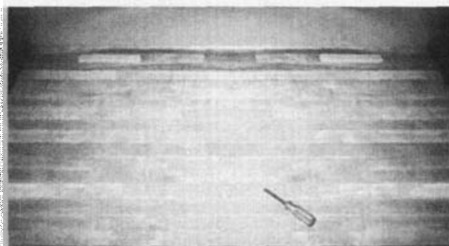


FIGURE 3g

[21] Again, see Treisman and Gelade, 1980 and Wolfe, 1994.

[22] Schneider, W.X. 1995. VAM: A Neuro-cognitive Model for Visual Attention Control of Segmentation, Object Recognition and Space-based Motor Action. *Visual Cognition*, 2:2/3, 331-376.

[23] Biederman, I., R. Mezzanotte and J. Rabinowitz. 1982. Scene Perception: Detecting and Judging Objects Undergoing Relational Violations. *Cognitive Psychology*, 14, 143-177. Boyce, S. and A. Pollatsfi. 1992. Identification of Objects in Scenes: The Role of Scene Background in Object Naming. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 18, 531-543.

[24] See C'Regan, 1992

[25] Simons, D. and D. Levin. 1997. Change Blindness. *Trends in Cognitive Sciences*, 1, 261-267. McConkie, G.W. and L. C. Loschky. (In press.) Change Blindness. *Encyclopedia of Cognitive Science*. MacMillan/Nature Publishing Group.

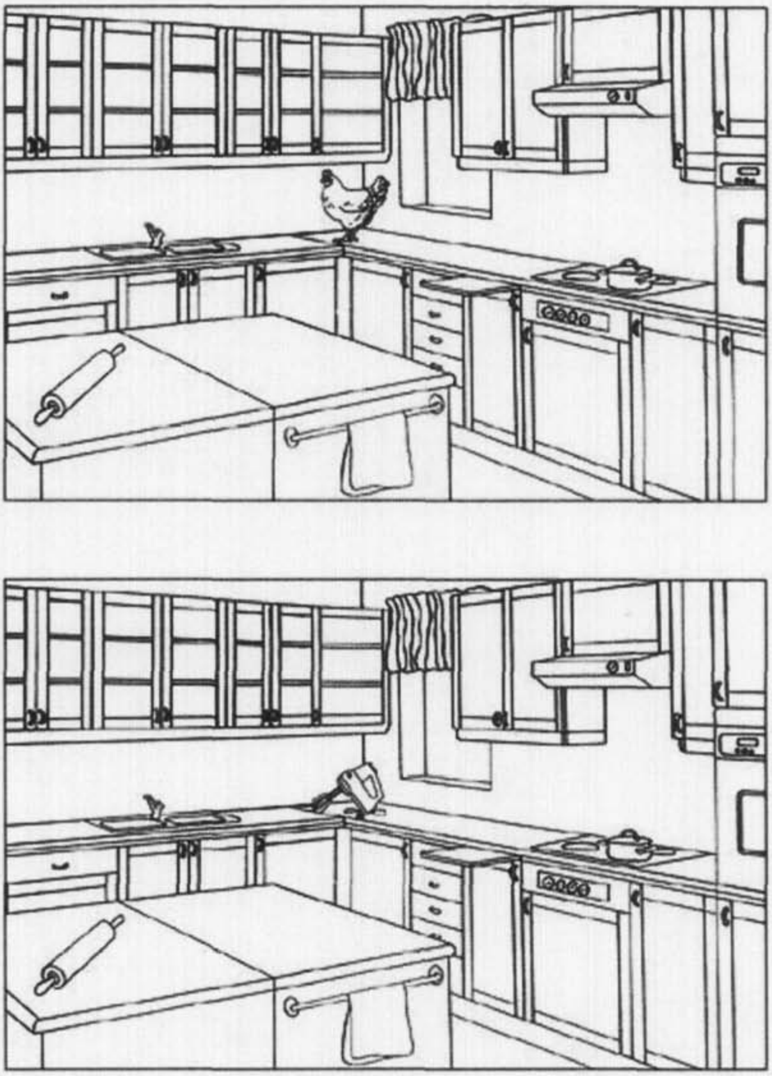


FIGURE 8

Two Versions of a Picture Differing in Their Typicality.

The top version of the picture contains an atypical item for a normal American kitchen scene, a live chicken. In contrast, the corresponding item in the bottom version of the picture, a mixer, is very typical for a normal American kitchen scene. When shown each of these versions of the picture, American college student viewers were much more likely to look long and repeatedly at the chicken than they were to look at mixer. This pattern of eye movements can be explained in terms of the role of working memory in visual comprehension. (Background scene and object images courtesy of John M. Henderson's Visual Cognition Laboratory, at Michigan State University.)

hension is intimately connected with the use of short-term or 'working memory.'[26] Most people who have studied until late at night have probably experienced the effects of fatigue on short-term memory and how this affects reading compr>en-sion. You can be reading along, arrive at the end of a sentence and suddenly realize that you can't remember what the beginning of the sentence was! Clearly, you cannot compr>end how a sentence fits in with the rest of the text if you cannot remember what it said. Luckily, because text does not change from moment to moment, the simple solution to such a problem is to refresh your memory by sending your eyes back to the beginning of the sentence and read it again. The same is true, of course, for other sorts of visual representations, such as pictures. Suppose, you are leafing through a popular magazine filled with pictures. You may spend a second or less looking at each picture as you "page surf" through the magazine. At some point, however, you might be tempted to go back to the last picture you had passed and examine it more closely. In this case, the physical picture on the page serves to refresh your visual memory. Obviously, if you had perfect memory there would be no need to look at the picture again — every detail would be etched in your mind's eye.[27] Thought of in this way, the short-term memory test in figure 3 can, as a whole, be considered another sort of 'defective picture.' This is because, by presenting all of the objects in separate images, it becomes more difficult to look back at objects you looked at before, and thus taxes your visual short-term memory (which makes it ideal for testing the limits of memory).

FIGURE 3h

[26] A classic paper on this topic follows. Just, M.A. and P.A. Carpenter. 1992. A Capacity Theory of Compr>en-sion: Individual Differences in Working Memory. *Psychological Reviea*. 99:1 122-149.

[27] Frequent claims to the contrary notwithstanding, there is only the scarcest evidence of true "photographic memory." The best test devised so far is the ability to fuse two random dot stereographic images (like the "Magic Eye" 3D images), while looking at one image with one eye and remembering the other image as previously viewed with the other eye, in order to see a complete image. So far, only a single person (under normal conditions) has passed the test. Stromeyer, C. and J. Psotka. 1970. The Detailed Structure of Eidetic Images. *Nature*, 225, 346-349. This is despite the fact that John Merritt published articles in many popular magazines along with a self-test asking readers to write-in if they passed. Out of the potentially vast number of readers of his articles (perhaps a million) thirty wrote him to report having been able to pass the self-test. However, none of them was able to pass a similar test at his lab. Thus, the proportion of people having this ability is probably less than one in a million. Merritt, John. 1979. None in a Million: Results of Mass Screening for Eidetic Ability Using Objective Tests Published in Newspapers and Magazines. *B>avioral & Brain Sciences*. 2:4. 612. Haber, R.N. 1979. Twenty Years of Haunting Eidetic Imagery: Where's the Ghost? *B>avioral & Brain Sciences*. 2:4, 583-629.

Importantly, visual short-term memory appears to be tightly connected with what I will call "visual compr>ension." Research has shown that when we look at a picture that is difficult to make sense of, we send our eyes repeatedly to the incompr>ensible parts. For example, in one study, viewers were shown, among other things, a picture of a kitchen scene. If, as shown in the top of figure 8, it contained a chicken standing on the counter, viewers were likely to look at the chicken longer and repeatedly. This is presumably because the American college student viewers in the study were not accustomed to seeing live chickens in contemporary American kitchen scenes. However, when the same scene was shown with a mixer in the same location on the counter, as in the bottom of figure 8, the viewers were far less likely to look repeatedly or as long at the mixer.[28] The skeptical reader might explain this simply by saying that, obviously, the chicken is odd in the kitchen context and the mixer is not. But the interesting question is, why should people look long and repeatedly at something that seems strange? First of all, longer looking times (i.e., fixation or gaze durations) in reading commonly indicate compr>ension difficulties, and the same is true for looking back at a word that was previously fixated.[29] More specifically, I would argue that in scene viewing, we look at odd things repeatedly because each time we are 'asking a different question' of the thing we are looking at. For example, "Is that a chicken??" "Is this a farm kitchen??" "Did this chicken just fly in the window?" Such 'questions,' however, must be occurring at a pre-verbal level because each one is accompanied by an eye movement to a relevant part of the image (e.g., the chicken or window), with multiple eye movements occurring each second — much faster than one could actually articulate such a question.[30] Importantly for our discussion, it is reasonable to suspect that this process of pictorial compr>ension is related to visual short-term memory or more specifically visual 'working memory.'[31] Thus, in reading, the need to look back at a previously fixated word, for example the referent of a pronoun (e.g., "Mrs. Jones.... She..."), can be explained in terms of a compr>ension difficulty specifically related to working memory limitations (i.e., forgetting who "she" is).[32] More generally, in the process of compr>ension, when we need further information about an item we looked at before, we can either refer to our working memory of it, or send our eyes back to it and get the information 'brand new.'

In fact, if we need to, most of us are capable of relying more on our visual working memory and reducing the number of times we look back at a particular object we are thinking about.[33] But we find it much easier to look at the relevant picture areas repeatedly, and there is usually little 'cost' to making more eye movements. This is much the same as when we reread a complex sentence several times in order to grasp its meaning. If we could remember every word from the first time we read the sentence, there would be no need to reread it as we attempted to understand it — we could simply review the sentence in our head. But doing so would require that we hold the entire sentence in memory while we also attempted to analyze it. In order to reduce the burden on our memory, we simply refresh our memory for the words we are concerned with by looking at them again. Thus, another important function of pictures is to allow us to repeatedly refresh our visual working memory in order to make sense of what we see. By relying on information in the picture, rather than our memory of it, we can free up our memory resources to use for other processes, such as relating information from the object we are looking at to other information in working memory in order to compr>end the picture. Once we have abstracted the information we need from the picture, we are less likely to look at it again.[34] Of course we do exactly the same thing while looking at objects in the real world. But a special characteristic of pictures is that they allow us to repeatedly refresh our short-term memories of objects without their having changed or disappeared in the mean time. Of course, we cannot make use of this attribute of pictures if they are dynamically changing, for example on television, unless we have the option of reviewing them, as in video or digital media.

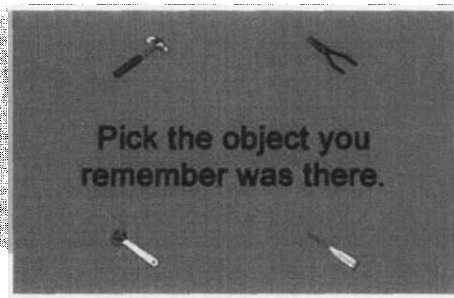


FIGURE 3i

[28] Henderson, J.M., P.A. Wefis and A. Hollingworth. 1999. The Effects of Semantic Consistency on Eye Movements During Complex Scene Viewing. *Journal of Experimental Psychology: Human Perception & Performance*. 25:1, 210-228. Figure 8 (top and bottom) come from a set of scene and object images publicly available on Dr. Henderson's laboratory website (http://eyelab.msu.edu/VisualCognition/).

[29] See, for example, Rayner, K., G.E. Raney and A. Pollatsfi. 1995. Eye Movements and Discourse Processing. In O'Brien, R.F. Jr. and E. J. O'Brien, editors. *Sources of Coherence in Reading* (Vol. xiv, 9-35). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Inc.

[30] For research looking at the relationship between the eye movements people make and their speech while they are describing pictures they are looking at see the following. Griffin, Z.M. and K. Bock. 2000. What the Eyes Say About Speaking. *Psychological Science*. 11:4, 274-279.

[31] Working memory is thought of as including both the short-term storage of information and the concurrent manipulation of that or other information for some purpose, for example compr>ension, reasoning, etc. See Logie, R. H. 1996. The Seven Ages of Working Memory. In Richardson,J., R.W. Engle, L. Hasher, R.H. Logie, E.R. Stoltzfus and R.T. Zacks, editors. *Working Memory and Human Cognition*. New York: Oxford University Press, 31-65.

[32] Rayner et al., 1995.

[33] For more evidence and a detailed discussion of this see the following. Ballard, D.H., M.M. Hayhoe, F. Li and S.D. Whit>ead. 1992. Hand-eye Coordination During Sequential Tasks. *Philosophical Transactions of the Royal Society of London*. B(337). 331-339.

[34] For example, people are less likely to look at an object in a picture that they are describing if they refer to it using a pronoun (e.g., he, she, it) than if they refer to it using a noun (e.g., the tall man). van der Meulen, F.F., A.S. Meyer and W.J.M. Levelt. 2001. Eye Movements During the Production of Nouns and Pronouns. *Memory & Cognition*, 29:3, 512-521. Referring to an entity by using a pronoun suggests a level of abstraction in our mental representiation of it.

Conclusion

Pictures enable us to lead richer visual lives by expanding our options in dealing with our limited attentional and short-term memory faculties. Pictures can be an extremely effective means of directing a person's attention to things they might not have noticed. And because pictures can freeze visual information in an unchanging state, they allow a viewer more time to carefully explore (i.e., pay attention to) visual details that otherwise might have disappeared in the continuous flux of the visual world. Furthermore, because pictures can store information in a stable state, our leaky short-term memories don't have to struggle to hold onto their contents. This then frees up mental resources for contemplating (i.e., more deeply processing) information in the image. Whenever we need to go back to get further information from a picture, in order to think about it in relation to other information, it is still there, provided the pictorial information is stable or dynamically retrievable.

An interesting related question is whether our habits in looking at pictures can alter the ways in which we look at the world. Watching television is commonly argued to result in shortening viewers' attention spans, since the images change at a more rapid pace than much of the non-televised world (see a detailed discussion of this in Mathew McClain's article in this issue). Whether this is true is an empirical question, but it certainly makes sense within the framework I have presented here. It has also been frequently said that the most important thing in becoming a visual artist is to "learn how to see." Within the context of this article, I would argue that a critical factor in such "sight" is how much and what we pay attention to in our visual surroundings, and how deeply we process that information. And finally, although I have been speaking about visual attention and memory as two separate things, they are tightly interconnected: in general, we remember only those things that we pay attention to.[33] Thus, although we are all faced with basic attentional and memory limitations, pictures can help us by guiding our attention, allowing us the luxury of attending to details and enabling us to more deeply process the contents of the visual world.

Some Things That Pictures are Good For: An Information Processing Perspective

Lester Loschky
University of Illinois at Urbana-Champaign
Human Perception and Performance Group
Beckman Institute for Advanced Science and Technology
405 North Mathews
Urbana, IL 61801
Visible Language 35.3
Loschky, 244-265
loschky@uiuc.edu

[35] Irwin, D.E. and R.D. Gordon. 1998. Eye Movements, Attention and Trans-saccadic Memory. *Visual Cognition*, 5:1, 127-155. Neisser, U. and R. Becklen. 1975. Selective Looking: Attending to Visually Specified Events. *Cognitive Psychology*, 7:4, 480-494.