JVL

Volume 21 Number 1 Spring 2001

Official Journal of the International Visual Literacy Association, Inc.
ISSN: 1051–144X © 2001
••Indexed in CIJE beginning with Vol. 11, No. 1••

CONTENTS

IVLA Website. Online abstracts of JVL are available at **http://www.ivla.org**

Journal of Visual Literacy, Spring 2001
Volume 21, Number 1, 1-14

Depth Perception in Media Design: From Sensory Psychology Cues to Interactive Tools

Kevin Oliver, Ph.D.
Virginia Tech
Blacksburg, Virginia, U.S.A.

Abstract

Considerable research in sensory psychology has demonstrated the ability to generate depth perception in media by manipulating the physical presentation, relative motion, or pictorial qualities of objects. These research results can enable educational developers to create more realistic educational materials when depth perception is key to comprehension. Developers may also utilize research results and modern software tools to create exploratory environments for learners to alter, position, or interact with spatial elements in a scene.

Importance of Depth Cues in Education

Depth perception is important to avoid potential misunderstandings by students learning to gauge the relative physical distance between objects in a defined space. Across numerous disciplines, it is important to realistically represent the relative position of objects to one another: internal organs for those studying medicine; molecules for those studying chemistry; planes for those studying air traffic control; atoms for those studying attractional forces; plants for those studying landscape design; buildings for those studying urban planning; floors, walls, and ceilings for those studying architectural spaces; earthquake waves and differential ground movements for those studying seismology.

Sensory psychology research indicates three types of cues help to promote depth perception: physical, motion-based, and pictorial. Media designers can apply the research evidence to create more realistic teaching materials. Physical cues are most appropriate for those developers with access to high-end virtual reality systems. Motion-based cues can be employed by those with access to animation or video generating software. Pictorial depth cues can be readily employed by any media designer struggling to create the illusion of three-dimensions on standard, two-dimensional computer screens. Descriptions of specific depth cues are now provided with supporting research evidence.

Physical Depth Cues

Media designers with access to high-end virtual reality systems can create the illusion of depth by manipulating the perspective or clarity of images on the eyes.

Disparity, Stereopsis, or Binocular Parallax

Binocular disparity is the physical separation of the eyes, allowing an individual to see the world from two slightly different perspectives. This physical difference in retinal stimulation creates a perception of depth referred to as stereopsis and sometimes binocular parallax. Disparity represents a binocular depth cue as vision with two eyes is required for the cue to take effect. Complex virtual reality systems support stereopsis by rendering two distinct images that are slightly displaced. Each image is fed to one eye. These systems utilize head-mounted display units with two separate screens receiving the two different images. Stereoscopic images can be generated through two synchronous computers or using a single computer with dual graphics hardware (Pimental & Teixeira, 1993). Single computer systems are more affordable, but considerably slower in rendering speed. Before computers, simple instruments known as stereoscopes allowed images to be seen in 3-D by presenting two slightly displaced photographs to the eye (National Stereoscopic Association, 2000).

Accommodation, Convergence

At distances closer than five or ten feet, the perception of depth can also be cued by accommodating or convergent eye motions (Goldstein, 1996). Accommodation is a physical process whereby the ciliary muscles in the eye tighten and loosen to increase and decrease the curvature of the eye's lens. Accommodation keeps near or far objects in focus by either fattening or reducing the width of the eye's lens. Light from the target image is bent appropriately to fall directly onto the retina. Convergence is the physical movement of our eyes inward toward one another and outward away from one another to facilitate focusing. Both accommodation and convergence provide depth information to the brain regarding the distance between an observer and a target object.

Motion-Based Depth Cues

Motion has also been shown to provide depth information through both stereoscopic and standard, monoscopic computer displays. Media designers may manipulate motion through animation or video to provide depth cues. Motion parallax and the kinetic depth effect are two cues created by underlying motion processing systems: the image-retina/eye-head motion recognition system, and the short-range/long-range motion processing system

Motion Parallax

Motion parallax refers to the different speeds in which objects move relative to one another. When we view far objects related to near objects, the far objects appear to move more slowly than the near objects and this cue facilitates the perception of depth. Motion parallax can be experienced by motor action (i.e., the eye-head system) in which we move our eyes or head from side to side while physically walking through some room or virtual environment. Motion parallax can also be experienced when we are stationary (i.e., the image-retina system) and some stimulus motion sweeps across our retina without any action on our part (e.g., an animated dog runs across our field of view on a computer monitor, we view the countryside in motion as we stare out the window of a car or train).

Rivest, Ono, and Saida (1989) conducted experiments to determine the role of motion parallax in facilitating depth constancy. Depth constancy refers to our perceptual ability to see objects as located in approximately the same depth no matter what distance we view these objects from. Motion parallax information changes at varying viewing distances, thus the farther away we view objects, the less motion information is available to serve as a cue to depth. Other cues, then, must be used in conjunction with motion parallax information to regulate depth constancy when motion parallax information is unavailable or lessened. Whether these "other" cues were absolute (i.e., original convergence information) or apparent (i.e., "eye rotation relative to head movement during fixation") was the subject of a number of experiments (Rivest et al., 1989, p. 407). Results show absolute cues do not regulate depth constancy with motion parallax but apparent cues are involved. Results suggest the visual system uses apparent, processed information rather than absolute, original information to calibrate motion parallax and binocular cues are not necessary for depth constancy since apparent information can be derived monocularly (Rivest et al., 1989).

In other experiments, Cavanagh, Saida, and Rivest (1995) discovered color also contributes to the perception of depth in motion parallax conditions. Using heterochromatic (i.e., red/green) and monochromatic (i.e., yellow) gratings in motion, the perception of depth was achieved for both at high luminance contrasts but was decreased for both as luminance decreased (Cavanagh *et al.,* 1995). Most significantly, Cavanagh *et al.* (1995) found observers were more likely to see depth from the heterochromatic gratings than from the monochromatic gratings. This finding suggests the "presence of color adds signal to depth from parallax" (Cavanagh *et al.,* 1995, p. 1872).

Kinetic Depth Effect

The kinetic depth effect refers to the perception of objects as two-dimensional or flat when stationary, but three-dimensional when in motion. A series of dots that form the image of a flat sphere may suddenly appear to take 3-D shape when set to motion in similar directions (Mather, 2000). Mather (1989, p. 184) notes the kinetic depth effect "must be mediated by a visual process that signals movement in the image," and argues for two motion processing systems – short-range and long-range. The short-range process allows us to perceive motion over restricted ranges, whereas the long-range process allows us to perceive motion over wider ranges. While the long-range process is thought to be the key mediator of kinetic depth, since it may take longer periods of time to compute depth from moving objects, Mather (1989) conducted experiments to determine whether the short-range process might also play a role. Results indicated subjects were more likely to see depth from a kinetic stimulus when the rotation angle (i.e., displacement between frames) was small and when inter-frame intervals (IFI) were short (i.e., fast). These parameters are closely aligned with the short-range apparent motion system, supporting Mather's hypothesis that early motion processes also mediate the kinetic depth effect.

Day (1989) also experimented with the kinetic depth effect, utilizing a moving stimulus that was viewed through a narrow aperture. Results show shadows alone are perceived in only two dimensions, but that shadows in motion allow for the perception of object depth. This perception of depth occurred through the narrow aperture both when the axis of rotation was centered on the aperture and when the axis was moved to the left and right of the aperture. These results suggest the visual system is highly sensitive to motion, even when the stimulus is partly occluded.

Pictorial Depth Cues

Pictorial depth cues are often referred to as monocular depth cues, because they are just as effective in creating the perception of depth with only one eye or monoscopic media as they are with two eyes or high-end, stereoscopic systems. Pictorial depth cues were first applied by artists to create the perception of depth. These cues allow media designers limited to low-end software tools to create depth perception.

Interposition or Occlusion

Interposition refers to the partial overlap of a more distant object by a closer object. Interposition is also referred to as occlusion. This cue provides information about relative depth but does not give a strong impression of depth or provide any information about an object's distance (Krantz, 1999; Goldstein, 1996) (see Figure 1).

Marshall, Burbeck, Ariely, Rolland, and Martin (1996) conducted experiments to determine whether a sharply focused or a blurred occlusion boundary between a sharp and blurred surface served as a relative depth cue. Despite variability in strength of effect, "observers consistently used the blur of the boundary as a cue to relative depth" (Marshall et al., 1996, p. 681).

Despite the fact that interposition is seen by most as a monocular depth cue (Goldstein, 1996), Nakayama and Shimojo (1990, p. 1811) note that "distant surfaces are occluded by nearer surfaces to different extents in the

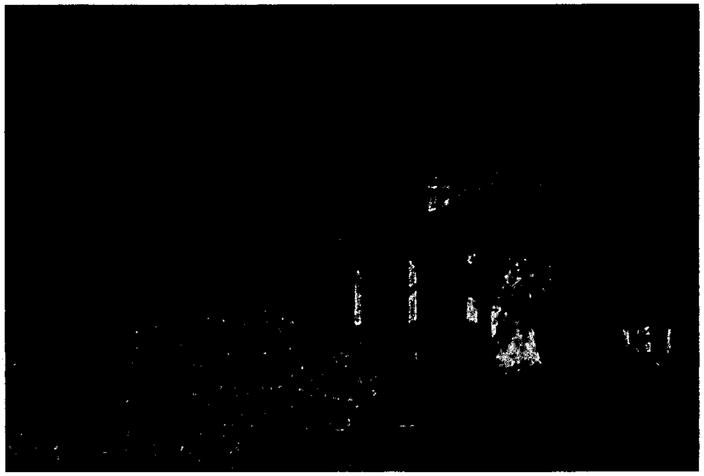

Figure 1. Occlusion is illustrated by the car behind the bush and the house behind the trees.

two eyes." In experiments, Nakayama and Shimojo (1990) discovered the resulting unpaired image points facilitate the perception of depth, contour, and surface. Since unpaired image points are clearly related to stereopsis and require binocular viewing, the contribution of occlusion to depth perception can not be realized by standard, monocular computer displays.

Linear Perspective

Linear perspective provides an illusion of depth through the use of lines converging onto a vanishing point. One of the most cited examples of the linear perspective is our perception of parallel railroad tracks converging in the distance (Pimentel & Teixeira, 1993; Goldstein, 1996). Architects and artists use the linear perspective to depict realistic scenes. To illustrate a hallway, for example, one could draw a large square in the foreground, a small square in the distance, then connect the four corners of each square with converging lines. The resulting "hallway" is seen in depth on its two-dimensional surface. The receding, narrowing sidewalk in Figure 2 illustrates the linear perspective.

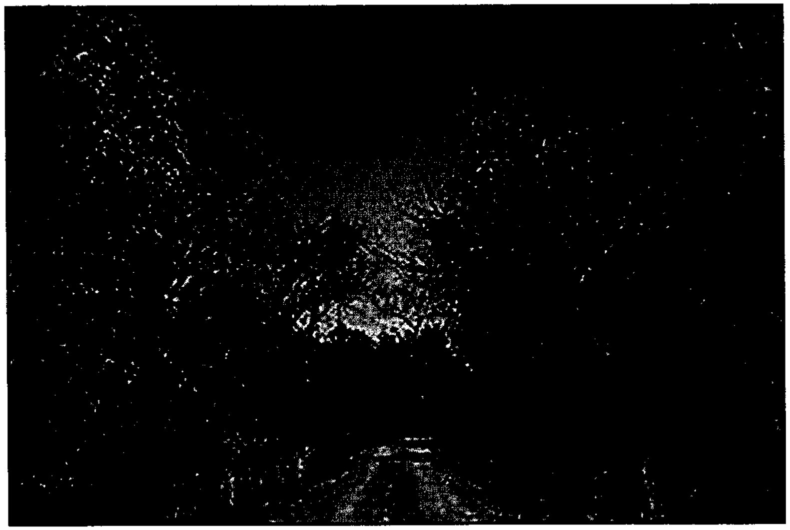

Figure 2. The linear perspective is illustrated by converging edges on the road.

Relative Height and Size

Relative height is based on the horizon line in a photograph, image, or real-world scene. As illustrated in Figure 3, objects that lie nearer the horizon line (e.g., clouds, trees) will appear more distant than objects that lie farther away from the horizon line. Relative height often works in conjunction with the relative size depth cue. According to the relative size depth cue, objects that appear smaller in a field of view will generally appear more distant than objects that appear larger. Although we tend to perceive smaller objects as lying farther away than larger objects, we don't necessarily perceive smaller objects as being smaller than larger objects. This phenomena is related to size constancy and our perceptual ability to correctly estimate an object's size regardless of distance. A mechanism known as "size-distance scaling" takes into account both the size of an object on our retina as well as the distance of that object from us in determining the object's true size (Goldstein, 1996, p. 257). This mechanism describes how the visual system takes into account both relative height (i.e., distance information) and relative size depth information in correctly estimating an object's size.

Figure 3. Trees and clouds at a relative height farther from the horizon line appear closer than trees and clouds near the horizon line.

Texture Gradient or Surface Composition

Another depth cue related to relative size and linear perspective is texture gradient or surface composition. Texture gradient refers to surface properties (e.g., borders between bricks, cobblestones, floor tiles) and their tendency to get smaller, finer, and smoother as they recede into the distance. Floor tiles in an image, for example, appear packed more closely together as the surface recedes into the distance. According to Pimentel and Teixeira (1993, p. 116) "texture mapping can dramatically improve the sensation of depth" in media. These cues provide the eye with valuable details for resolving relationships. Norman, Lappin, and Zucker (1989, p. 252) concur, "...stereopsis is sensitive to the differential structure of surfaces, responding to gradients, curvature, and/or discontinuities of the disparity field."

Atmospheric or Aerial Perspective, Contrast

Objects in the foreground of a scene often look sharp and clear in comparison to objects in the background that looked washed out and blurred (refer to Figure 4). This phenomena has been described as the aerial perspective and is a cue for the perception of depth. The aerial perspective is also referred to as the atmospheric perspective, because particles of dust and water in the atmosphere are partly responsible for the blurring of distant objects (Goldstein, 1996). In addition to blur, distant objects may also appear more blue than objects in the foreground (Pimentel and Teixeira, 1993). This feature is caused by the scattering of blue light wavelengths which are easily scattered by atmospheric particles (Krantz, 1999). As light travels from distant objects, those wavelengths will be scattered more than light travelling from nearby objects.

One "visual consequence" of light from distant objects being scattered more is "a reduction in contrast of an object with distance" (O'Shea, Blackburn, & Ono, 1994, p. 1595). O'Shea *et al.* (1994) tested the value of contrast as a depth cue, and not surprisingly, found that subjects reported areas of low contrast were farther away than areas of high contrast.

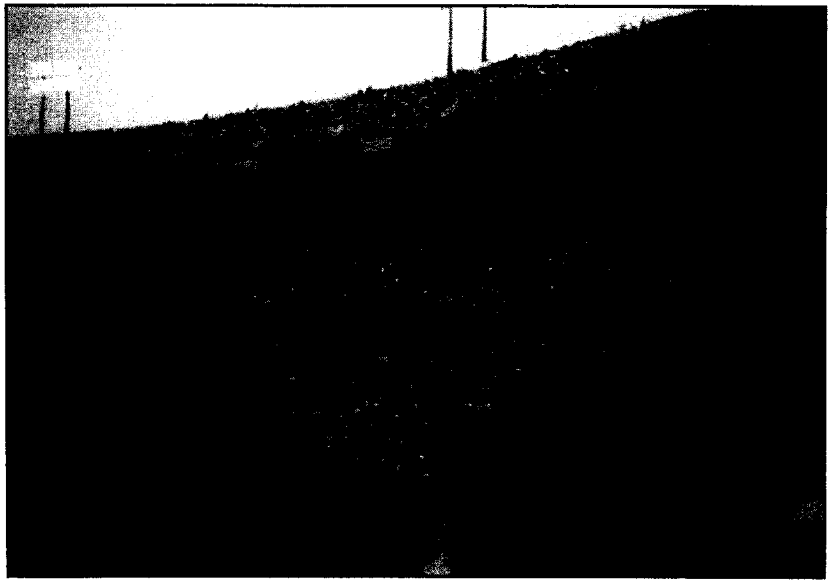

Figure 4. Texture gradient is illustrated by the ever-decreasing detail in persons and objects as distance increases.

Shading, Shadows

Shading is a powerful cue to depth perception most often related to the shadows an object casts when illuminated by a light source. According to DeHaan, Erens, and Noest (1995, p. 2985), "shading is fully specified by local surface reflectance properties and the angles between a surface element and the directions of light sources and the viewer." Certain accepted rules for shadows exist, allowing us to infer which objects are at depth: light sources typically illuminate and cause shadows from above, an object's shadow side is away from the light source, objects close to light sources cast shadows on more distant objects, and objects in shadow are farther from light sources than objects not in shadow (Krantz, 1999, chap. 6). Kleffner and Ramachandram (1992, p. 18) endorse these general rules in noting, "the extraction of shape from shading information incorporates at least two 'assumptions' or constraints--first, that there is a single light source illuminating the whole scene, and second, that the light is shining from above...." Shadows cast by objects and shadows cast onto objects are both capable of producing stereoscopic depth perception (Puerta, Guttmann, and O'Day, 1989).

Experimental evidence indicates the "extraction of shape from shading" occurs early in visual processing and that "neural elements" or "tokens" may exist to serve as a basis for "perceptual grouping and segregation" (Kleffner & Ramachandram, 1992, p. 18). Since shading is thought to cue shape and depth perception when the viewer is aware of light sources, viewing direction, knowledge of object shape, etc., DeHaan et al. (1995) tested if "pure shading" would cue shape and depth in the absence of these other cues. Results indicated shape can be determined from pure shading information alone, despite biases stemming from the human predisposition for familiar cues (i.e., bias for right or overhead illumination) (DeHaan *et al.,* 1995).

Color

Color can also serve as a depth cue under certain conditions. Troscianko, Montagnon, LeClerc, Malbert, and Chanteau (1991) experimented with texture gradient and no texture gradient stimuli and found that the inclusion of color gradient information increased the perception of depth across all texture conditions. Using color gradients alone (e.g., red at bottom, grey at top), however, did not contribute to the perception of depth unless texture gradient information was also available. The findings suggest "...while colour can code depth, its contribution is contingent upon the presence of texture cues" (Torscianko et al., 1991, p. 1928).

Faubert (1994, p. 1165) notes "the most compelling example of seeing depth with coloured stimuli is evident in conditions of chromostereopsis." Chromostereopsis refers to the ability to see two colors (i.e., red-blue or red-green) in the same depth plane as lying in separate planes. This color depth cue is caused by "chromatic aberration," or a difference in refraction causing some light rays to converge on the eye more quickly and to fall on non-corresponding regions of the two eyes. This particular use of color as a visual depth cue is effective only in binocular conditions through stereoptical systems.

Image Blur

Already noted was Marshall et al.'s (1996) experiments where observers used occlusion boundary blur to help with depth perception. In addition to these findings, Mather (1996) notes photography and videography often generate different regions blurred in differing amounts due to limitations in depth of focus. Experimental data indicates that the visual system uses blur between borders of sharp and blurred regions to order the depth of different regions (Mather, 1996). Since limitations exist in the human eye's depth focus, depth focus limitations in reproduced images may produce a natural blur that the eye can already work with in perceiving depth. Selectively blurring images, then, across different regions in depth may contribute to the perception of depth in computer-generated images and virtually developed environments.

Luminance

A final pictorial depth cue is luminance. Without luminance, experiments have shown texture gradients and colored gradients to be less effective in cueing depth, thus luminance information must help to signal depth in these otherwise effective depth cues (Troscianko *et al.,* 1991). Further research evidence suggests three crowded lines may be discriminated in depth when they have different luminances (Kumar & Glaser, 1995). Luminance makes no difference in the perception of depth when lines are less crowded. In crowded stimulus conditions, then, differing luminance values can contribute to the perception of depth.

Summary and Application

The illusion of depth promotes clarity of meaning for numerous disciplines in which relationships between objects in space must be defined, manipulated, or judged. Stereopsis is the physical mechanism by which we perceive depth in the real world. Those media designers with access to expensive stereoscopic systems are able to simulate stereoptical viewing and the illusion of depth. Until such virtual representations are commonplace and accessible to more individuals, however, most media designers will continue to rely upon pictorial and/or motion-based depth cues to provide depth information through more standard, two-dimensional computer displays.

Pictorial depth cues are perhaps the most common and easily employed treatment, manipulated on graphics and photographs by image editing software. Freehand software (Macromedia, 2000), for example, allows a media designer to generate a new graphic, employing depth cues such as the linear perspective with easily manipulated perspective grids. Photoshop (Adobe, 2000) readily allows a media designer to occlude objects, alter height and size, add filter gradients to parts or backgrounds, blur more distant items, shade objects, combine color gradient with texture information, and vary luminance values. The ability to manipulate multiple pictorial elements is important since some pictorial depth cues work in combination (e.g., linear perspective, relative size, and texture gradient; relative height and relative size).

Not only can two or more pictorial depth cues work in combination, but the mind is capable of processing pictorial, motion-based, and physical depth information simultaneously. Research studies suggest the existence of visual pathways that take apart and process different visual information (e.g., motion, color, shape). This information is recombined into a "whole" at a point in time when we perceive depth. Electrical signals leaving our retina are processed along several pathways, and that these pathways carry different information: the blob pathway processes color information, the magno pathway processes motion and stereoscopic depth information rapidly, and the parvo pathway processes detailed images and shapes more slowly (Livingstone, 1988). Experiments by Ingling and Grigsby (1990) indicate components of the color blind parvo channel also can perceive motion-based depth.

Motion-based depth cues are somewhat more difficult to employ, but numerous hardware and software tools are available to design and produce media with depth effects. To produce video with depth information, considerable pre-planning should be undertaken to determine how the preceding pictorial cues can be combined with the motion-based cues of parallax or kinetic depth for the shoot. For example, a black field of white stars in time-lapse motion against a stationary moon with a blur boundary might promote the illusion of depth through both pictorial and parallax cueing. Premiere (Adobe, 2000) is one software tool allowing for the digitization of raw video footage and the editing of scene sequences. Three-dimensional modeling and animation software also can effectively combine pictorial and motion-based depth information. Grunin (1996) provides an overview of eight three-dimensional imaging and animation software programs, including Ray Dream Studio (Fractal Design Corporation, 2000) and LightWave 3D (Newtek, 2000). Increasingly, animations can be created for and viewed on the Internet. Flash software (Macromedia, 2000) is useful for designing vector-based animations for the Internet which can include interactive components for manipulation.

The interactive factor is a critical one, since depth perception may be aided not only by pictorial and motion-based cues, but also by increased functionality in viewing. Is depth perception aided by the ability to view a scene from different angles (e.g., bottom, top, left, right), virtually "walk" into and around a scene, or even stretch or combine scene elements? Can simulated environments aid depth perception by allowing the learner to redefine variables and view resulting changes (e.g., resizing a heart to understand its fit within the chest or to view underlying structures; rapidly growing virtual plants and shrubs to determine adequate spacing from structures)? Modern software tools can be generated for students to manipulate scenes and rapidly generate alternate versions of vector-based animations and images. While promoting depth perception through cueing, developers of educational media should also study the potential benefits of flexible tool functions.

References

Adobe Systems Incorporated. (2000). [On-line]. Available: http://www.adobe.com/

Bishop, P. O. (1996). Stereoscopic depth perception and vertical disparity: Neural mechanisms. *Vision Research, 36* (13), 1969-1972.

Cavanagh, P., Saida, S., Rivest, J. (1995). The Contribution of Color to Depth Perceived from Motion Parallax. *Vision Research, 35* (13), 1871-1878.

Day, R. H. (1989). Apparent depth from progressive exposure of moving shadows: The kinetic depth effect in a narrow aperture. *Bulletin of the Psychonomic Society, 27* (4), 320-322.

DeHaan, E., Erens, R. G. F., Noest, A. J. (1995). Shape from shaded random surfaces. *Vision Research, 35* (21), 2985-3001.

Faubert, J. (1994). Seeing depth in colour: More than just what meets the eyes. *Vision Research, 34* (9), 1165-1186.

Fractal Design Corporation. (2000). [On-line]. Available: http://www.fractal.com/

Goldstein, E. B. (1996). *Sensation and perception* (4th ed.). Pacific Grove, Ca: Brooks/Cole.

Grunin, L. (1996). Ideas taking shape. *PC Magazine Online*, [On-line]. Available: http://www.zdnet.com/pcmag/features/software/1519/_open.htm

Ingling, C. R. Jr., Grigsby, S. S. (1990). Perceptual correlates of magnocellular and parvocellular channels: Seeing form and depth in afterimages. *Vision Research, 30* (6), 823-828.

Kleffner, D. A., Ramachandram, V. S. (1992). On the perception of shape from shading. *Perception and Psychophysics, 52* (1), 18-36.

Krantz, J. H. (1999). Vision and art. Sensation and Perception Tutorials [On-line], Available: http://psych.hanover.edu/Krantz/art/

Kumar, T., Glaser, D. A. (1992). Depth discrimination of a line is improved by adding other nearby lines. *Vision Research, 32* (9), 1667-1676.

Kumar, T., Glaser, D. A., (1995). Depth discrimination of a crowded line Is better when it is more luminant than the lines crowding it. *Vision Research, 35* (5), 657-666.

Livingstone, M. S. (1988). Art illusion and the visual system. *Scientific American, 258* (1), 78-85.

Macromedia. (2000). [On-line]. Available: http://www.macromedia.com/

Mansfield, J. S., Simmons, D. R. (1989). Contrast threshold for the identification of depth in bandpass-filtered stereograms. *Investigative Opthamology & Visual Science, 30* (3), 251.

Marshall, J. A., Burbeck, C. A., Ariely, D., Rolland, J. P., & Martin, K. E. (1996). Occlusion edge blur: A cue to relative visual depth. *Journal of the Optical Society of America, A* (13), 681-688.

Mather, G. (1989). Early motion processes and the kinetic depth effect. *Quarterly Journal of Experimental Psychology and Human Experimental Psychology, 41* (1), 183-198.

Mather, G. (1996). Image blur as a pictorial depth cue. *Proceedings of the Royal Society of London Series B Biological Sciences, 263* (1367), 169-172.

Mather, G. (2000). Motion perception. [On-line]. Available: http://www.biols.susx.ac.uk/Home/George_Mather/Motion/

Miller, R. J., Patterson, R. (1995). Influence of flicker on perceived size and depth. *Perception & Psychophysics, 57* (5), 604-613.

Nakayama, K., Shimojo, S. (1990). DaVinci stereopsis depth and subjective occluding contours from unpaired image points. *Vision Research, 30* (11), 1811-1826.

National Stereoscopic Association. (2000). [On-line]. Available: http://nsa-3d.org/

Newtek. (2000). [On-line]. Available: http://www.newtek.com/

Norman, J. F., Lappin, J. S., & Zucker, S. W. (1989). Stereoscopic discrimination of smooth vs. discontinuous surface structure. *Investigative Opthamology & Visual Science, 30* (3), 252.

O'Shea, R. P., Blackburn, S. G., Ono, H. (1994). Contrast as a depth cue. *Vision Research, 34* (12), 1595-1604.

Parks, T. E. (1987). On the relative frequency of depth effects in real versus illusory figures. *Perception and Psychophysics, 42* (4), 333-336.

Journal of Visual Literacy, Spring 2001
Volume 21, Number 1, 15-30

Reductive Reasoning: A Cognitive Barrier to Visual Literacy

George H. Perino, Jr. Ph.D.
Defense Systems Management College
Fort Belvoir, Virginia, USA

Abstract

The human mind, on its own, is easily overwhelmed by the multi-faceted relationships inherent in complicated management issues. Graphical displays are often used in an attempt to reduce cognitive overload. Results of an empirical investigation into the effectiveness of relational diagrams designed for use in the management of complexity are presented. These graphical displays provide a figurative, rather than literal, portrayal of reality. Their logic is based on inference rather than causality. We tested first-time viewer ability to interpret a modestly sized diagrammatic display of problem elements and their interrelationships. Results showed significant misinterpretation of the diagram. Study participants were highly schooled technically oriented individuals attending a mid-career course in management. Analysis of responses to questions about the meaning of the display indicated that participants were predisposed to reductive reasoning with emphasis on cause and effect as a principle mode of thought.

The inability of the human mind to process more than a few bits of information simultaneously is well known (Miller, 1956; Simon, 1974; Warfield, 1988). The tendency is to under-conceptualize complexity, avoiding cognitive overload. The use of graphical displays to make visible the structural nature of complicated issues is one technique that appears to