A COMPONENT TASK ANALYSIS OF STEREOSCOPIC DISPLAYS

by

Robert Howard Miller

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Industrial and Systems Engineering

APPROVED:

R. J. Beaton, Chairman

C. P. Koelling

A. M. Prestrude

H. L. Snyder

R. D. Williams

April, 1994
Blacksburg, Virginia

A COMPONENT TASK ANALYSIS OF STEREOSCOPIC DISPLAYS (ABSTRACT)

by

Robert Howard Miller

Committee Chairman: Robert J. Beaton
Industrial and Systems Engineering

Considerable research has centered around the issue of whether stereoscopic (3D) viewing allows improved viewer performance for tasks that involve three-dimensional information. Taken as a whole, such previous research indicates that the potential stereoscopic advantage may be dependent on the nature of the task being examined. This task dependency makes it difficult to predict whether stereoscopic viewing will improve viewer performance for a given untested task. By measuring performance over a variety of component tasks, this research examined the potential task-dependent nature of the stereoscopic advantage. In addition, a method was proposed to employ such component-task data for predicting the stereoscopic advantage within future unknown tasks.

A set of 12 component tasks (in six task groups, each with two representative tasks) was developed to represent the various task demands of processing 3D visual information. Participants performed each of the 12 component tasks in both a monoscopic (2D) and a stereoscopic (3D) viewing condition. Performance was measured in terms of viewing time, percent accuracy, and a generic mental effort rating.

Results indicate that when certain display guidelines are not violated, stereoscopic display improves or at least maintains the overall level of viewer performance for most tasks. Furthermore, the results clearly indicate that the stereoscopic advantage is dependent on the nature of the task. Although further refinement to the set of component tasks is necessary before the precise nature of the task dependency can be determined, the component task method displays considerable promise for being able to predict the stereoscopic advantage for any number of complex 3D tasks.

ACKNOWLEDGEMENTS

This research was funded in part by Texas Instruments, Inc. The stereoscopic display apparatus used in this research was supplied by Tektronix, Inc.

I would like to thank each of the advisory committee members for their efforts in guiding this research. Particular thanks go to Dr. Robert J. Beaton, whose infinite patience and ability to maintain sight of the larger issues were instrumental in the completion of this work. Thanks also go to Dr. Gregory J. Buhyoff, who was unable to continue serving on the advisory committee. Each of these gentlemen will be long appreciated for their help in shaping my academic outlook.

Endless appreciation goes to my parents, Mr. and Mrs. Robert Charles Miller. They have always been my strongest examples of fortitude in any endeavor. From their teachings, I have found the perseverance and motivation to complete this arduous process.

I would like to express my gratitude to Dr. Robert J. Logan and Louis Lenzi of Thomson Consumer Electronics. Their patience over the last year has been most generous.

Thanks go to Willard Farley, who always has advice on how to get things done.

Finally, I cannot begin to measure the amount of personal support provided by my love, Kerry Aronoff. Her words of encouragement have always been the needed spark in times of despair.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

It is generally accepted that an important means for improving the transfer of information from a visual display to its viewer is to render visual information in a manner compatible with the viewer's mental image. Because different mental images coordinate with different types of information transfer, there are many techniques for achieving this display-viewer compatibility. The topic of this research is the information contained within a three-dimensional (3D) space, the coordinating mental image of such space, and the best means to visually display such information. Specifically, this research focuses on one elementary question: whether stereoscopic (3D) displays transfer spatial information in a manner inherently superior to two-dimensional (2D) displays. This potential superiority is referred to as the *stereoscopic advantage*.[1]

This is not the first attempt to answer the 2D versus 3D question. However, this work employs a technique not found in previous evaluations, which is to isolate the basic component tasks involved in interpreting 3D visual scenes and to base the 2D/3D comparison on these tasks as a group. The reason for this approach is simple. When taken as a whole, previous 3D research leads to one primary suspicion: the stereoscopic advantage appears to be task-dependent. Consequently, although many have demonstrated a stereoscopic advantage or disadvantage for a variety of *specific* visual tasks, such research did not determine whether the retinal disparity depth cue is universally potent. Nor did this research lead to an understanding of the nature of the task dependency. Thus, it has been impossible to predict empirically the stereoscopic advantage for a system which does not directly match one studied in previous research.

This component task research helps to define more narrowly the efficacy of retinal disparity in relation to other depth cues. Furthermore, this research provides data that may help predict the stereoscopic advantage for any complex 3D task. By combining the component tasks, a more complex 3D task may be modeled. Based on the level of stereoscopic advantage for each of the component tasks, the expected stereoscopic advantage for the complex task may be determined.

---

[1] The term *stereoscopic* advantage should not be confused with *binocular* advantage. While both terms relate to increased visual performance as a result of viewing a scene with two eyes, a *binocular* advantage is due to redundancies between the two retinal subimages, whereas a *stereoscopic* advantage is due to horizontal disparities between the two retinal subimages.

DEPTH CUES

The following section includes a brief description of monocular and binocular depth cues. For a more detailed description, consult publications by Miller (1991) or Reinhart (1990).

Monocular Depth Cues

Monocular cues are those that require only one eye for perception. Monocular cues are typically used in 2D displays to achieve pseudo-3D appearance. Such displays often are called $2^1/_2$D displays in the displays literature. Indeed, proper use of monocular cues can create convincing depth in many scenes.

Luminance contrast. In most real-world scenes, bright objects are perceived as closer to the viewer than dim objects. However, Coules (1955) found that the direction of this effect is dependent on background luminance. Objects of high relative luminance appear closer when the background is darker than objects of interest, and objects of low relative luminance appear farther when the background is brighter than objects of interest.

The method of coding depth via luminance cues is often referred to as Proximity-Luminance Covariance. In this method, the luminance of objects increases in inverse proportion to viewing distance.

Accommodation. Accommodation of the lens of the eye provides a cue to the distance of the object being viewed. The ciliary muscles of the eye, which contract or relax around the lens, serve to focus images on the retina. The proprioceptive feedback associated with these muscle contractions contributes to the sensation of depth. However, Julesz (1971) pointed out that accommodation is an insensitive depth cue for two reasons: 1) the evaluation of blur is dependent on the contour properties of the stimulus, and 2) in bright surroundings the reduced pupil diameter results in an increased depth of field.

Interposition. If an opaque object is interposed between the viewer and a more distant object, then that opaque object will occlude at least part of the more distant object. Since occlusion of an object very clearly indicates its relative distance, interposition is a potent depth cue, as shown in Figure 1. One limitation of interposition is dependence on predictable object contours. In Figure 2, the shapes of the objects are ambiguous and so is their depth order (Buffett, 1980).

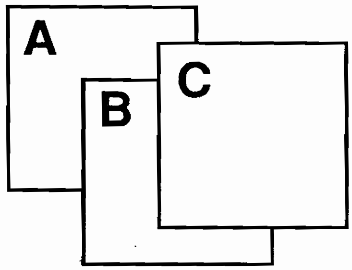


Figure 1. *An illustration of the depth cue of interposition. The occlusion of squares A and B yields potent information about their position behind object C.*

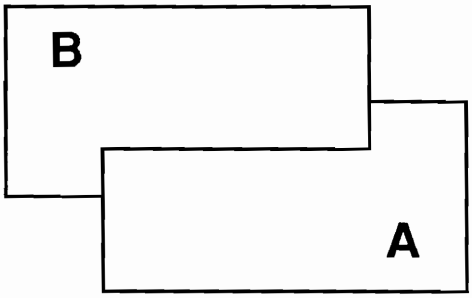


Figure 2. *An illustration of interposition ambiguity. Both object contours and depth order are uncertain. (Adapted from Buffett, 1980)*

Linear perspective. Linear perspective refers to the diminishing visual angle between two parallel lines as viewing distance increases. The depth cue of relative size refers to the effect of linear perspective on objects that lie at different distances along a viewer's line of sight.

Shading and shadowing. Shading, a form of luminance gradient, can indicate the shape of objects if the direction of the light source is either known or can be derived. Shading also can resolve depth ambiguity in sterile images such as the Necker Cube (Figure 3). Shadowing, also a luminance gradient, is dependent on the direction of light sources, the shape of objects, and the terrain underlying the object (Cavanaugh and LeClerc, 1989). The depth effect of shadowing is illustrated in Figure 4.

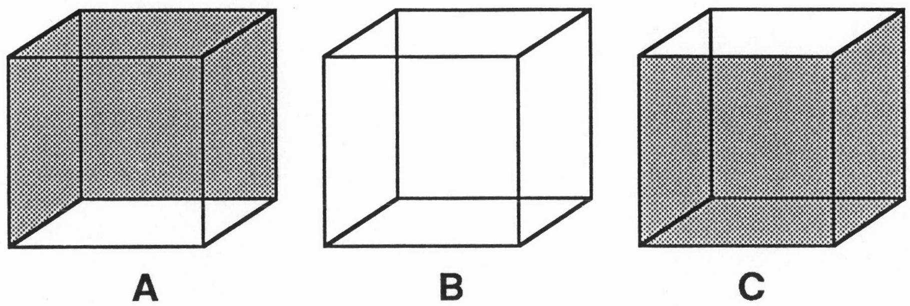


Figure 3. *The depth cue of shading. The orientation of box B is ambiguous, but the orientations of boxes A and C are more apparent.*

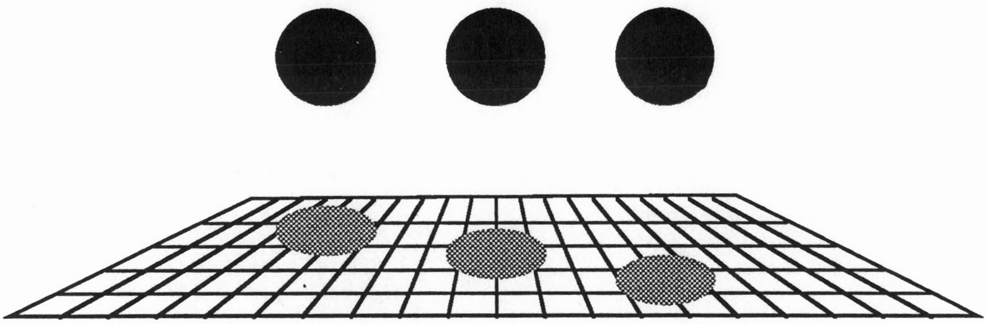


Figure 4. *The depth cue of shadowing. Without shadows, the relative depth order of the three spheres would be ambiguous.*

Detail perspective. Detail perspective results from the limitations of human visual acuity. At a distance, high spatial frequencies are attenuated, causing loss of visible detail and smoothing of textures.

Aerial perspective. Apparent when viewing a distant horizon such as a mountain range, aerial perspective has two causes: (1) airborne particles (aerosols) result in the loss of luminance contrast, detail, and color purity, and (2) the greater transparency of the atmosphere for shorter wavelengths of light makes objects appear bluish.

Motion Parallax. A potent depth cue, motion parallax is apparent when either the viewer or the object(s) being viewed moves perpendicular to the viewer's line of sight. The apparent rates of movement of objects in the visual field provides a cue as to each object's relative distance from the viewer, as closer objects appear to move faster than farther objects.

Binocular Depth Cues

Binocular depth cues are those that require two eyes to be seen. More precisely, binocular cues exist only when different subimages are presented to each eye. Of course, this situation occurs in normal viewing, but also can be simulated via stereoscopic presentation.

Vergence. Vergence refers to the proprioceptive feedback associated with the inward (convergent) or outward (divergent) rotation of the eyes as the fovea meet at a common point. Vergence is associated closely with accommodation and is limited in a similar manner. Beyond a few meters, vergence cues become ineffective as the optical axes approach parallel.

Retinal disparity. Retinal disparity refers to the degree of angular separation between the left and right retinal subimages. Since the eyes are separated horizontally in the head, the two retinal subimages differ, and the difference between these subimages is a function of the stimulus distance. Disparate retinal images are fused in the brain, and the resulting perception is that of visual depth. All virtual 3D displays that render images on a plane utilize retinal disparity. By displaying a different subimage to each eye in the form of a stereo pair, a convincing depth illusion can be created.

Retinal disparity can take two forms, as illustrated in Figure 5. As left- and right-eye subimages are shifted horizontally, they may do so with or without crossing the line of sight for each subimage. If the subimages separate without crossing, then the apparent depth of the image will increase: this is uncrossed (positive) disparity. If the subimages cross so that the left-eye subimage is placed to the right of the right-eye subimage, then the apparent depth of the image decreases: this is crossed (negative) disparity. If right- and left-eye subimages occupy the same position, then there is no disparity and the virtual image rests on a neutral plane. Disparity may be quantified in terms of the distance between the two subimages. For example, if the left-eye subimage is 2 arcminutes to the right of the right-eye subimage (crossed disparity), then the image has a total disparity of -2 arcminutes.

Uncrossed (positive) disparity can progress only to a point at which the optical axes are parallel. Beyond this point divergence is required to fuse the image. Divergence, a phenomenon which should never occur during normal viewing, is tolerable up to approximately two degrees of rotation.

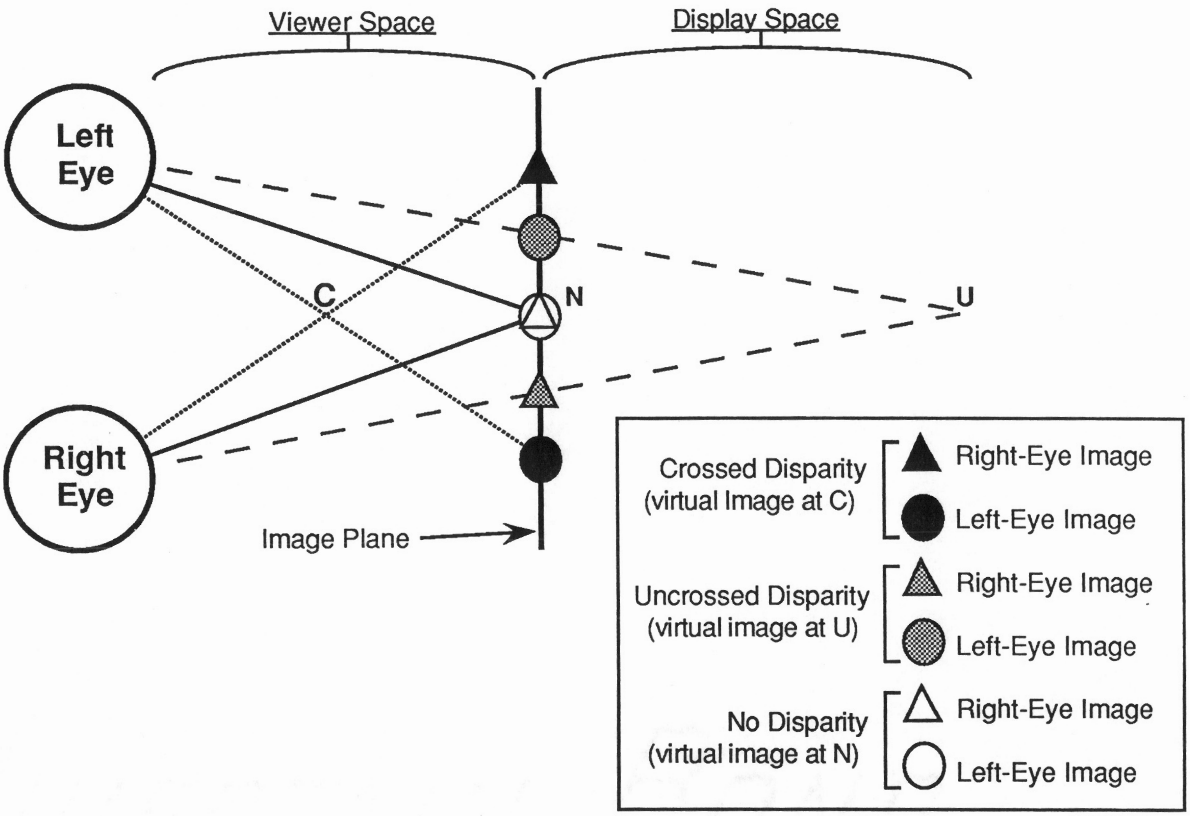


Figure 5. *An illustration of crossed and uncrossed retinal disparity as rendered stereoscopically.*

PERCEPTION OF VISUAL SPACE

X = Y = Z

Uttal (1983), based on the results of 16 experiments, concluded that visual space is homogeneous, symmetrical, and isotropic. In other words, visual perception deals with all three dimensions equivalently, and equal stimulus changes in any direction correspond with equal perceptual effects. If one considers the virtual two-dimensional nature of the retina, it seems unlikely that the optical system would respond equivalently to changes in depth (Z) as to vertical (Y) or horizontal (X) changes, for there is no inherent three-dimensionality coded within retinal images. This, however, is the power of binocular vision, as two disparate retinal images are integrated to achieve depth perception. This integration suggests the existence of a high-level cognitive process which fuses the retinal images into a single spatial perception. Given the demonstrated perceptual equivalence of X, Y, and Z space, Uttal concluded that X and Y perception also must belong to a higher-order process, which further suggests that spatial perception is performed as a single process, rather than an integration of various visual cues.

The above argument leads to a significant conclusion: retinal disparity is a spatial cue as fundamental to the perception of three-dimensional space as are the horizontal and vertical spatial cues. With regard to visual displays, this logic suggests an inherent power in stereoscopic displays—a power readily apparent to many who have viewed such a display. However, if the retinal disparity cue is so fundamental, then why are stereoscopic performance effects typically task-dependent? Why doesn't stereoscopic viewing improve information transfer in any three-dimensional scene? Surely, the removal of either the X or Y spatial cue would render a visual display impotent with regard to information transfer, yet viewers frequently perform adequately without the retinal disparity (Z) cue.

There are two predominant answers to this dilemma: either the retinal disparity cue is not as potent as argued above and is not always a dominant depth cue, or there are many three-dimensional tasks that do not require retinal disparity information. Although either option points to the relative importance of other depth cues, the decisive issue is whether these cues are monocular or binocular, and in which specific three-dimensional tasks these differences exist.

Retinal Disparity Issues

Random dot stereograms. It was once thought that binocular depth perception was dependent upon vergence as a form of triangulation, via oculomotor feedback. Evidence to the contrary lies in the work of Julesz (1960), who developed the random dot stereogram (RDS). An RDS consists of two fields in a stereo pair, each with an identical random dot pattern throughout the field (Figure 6). Within a certain area of the field, retinal disparity is applied. The resulting stereoscopic image contains an area of random dots that appears closer or farther than the rest of the field, depending on the direction of the disparity. If depth perception of the RDS scene were dependent upon vergence, then recognition of the hidden object would necessarily precede perception of depth. In an RDS, however, recognition of the object is achieved through perception of its depth. The object cannot be disclosed without sensing its visual depth relative to its background. Therefore, it must be the binocular fusion of the two retinal subimages, followed by recognition of the components of a visual scene, which leads to understanding of the complete scene.

Based on his work with RDSs, Julesz claimed that the retinal disparity depth cue is the most potent depth cue. Although the RDS research clearly eliminated vergence as the depth cue of choice, the claim of retinal disparity dominance is not universally accepted.

Task dependency. Many studies (Cole and Parker, 1989; Dosher, Sperling, and Wurst, 1986; Lippert and Benser, 1987; Mountford and Somberg, 1981; Reinhart, 1991; Smith, Cole, Merritt, and Pepper, 1979; Spain and Holzhausen, 1991; Woodruff, Hubbard, and Shaw, 1986; Yeh and Silverstein, 1990, 1992; and Zenyuh, Reising, and Walchli, 1988) have demonstrated an overall stereoscopic advantage for the improvement of viewer performance. However, retinal disparity cues do not always provide for improvements in viewer performance, a fact that suggests the potency of other depth cues. Aside from other depth cues, a more significant difference is the type of task used to evaluate the efficacy of retinal disparity cues. Accordingly, the following sections describe studies which demonstrate or address the task dependency of retinal disparity.

Wickens, Todd, and Seidler (1989) concluded that "the advantages gained by 3D technology are probably somewhat task specific and that tasks in which a holistic awareness is critical may be facilitated by the use of 3D technology." This conclusion was based on the identification of three risks associated with 3D displays, which apply equally to both pseudo-3D and stereoscopic displays:

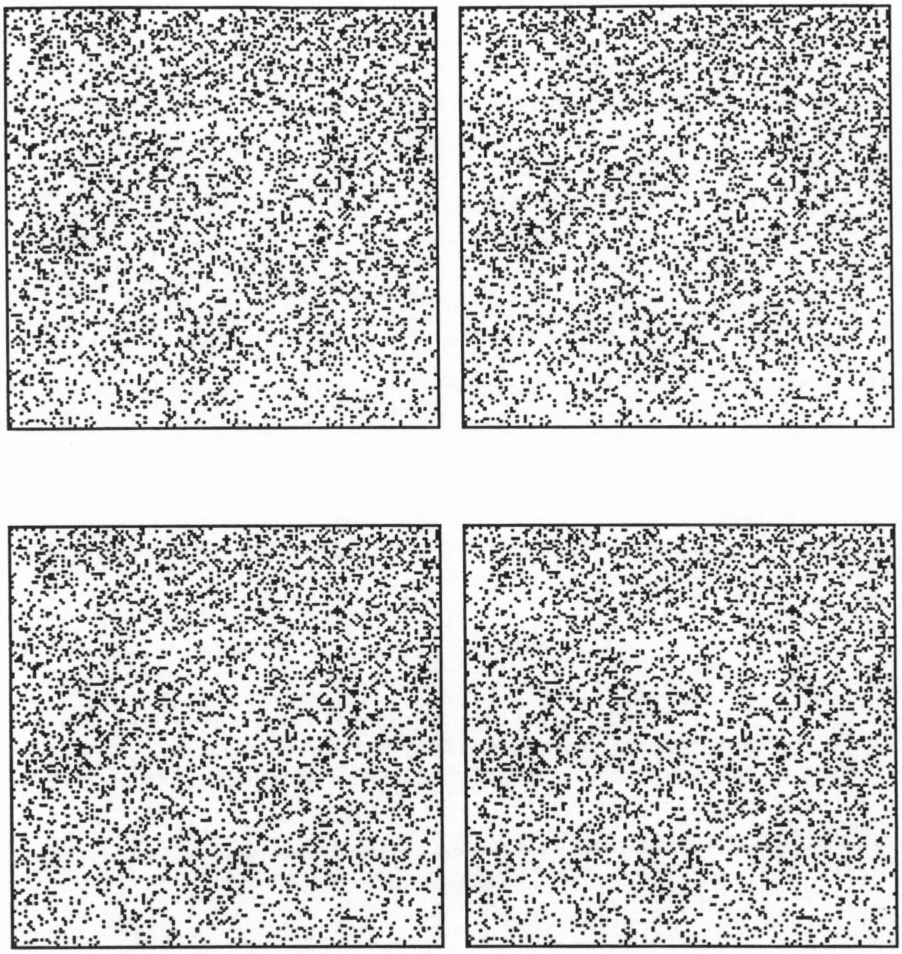


Figure 6. *Two random dot stereogram pairs. When viewed stereoptically, a T-shaped figure appears to float above the background (upper pair) and behind the background (lower pair). To view these images, use the viewing glasses located at the back of this document.*

(1) Perceptions of absolute distances parallel to the line of sight are inherently less accurate than perceptions of such distances perpendicular to the line of sight. Although this is a limitation of 3D displays, Wickens *et alia* noted that the limitation also applies to real-world viewing.

(2) Portraying a solid in an integrated 3D manner (as opposed to separate 2D elevations) may result in reduced precision in absolute spatial perception along *any* one particular axis (Carswell and Wickens, 1987).

(3) 3D displays can complicate the proper generation of imagery, as they introduce additional design issues, such as determining the optimum field of view and viewing angle, as well as additional hardware issues, such as frame rate, viewing spectacles, et cetera.

Kim, Ellis, Tyler, Hannaford, and Stark (1987), in a comparison of several forms of perspective display enhancements, compared monoscopic and stereoscopic displays. Their results indicate that without perspective enhancements, stereoscopic viewing results in improved performance. However, in a visually enhanced setting, where vertical reference lines and a textured ground surface added monocular depth information, stereoscopic presentation demonstrates no significant advantage. The researchers concluded that the vertical reference lines, and not the grid, are the actual cause of the improvement.

Nataupsky and Crittendon (1988) compared pilot performance under two types of monocular depth information and found that a "monorail" display is less effective than a "signpost" display. Replication of the comparison under stereoscopic viewing revealed an improvement for the inferior monorail display, but no stereoscopic advantage for the superior signpost display.

In a simplified flight-simulation environment, Way (1988) found that addition of retinal disparity cues reduced both response time and error frequency for a three-dimensional tracking task. However, when introduced as a means for highlighting a selected object on an otherwise "flat" display, retinal disparity cues did not produce a difference in either error frequency or response time.

Way, Hobbs, Qualy-White, and Gilmour (1989) conducted a full-mission fighter simulation of both air-to-air and air-to-ground scenarios, in which various flight-information displays (ground and air perspective situations, a HUD "highway in the sky" display with a top-down map, and a horizontal situation display) were presented in both 2D and stereoscopic modes. Results of pilot performance showed no stereoscopic advantage. Interestingly, the researchers reported that pilots felt the display resolution was degraded under stereoscopic viewing conditions, which suggests improper implementation of the 3D hardware.

Way (1989) conducted a study designed to assess the efficacy of retinal disparity depth cues in a context involving a complete cockpit and mission simulation. Way used a perspective display to present position, course, and enemy threat information. The display was enhanced with vertical reference links from airborne objects to the ground plane of the airspace. Out of 15 performance measures, only 1 showed a statistically significant advantage. Way concluded that the effect of retinal disparity in the perspective display format is not significant.

Miller and Beaton (1991) evaluated the potency of retinal disparity cues for two tasks in each of three display formats: (1) a plan view, (2) a simple perspective view, and (3) an enhanced perspective view. Stereoscopic viewing had no effect on search times or error levels for either perspective view. However, stereoscopic viewing did reduce errors for one task in the plan view display.

Merritt (1984), in a discussion of tasks which *require* stereoscopic viewing, claimed that "it is not possible to observe the benefits of 3-D displays if they are not tested and properly evaluated for a given type of task." Although this statement was made in defense of stereoscopic displays, it actually burdens the display system designer who, according to Merritt's logic, could not predict the inherent utility of stereoscopic displays without first evaluating them for the task at hand. Of course, it would be undesirable to evaluate each and every task to which stereoscopic displays might be applied. Therefore, there exists the need for a generalizable prediction and some method that could predict stereoscopic utility for any known task.

Cognitive workload. Livingstone and Hubel (1988) presented evidence of a neural pathway for the processing of stereopsis, a pathway parallel to those for color and form information. Based on this evidence, Gooding (1991) hypothesized that the presence of retinal disparity depth cues would, according to the theory of multiple resources, spread cognitive workload demands over a greater number of channels, thereby decreasing overall cognitive workload. Using a 3D spatial task incorporating varying levels of perspective information and scene complexity, Gooding analyzed the effect of retinal disparity cues on cognitive workload by using both a primary-secondary task paradigm and subjective ratings via the Subjective Workload Assessment Technique (SWAT). Results indicated that the main effect of retinal disparity on cognitive workload was not significant for either the secondary task or the SWAT measure. Interactions with scene complexity and perspective orientation also were not significant. Based on these results, Gooding speculated that although retinal disparity may provide the viewer with more depth information, such information requires additional processing and, thus, does not decrease overall cognitive workload. Although this hypothesis does not agree with the aforementioned independent channel theory, it was supported by the findings of Gooding's second experiment. In her second experiment, Gooding used a target identification task that did not require depth information and found that retinal disparity cues increased the mean SWAT rating for some conditions.

The results of Gooding's study may be explained as follows: although the presence of retinal disparity cues may reduce the portion of cognitive effort required to process other spatial cues, this effect is counteracted by the additional portion of cognitive effort required to process the retinal disparity cue itself, and so the overall cognitive workload does not change according to the presence or absence of retinal disparity cues. However, just as the results of most 3D studies are lacking in generalizability due to the task-specific nature of the potential stereoscopic advantage, we must also question the generalizability of the workload results from Gooding's experiments. Accordingly, this experiment included a simplified unidimensional cognitive workload measure.

COMPONENT TASKS

Rationale

If it is determined that the stereoscopic advantage is indeed task dependent, then it would be difficult to determine whether stereoscopic display is appropriate for a given application. Accordingly, this study proposes a method of analysis through component tasks that seeks to improve the generalizability of results toward any potential 3D display application. This method assumes that when interpreting a visual scene, the viewer performs (based on the task and the scene) some finite number of cognitive analyses to interpret the spatial information content of the scene. Further, these cognitive analyses may be described using component tasks; each epitomizing one type or class of cognitive analysis.

Most discussions of spatial perception are concerned with the various spatial cues and depth cues, both monocular and binocular. In fact, the spatial perception literature is filled with discussions of spatial cues and depth cues and their various interactions. The goal of such discussions is to better understand the components of a visual scene and how those components contribute to veridical space perception. What is not found in these discussions is the actual cognitive needs that may or may not be met by the cues. In other words, what are the various tasks that, when combined, comprise the process of perceiving spatial information, and how are these tasks supported by the various spatial cues? The component tasks method is an attempt to represent these fundamental spatial decisions.

To restate this argument, it is known that certain visual cues exist in a given scene and these cues yield spatial information about that scene. However, what has not been well described are the perceptual needs that are fed by these visual cues. To say that linear perspective provides enough information to perform a visual task is a different matter than describing the perceptual task itself. Consider the following analogy: humans receive auditory information through cues such as pitch, amplitude, timbre, et cetera, and the perception of these cues is important in replicating a desired auditory "scene." However, the component task method would describe the various listening tasks that are supported or not supported by the various cues. For example, to isolate the difference between a trumpet and a violin, timbre is important. But to hear the difference between a bass and a violin, pitch cues are more important. To know the overall importance of a given cue, its contribution to each of many auditory tasks must be evaluated.

With respect to the goal of isolating the nature of the stereoscopic advantage, the component task method attempts to determine the specificity or universality of retinal disparity cues toward the various component tasks. In other words, retinal disparity may or may not supply the spatial information necessary to complete a given component task. Once we have evaluated the contribution of retinal disparity cues to each task, the overall efficacy of retinal disparity cues may be determined, either globally (by considering all component tasks equally) or specifically for a complex 3D task (by considering a weighted subset of component tasks).

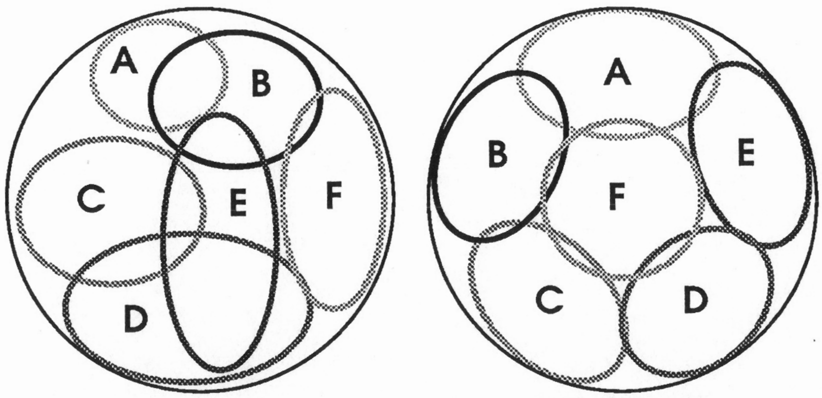
For full interpretation of a real-world scene, the number and variety of component tasks is very large, without question. However, when the scene content is limited[2] to graphic imagery, the number and variety of component tasks is reduced drastically. When the focus is further limited to those mental analyses that deal exclusively with the interpretation and understanding of 3D spatial information, the set of component tasks that effectively approximate those analyses becomes tractable.

The interpretation of graphic 3D information may be thought of as a set. Within this set, the 3D component tasks are subsets. If the component task subsets are defined optimally, they will represent all of the larger set with very little overlap among subsets. If not optimal, at least we can approximate the set with subsets which represent the vast majority of the set and have only marginal overlap (Figure 7). The 3D component tasks developed for this study are believed to approximate the optimal task definitions.

This breakdown of cognitive processes into smaller, more easily describable elements is not a novel endeavor. For example, Wickens (1984) developed a type of component analysis in his notable model of human information processing. Uttal (1983), in prefacing his taxonomy of visual processing levels, discussed two assumptions underlying this type of piecemeal analysis: (1) perception is an aggregate of virtually independent sub-processes or levels, and (2) it is possible to separately assay these independent processes by appropriate experimental designs. Uttal further noted that these assumptions, although necessary to proceed with his research, are at best hypothetical and at worst fallacious. In spite of this, the breakdown of cognitive processes is necessary to gain even a small understanding of the complex entity that is cognition.

---

[2] This limitation is discussed in detail in the section entitled "Graphic 3D Imagery."



Current Tasks Optimal Tasks

Figure 7. *Pictorial representation of 3D component tasks. Each large circle represents the set of mental tasks performed in the interpretation of 3D graphic imagery, and each of the smaller areas (A–F) represents a component task. The diagram on the left represents less than optimal task definitions—notice the overlap between areas and the amount of space not represented by component tasks. The diagram on the right shows more optimal task definitions—there is little overlap and most of the larger set is represented by the component task elements.*

Component Task Development

There are several simple elements of spatial perception, and the link between these elements and each of the defined tasks will be discussed in turn. In brief, the perception of visual space might be defined as follows: the viewer must (1) perceive the distances between objects, (2) predict future object positions given past positions over time, (3) perceive the dimensions of an object, (4) perceive absolute distances between an object and the viewer or between two objects, (5) perceive relative distance differences between two objects, and (6) identify and discriminate among objects based on their shape and orientation.

The component tasks used in this experiment were developed based on an analysis of the types of visual depth cues available to humans, the types of depth cues presented by different 3D visual displays, the quality of these depth cue presentations, the limitations of graphic imagery, and typical real-world tasks involving 3D viewing. Primarily, these tasks represent the basic functions of spatial analysis as described by Uttal (1983), which are to detect, discriminate, recognize, and perceive. In all, there were 12 tasks developed for this experiment—two for each of the six basic component tasks described above. These tasks are described in detail in the section entitled "Experimental tasks."

Describing a Complex 3D Task

The true utility of the component task method resides in its potential to predict the stereoscopic performance advantage for more complex 3D tasks. The only limitation is that the complex 3D task must use graphic imagery because the generalizability toward real-world imagery is not known. Although analysis of an actual complex task will not be conducted here, a potential method for applying the results of this experiment to more complex systems will be described.

For each component task, the results of this experiment predict the performance advantage or disadvantage associated with the inclusion of retinal disparity cues. These performance differences are with regard to speed or accuracy and may be calculated as the ratio of performance with retinal disparity cues over performance without retinal disparity cues. For example, if speed is doubled in the presence of retinal disparity cues, then the performance advantage in that case is 2. If accuracy in the retinal disparity condition improves by 30%, then the performance advantage in that case is 1.3.

Consider a complex system in which an operator is required to perceive a 3D space (rendered with graphic imagery) and then to make decisions about the objects in that space. To determine the efficacy of retinal disparity cues within this system, it would be necessary to determine the basic tasks performed by the operator. Next, for each of these basic tasks, it would be necessary to determine the importance of each of the six component tasks described above. For example, consider a task in a Computer-Aided Design environment, where the operator must choose a 3D object of a particular shape and place it at the same depth as other similar objects. To complete this task, the operator must recognize an object by its 3D shape and compare its depth relative to other objects. Accordingly, the component tasks of object recognition and relative depth judgment would receive high importance scores. Other component tasks which contribute little to the task would receive low importance scores. The contribution of each component task towards completion of the whole task would be weighted according to its importance rating and scaled so that all weights sum to 1.

Once each task has been analyzed, the importance weighting of each component task would be multiplied by the performance advantage associated with that task, as determined by the results of this experiment. The sum of these values indicates the expected performance difference given the inclusion of retinal disparity cues. This procedure is illustrated in Figure 8.

The above procedure could be conducted separately for speed and accuracy, or use both measures in a combined fashion. The final values for each basic task then could be combined in weighted fashion (based on task frequency, for example) to make a prediction for the system as a whole. Predicted values greater than 1 indicate a performance improvement, whereas values less than 1 indicate a performance decrement. The practical significance of deviations from 1 is situation-dependent. That is, for some systems a 5% improvement may be considered nominal, whereas for other systems a 5% improvement may be outstanding.

For a given task, there may be component tasks other than those used in this experiment. If we assume that addition of retinal disparity is neither advantageous or disadvantageous for component tasks not included in this experiment, then such component tasks could be added to the above procedure by simply using a value of 1.0 to represent the observed performance effect. On the other hand, there may be some component tasks for which the introduction of retinal disparity cues is detrimental. In this case, performance values less than 1.0 would be required—these values may be determined by experimentation or derived from previous research.

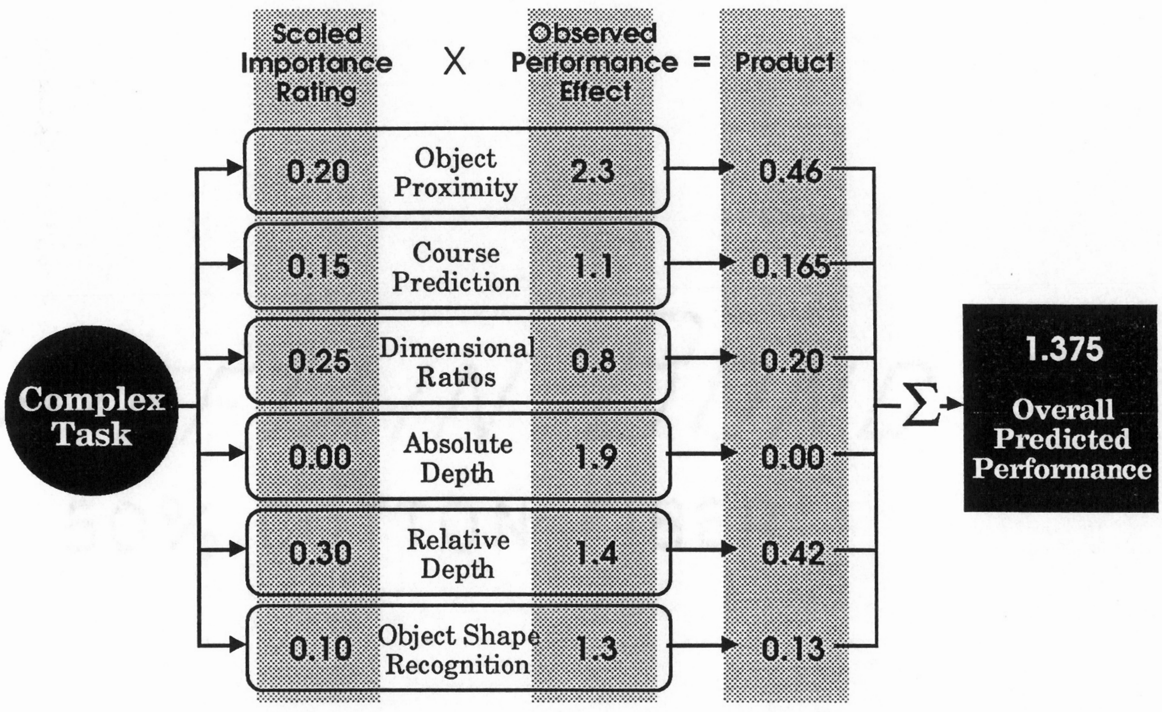


Figure 8. *Example of a method for predicting the stereoscopic advantage in complex 3D tasks. The importance of each component task is measured, then scaled so that the six importance ratings sum to 1. Next, each importance value is multiplied by the performance effect as measured in this experiment. The products are summed, and the total represents the predicted performance level given the introduction of retinal disparity cues.*

Graphic 3D Imagery

This study is limited to graphic images. The reasons for this limitation are due more to practicality than to any scientific rationale. The 3D hardware used in this experiment is limited, and so the scope of this work was similarly limited. Furthermore, graphic images are simpler in content and may be generalized further than complex images. Graphic images are common in conditions where 3D displays might be applied (e.g., computer-aided design, air traffic control) so there is obvious value in a graphics-only evaluation. Of course, the obvious limitation is that no claim can be made for the application of these results towards 3D real-world imagery.

Within this study, the term "graphic imagery" refers to scenes that are created artificially or are abstracts of scenes occurring in the real world. Such images typically are computer-based and consist of lines, curves, and points. The term "real-world imagery" refers to scenes that exist in real space and have been captured through photographic means, or man-made scenes that attempt to replicate such real-world imagery. Examples of graphic imagery are bar charts or line drawings. Conversely, a scanned photograph exemplifies real-world imagery.

Davies (1985) drew a potent analogy between graphic images and written communication. Just as a minimum number of properly expressive words is required before a reader can understand a sentence, so is a minimum number of picture elements (lines, dots, etc.) required before a viewer can interpret an image. As the potency and combination of the words improve, so does the transmission of the concept to the reader. As the number and type of picture elements improve, so does the transmission of the image to the viewer. Davies noted that "there must be a limit to the number and type of parts of speech which we can neglect before our conversation level of English becomes unintelligible. Likewise, there must be a limit to the number and type of depth cues missing from a line drawing which, if exceeded, will result in either an ambiguous figure or a nonsense shape." When taken to this limit of unintelligibility, stereoscopic presentation may fill in otherwise missing information that would have made a scene intelligible. This reasoning suggests a need for stereoscopic display with graphic images, since by nature they can lack potentially important visual cues. Conversely, this reasoning suggests that stereoscopic presentation is less helpful for enhancing high-fidelity "real-world" scenes, a concept which is supported by the fact that stereo-blind or one-eyed people are not considered to be disabled in any real-world context.

It is possible to display images that contain no depth cues (except accommodation to the display). It would be further possible to examine the efficacy of retinal disparity depth cues by comparing stimuli rendered with retinal disparity cues and stimuli rendered with no depth cues. In such a comparison, however, the viewer performance measured under conditions of no depth cues is of little value. The lack of cues will, for some tasks, make impossible the very completion of the viewing task. Furthermore, such scenes are very artificial and uncommon in everyday viewing. A more useful comparison is to evaluate retinal disparity cues in a context where all scenery contains at least a moderate amount of monocular depth information. In other words, the comparison is between a $2^1/_2$D display and a 3D display.

In this experiment, all experimental stimuli contain three monocular cues: linear perspective (relative size), proximity-luminance covariance, and interposition. These cues were selected for three reasons: (1) they are clearly within the domain of graphic imagery, unlike other cues (e.g., detail perspective) which fall more within the domain of real-world imagery, (2) other cues (e.g., accommodation) are difficult or impossible to display using the intended stereoscopic technology, and (3) removal of either linear perspective or interposition may create anti-cues in a given scene. Since linear perspective is based partly on the viewer's assumptions or expectations regarding object sizes, scenes devoid of linear perspective can appear paradoxical or artificial. McBurney and Collings (1984) stated that when relative object sizes are interpreted incorrectly, the resulting anti-cue is potent enough to overpower other depth cues such as luminance, leading to a misinterpretation of depth information. Removing interposition cues from a scene also can create paradoxical images. Since interposition is a function of both object opacity and depth position, removal of the interposition cue not only removes the depth information, but alters apparent object opacity as well.

METHOD

Participants

Twenty-four university students were paid $5.00 per hour for participation in the study. Twelve participants were male, the other 12 were female. Participants ranged in age from 17 to 25 years of age, representing a diverse educational background with a heterogeneous skill set. Participants were screened for acuity, phoria, and stereo vision as described below.

Apparatus

Two-dimensional and stereoscopic stimuli were presented on a Tektronix SGS-620 Stereoscopic 3-Dimensional Graphics Display System. The SGS-620 system consists of a 48cm (19 inch) diagonal Trinitron color cathode-ray tube (CRT), a liquid-crystal (LC) optical polarizer, and passive circularly polarized spectacles. The CRT and LC shutter operated at a left-right switching rate of 120 Hz, and were controlled by a microcomputer (IBM PC AT compatible) equipped with a Tektronix Stereoscopic Graphics Adapter card. Participants viewed the display at a distance of 1 m.

Participants used a QWERTY keyboard. The space bar was used to advance through each trial, and the arrow keys were used to alternate views in Task 12 (described below).

Environment

The room was fully darkened for several reasons: (1) to avoid glare reflections from the planar surface of the liquid-crystal shutter of the SGS-620 display, (2) to compensate for the relatively low luminance level of the display, and (3) to represent a typical viewing environment for displays of this type (Tektronix recommends operation of the SGS-620 in a darkened room). A dim light was placed near the participant's keyboard and shielded so that it provided ambient illumination only for the keyboard. This light could not be seen directly from the participant's position, nor could its reflection be seen in the display. A hand-held dim light at the experimenter's station was also shielded from the participant's position.

Although the room was fully darkened, there was a small amount of ambient illumination from the SGS-620 and from the experimenter's display. Because the level of this ambient illumination would have allowed visibility of the bezel of the SGS-620, a rectangular matte black mask was positioned so as to obscure the bezel. This mask helped to eliminate undesirable anti-cues which could have negatively influenced stereoscopic viewing.

The work surface and surrounding walls were also covered with matte black material. Thus, the only clearly visible surfaces were the keyboard and the SGS-620 display.

Experimental Tasks

Object proximity (Tasks 1 and 2). The goal of the two object proximity tasks was to analyze linear inter-object distances in all three dimensions. In Task 1, the observer judged the distance between one reference object and four comparison objects; the observer then determined which of the comparison objects was closer to the reference object. This task is similar to one used by Miller (1991). Task 2 involved judgment of which two objects in a group were the closest pair. This task is similar to one used by Gooding (1991).

Course prediction (Tasks 3 and 4). The goal of the two course prediction tasks was to extrapolate object positions based on current heading. In Task 3, the observer viewed a single object moving toward a group of stationary objects, then identified which (if any) of the stationary objects would be hit by the moving object. This task is based on a task used by Miller (1991). In Task 4, there were two objects in motion; the viewer predicted collisions between the two objects.

Dimensional ratios (Tasks 5 and 6). The goal of the two dimensional ratios tasks was to identify inequalities in the various dimensions (width, height, and depth) of a given object. In Task 5, the participant viewed three orthogonally opposed objects and identified which object was unequal in length to the other two. In Task 6, participants viewed a simple rectangular solid and determined whether all dimensions of that solid were equal in size.

Absolute depth (Tasks 7 and 8). The goal of the two absolute depth tasks was to accurately judge the depth difference between two objects. Task 7 included an object which moved in depth; the viewer had to rate that depth movement according to a visual depth scale. Task 8 included two objects separated in depth. As in Task 7, the viewer had to rate the depth difference according to a visual depth scale. This task is similar to that used by Gooding, Miller, Moore, and Kim (1991).

Relative depth (Tasks 9 and 10). The goal of the two relative depth tasks was to judge which of two objects appears closer to the viewer. Although judgment of relative depth is a precursor to judgment of absolute depth, it is a separate task in that it becomes more difficult at smaller depth differences and less difficult at larger depth differences. In Task 9, the participant viewed two objects separated in depth and decided which appeared closer, and is similar to a depth judgment task used by Reinhart (1990). Task 10 involved comparison of two outer objects with a central reference object. Here, viewers had to decide which of the outer objects was closest in depth to the center object.

Object shape recognition (Tasks 11 and 12). The goal of the two object shape recognition tasks was to compare two or more three-dimensional objects and decide whether the objects are, in fact, the same exact shape. Task 11 was adapted from the Mental Rotations Test (MRT). The MRT is a paper-and-pencil test developed by Vandenberg and Kuse (1978) and is based on the original work of Shephard and Metzler (1971). For each test trial, the MRT portrays five wire-frame objects. The viewer is required to determine which of the last four are 3D rotations of the first object. In each case, only two of the latter four objects actually match the first object—the others only appear similar. For the purposes of this experiment, the MRT in Task 11 was diluted to a pair comparison, in which the participant viewed only two 3D objects and had to decide whether they matched exactly. Task 12, somewhat similar to Task 11, involved matching two 3D point clouds. The point clouds were random groupings of dots, and the observer had to determine whether they were identical in form.

Experimental Stimuli

Image Calculation. The experimental stimuli consisted of randomized left- and right-eye subimages displayed with linear perspective cues. Because it was impractical to render stimuli using simple pre-determined *X,Y* coordinates, a set of formulas was developed to convert Cartesian *X,Y,Z* coordinates into on-screen *X,Y* coordinates that were used to render the experimental stimuli. Such formulas yielded proper randomizations of rotation and translation for each experimental trial.

The first conversion of the original *X,Y,Z* coordinates was rotation, which was applied only to tasks which contained rotated scene elements (Tasks 5, 6, and 11). Three separate rotation calculations were performed, one for each axis. For example, rotation about the *Z*-axis (parallel to the viewer's line of sight) required a translation of *X* and *Y* as follows:

$$x_r = \sin(\theta_z + \theta_o) \times \sqrt{x_o^2 + y_o^2} \quad (1)$$

and

$$y_r = \cos(\theta_z + \theta_o) \times \sqrt{x_o^2 + y_o^2}\,, \quad (2)$$

where

θ_z = angle of rotation around the Z-axis,

θ_0 = $\tan^{-1}(x_o, y_o)$, the angle formed by x_0 and y_0,

x_0 = original position of object along the *X*-axis, and

y_0 = original position of object along the *Y*-axis.

Similar conversions were applied to rotations about the *X* and *Y* axes, yielding x_r, y_r, and z_r values.

After rotation, object coordinates were converted to present linear perspective. To accomplish this, object sizes were adjusted as a function of the object's position along the Z-axis. This function resized images to subtend a visual angle equal to that resulting from placing the object at the actual apparent distance. This technique is illustrated in Figure 9.

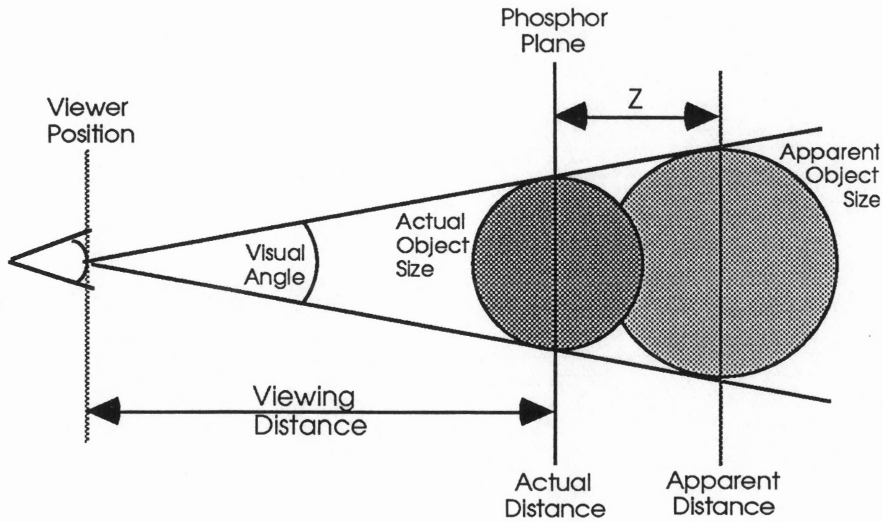


Figure 9. *Illustration of image resizing to present linear perspective cues. Here, the size of the actual object (dark gray circle) is reduced at the phosphor plane, yielding an apparent object (light gray circle) of the intended size which appears to exist at the farther distance.*

The visual angle of the actual object rendered at the phosphor plane is

$$VA_{actual} = 2\tan^{-1}\left(\frac{x_{perspective}}{2D_{viewer}}\right) \tag{3}$$

where

$x_{perspective}$ = position of *actual* object along the *X*-axis, and

D_{viewer} = distance from viewer to phosphor plane.

The visual angle of the apparent object is

$$VA_{apparent} = 2\tan^{-1}\left(\frac{x_r}{2(D_{viewer} + z_r)}\right) \tag{4}$$

where

x_r = rotated (or original) position of *apparent* object along the *X*-axis, and

z_r = rotated (or original) position of object along the *Z*-axis.

Applying the constraint that the actual and apparent visual angles are equal,

$$VA_{actual} = VA_{apparent}, \tag{5}$$

the resulting function to calculate linear perspective simplifies to

$$x_{perspective} = \frac{x_r D_{viewer}}{D_{viewer} + z_r}. \tag{6}$$

The conversion for Y-values included an adjustment for the aspect ratio of the CRT:

$$y_{perspective} = K_{aspect} \frac{y_r D_{viewer}}{D_{viewer} + z_r} \tag{7}$$

where

K_{aspect} = ratio of pixel width to pixel height (for the SGS-620, K_{aspect} = 1.12).

The final coordinate conversion was to adjust horizontal (*X*) values to present proper retinal disparity cues in the left- and right-eye subimages. First, the total disparity (in distance units) was calculated as:

$$Disparity_{total} = \frac{D_{IP}}{(D_{viewer} + z_r)z_r} \tag{8}$$

where D_{IP} = Interpupillary distance of viewer.[3]

Next, left- and right-eye x-values were calculated as:

[3] The interpupillary distance value used in this experiment was 6.32 cm, which is the 50th percentile as reported by Woodson (1981).

$$x_{left\ eye} = x_{perspective} - \frac{Disparity_{total}}{2} \tag{9}$$

and

$$x_{right\ eye} = x_{perspective} + \frac{Disparity_{total}}{2} \tag{10}$$

Finally, objects in the left-eye subimage were plotted at coordinates ($x_{left\ eye}$, $y_{perspective}$). Objects in the right-eye subimage were plotted at coordinates ($x_{right\ eye}$, $y_{perspective}$).

The 3D workspace. All 12 component tasks portrayed objects contained within a three-dimensional cube (Figures 10–21). Designed to appear to have equal width, depth, and height, the 3D cube defined the boundaries of the 3D "workspace." As such, the positions of task objects never exceeded the boundaries of this cube. At the top and bottom of the 3D cube were two 8x8 grids. These grids aided the viewer in perceiving the cube as an unchanging space—a cube with all right angles. Furthermore, in the absolute depth tasks (Tasks 7 and 8), participants used the grids as a depth scale to help in assigning absolute values to depth differences.

The next 12 sections describe the stimuli associated with each task. In viewing the figures below, the correct answer to the task may be obvious from just monoscopic viewing, and so the need for stereoscopic viewing may not be evident. However, the reader should keep in mind that Figures 10–21 are samples of the stimuli, created using the same randomization algorithms used to create the actual stimuli. During the actual trials, stereoscopic viewing was sometimes most critical toward achieving a correct answer and other times not so critical. This distribution mimics analogous situations in "real world" viewing, where stereoscopic viewing is not always critical for making accurate visual judgments.

In Figures 10–21, there are both a color monoscopic sample and a black-and-white stereoscopic sample. In many cases, the two renderings do not represent identical experimental scenes, but two different randomizations of the same task. The color figures were taken from actual photographs of the SGS-620 display: scanned, sharpened, and rendered in 8-bit color. The stereoscopic pairs were also scanned from actual photographs. However, to better present a clear stereo image in the necessary space, the scanned images were converted (by hand) to line-art images. Therefore, the luminance coding seen in the monoscopic images is not portrayed in the stereoscopic pairs.

All critical scene elements were rendered in red, since this color produces less severe ghosting effects on the SGS-620 (Beaton, 1989).

Task 1 stimuli. Task 1 portrayed five octahedrons located within the 3D cube (Figure 10). One octahedron (labeled X) was located at the center of the cube and the remaining four objects (labeled 1–4) were placed at random, each a different distance from the center object. The viewer's goal in this task was to determine which of the four outer objects was closest to the center object X.

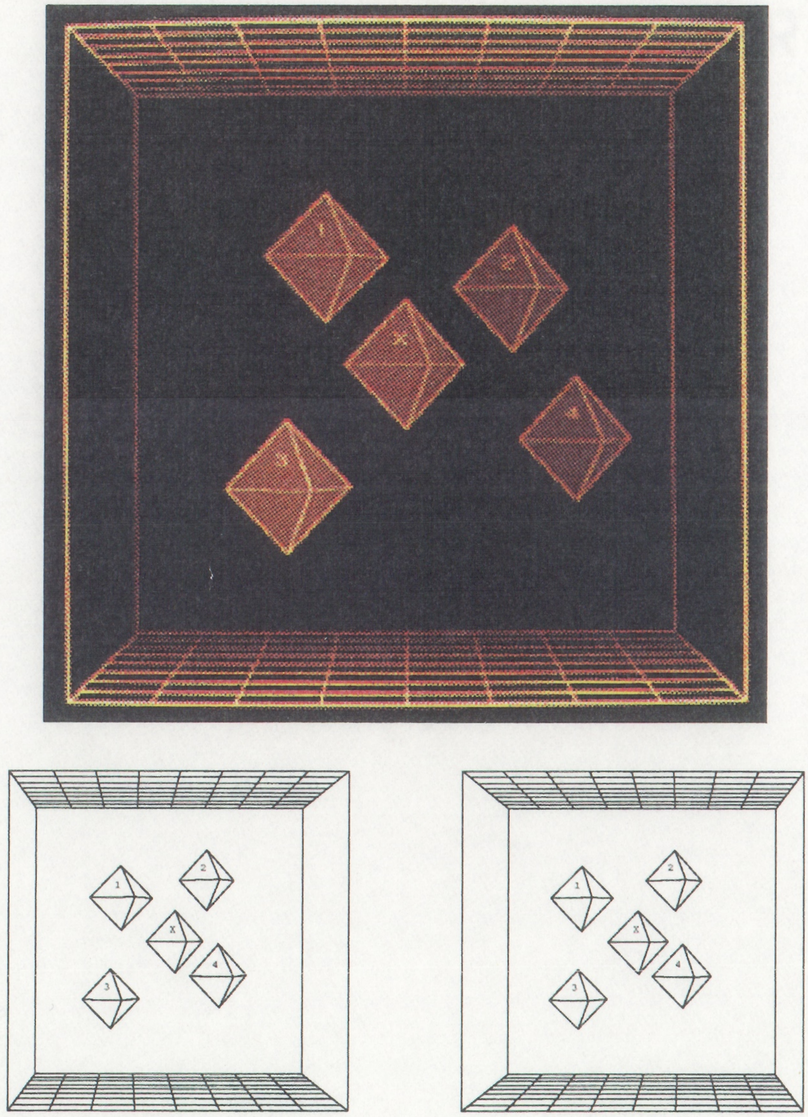


Figure 10. *Samples of experimental stimuli used in Task 1. The upper figure displays the linear perspective and proximity-luminance-covariance depth cues. The lower stereo pair displays the linear perspective and retinal disparity depth cues. To view the stereo pair, use the viewing glasses found in the sleeve on the last page of this document.*

Task 2 stimuli. Task 2 portrayed four octahedra within the 3D cube (Figure 11). Each of the four octahedra were randomly positioned in a different quarter of the cube (upper left, upper right, lower left, or lower right). This method both controlled the relative distances between the four octahedra and prevented overlap of multiple octahedra. The viewer's task was to determine which pair of objects (all possible pairs were potential choices) had the smallest distance between them.

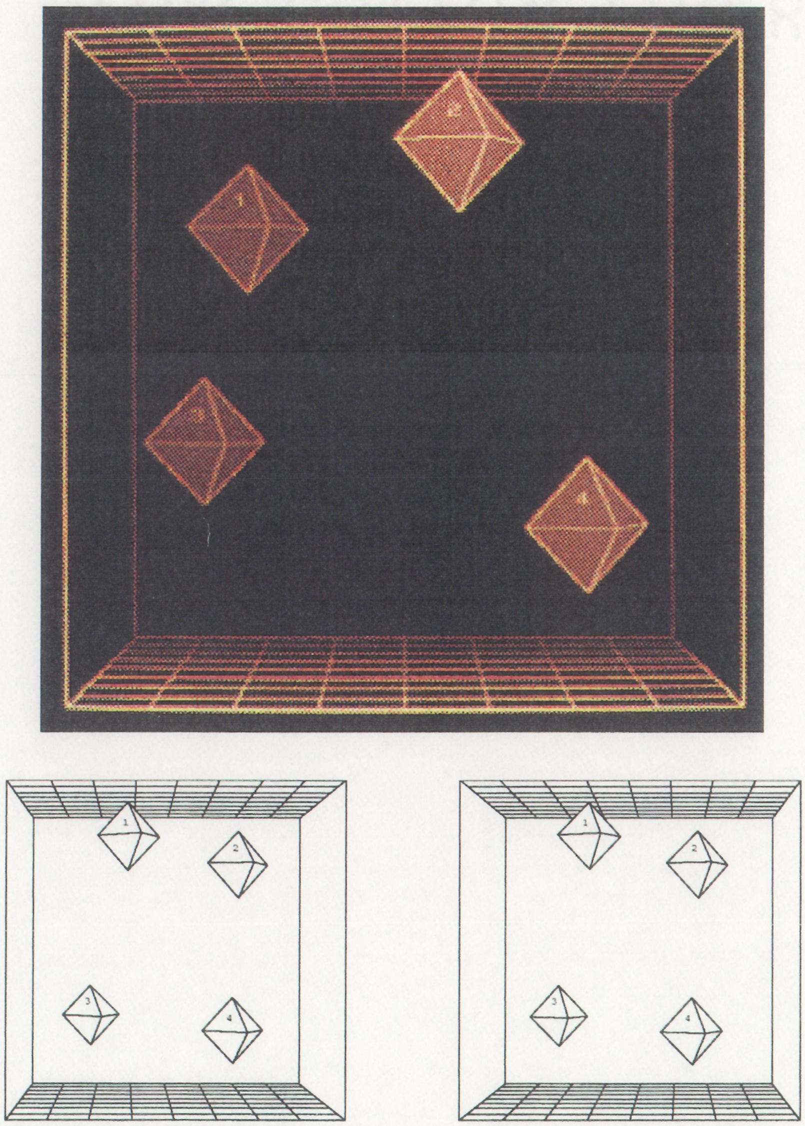


Figure 11. *Samples of experimental stimuli used in Task 2. The upper figure displays the linear perspective and proximity-luminance-covariance depth cues. The lower stereo pair displays the linear perspective and retinal disparity depth cues. To view the stereo pair, use the viewing glasses found in the sleeve on the last page of this document.*

Task 3 stimuli. Task 3 portrayed nine octahedra (Figure 12), eight of which were positioned in the eight corners. The ninth object started at the center of the cube and moved toward an edge or corner, retracing its path again and again. By viewing only half of the complete path, the viewer was required to predict any collisions with the eight stationary objects. All collisions were direct collisions and all misses were complete misses.

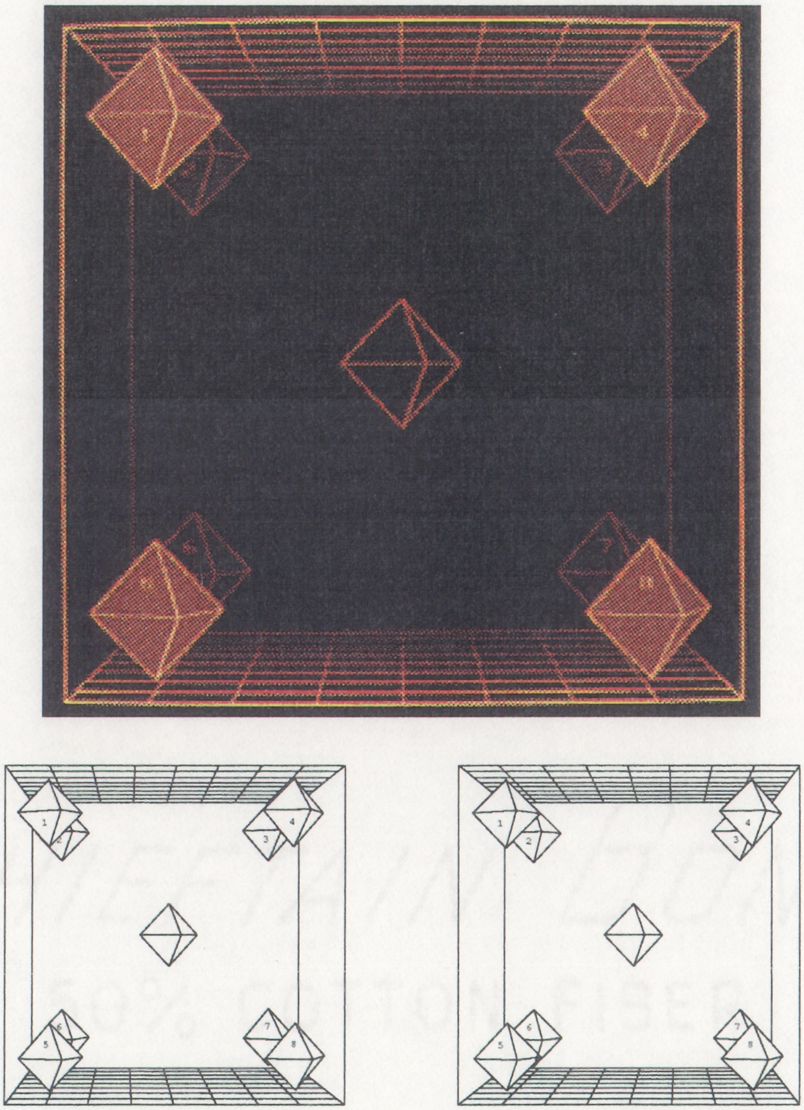


Figure 12. *Samples of experimental stimuli used in Task 3. The upper figure displays the linear perspective and proximity-luminance-covariance depth cues. The lower stereo pair displays the linear perspective and retinal disparity depth cues. To view the stereo pair, use the viewing glasses found in the sleeve on the last page of this document.*

Task 4 stimuli. Task 4 portrayed two moving octahedra that originated in two corners of the 3D cube (Figure 13). The objects were bound to either collide directly or miss each other completely. The viewer saw only half of the complete path; the remainder had to be predicted. The objects retraced their paths again and again, so viewers could continue watching if they were unable to confirm a collision or miss after the first pass.

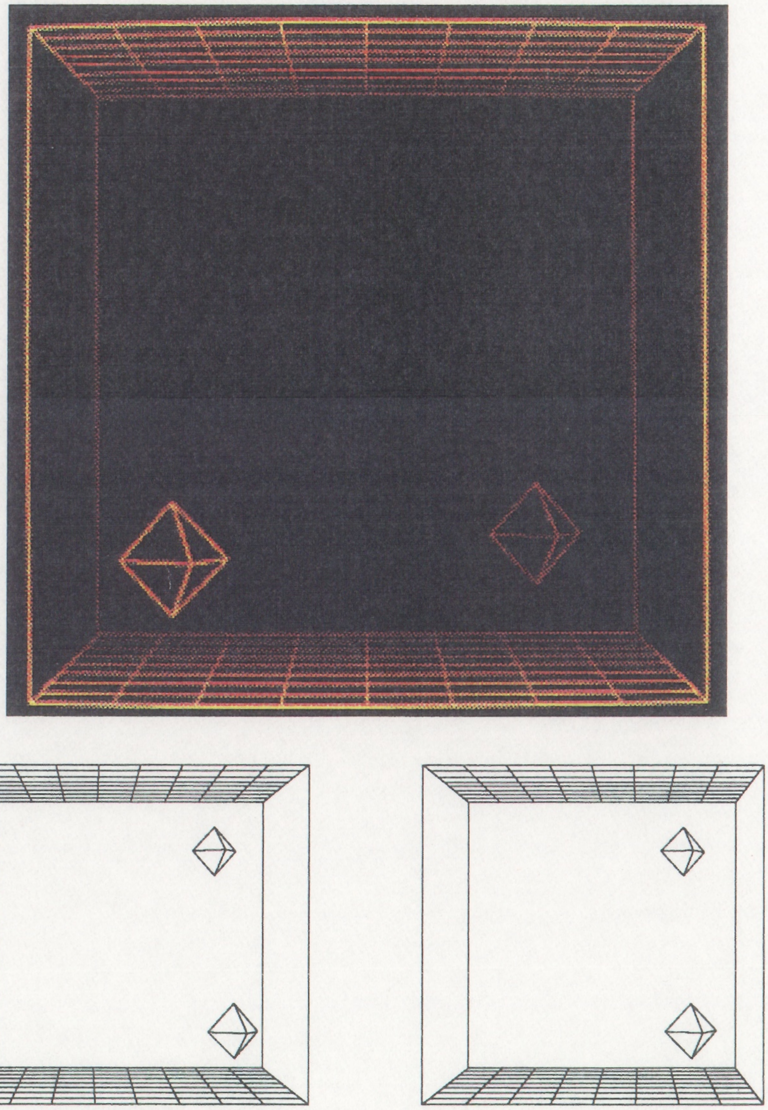


Figure 13. *Samples of experimental stimuli used in Task 4. The upper figure displays the linear perspective and proximity-luminance-covariance depth cues. The lower stereo pair displays the linear perspective and retinal disparity depth cues. To view the stereo pair, use the viewing glasses found in the sleeve on the last page of this document.*

Task 5 stimuli. In Task 5, participants viewed three elongated blocks that were positioned at right angles to each other (Figure 14). One of these three blocks was longer than the other two by 12–30%. The viewer's task was to determine which of the three blocks was the longest.

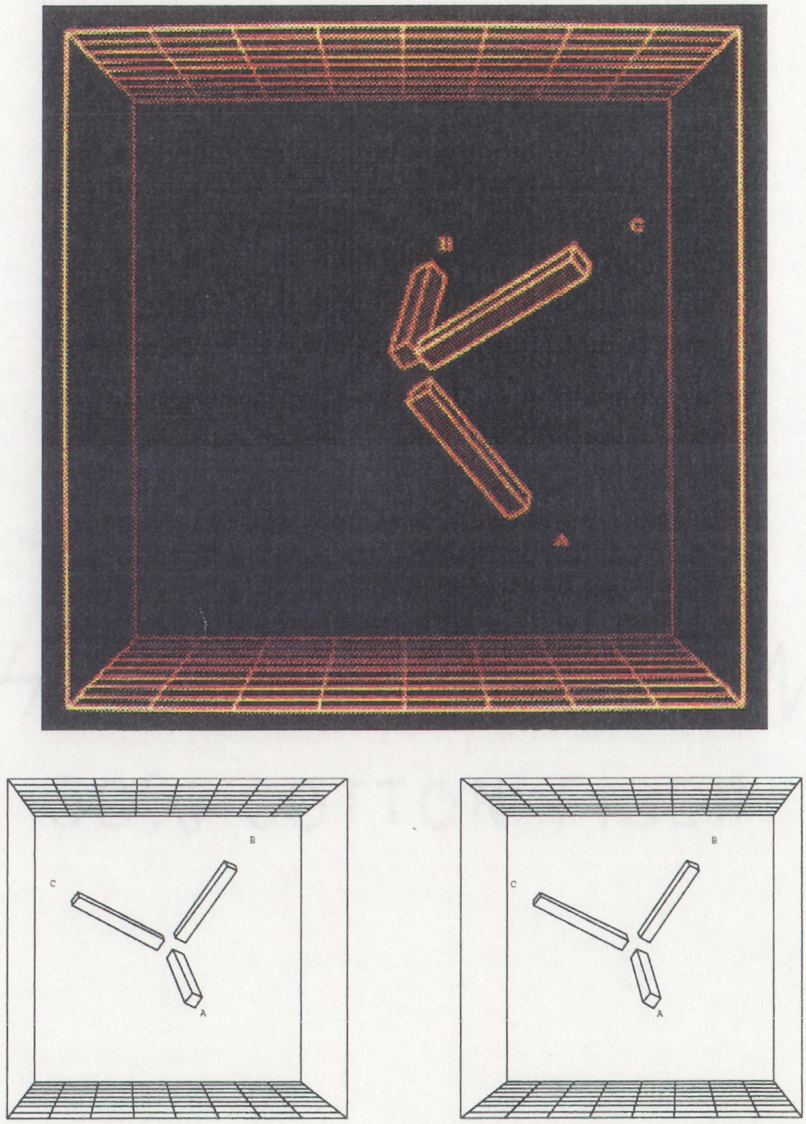


Figure 14. *Samples of experimental stimuli used in Task 5. The upper figure displays the linear perspective and proximity-luminance-covariance depth cues. The lower stereo pair displays the linear perspective and retinal disparity depth cues. To view the stereo pair, use the viewing glasses found in the sleeve on the last page of this document.*

Task 6 stimuli. Task 6 contained a rectangular solid (Figure 15) that was, in some cases, a perfect cube (all faces defined by perfect squares). In other cases, the solid contained faces which were rectangular so that the solid was not a perfect cube. For these rectangular faces, the longer side was 10–25% longer than the other side. The viewer's goal in Task 6 was to determine whether the solid was a perfect cube by analyzing all three dimensions and determining whether they are identical.

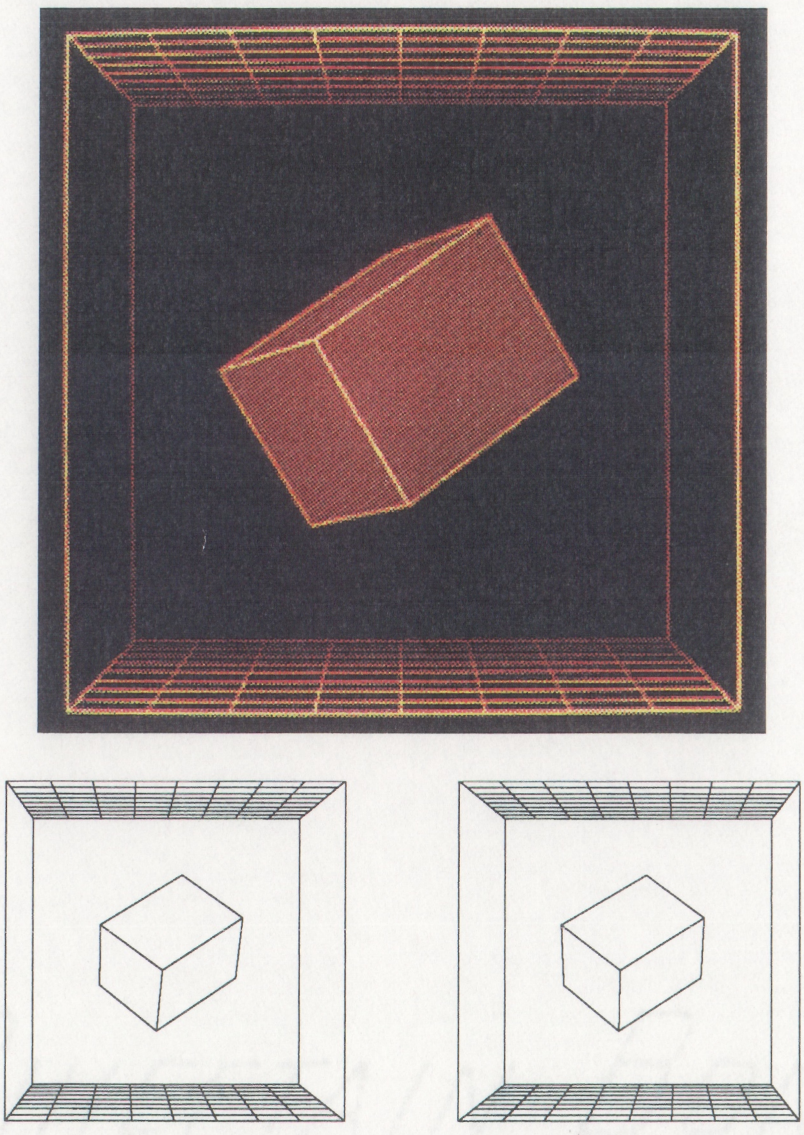


Figure 15. *Samples of experimental stimuli used in Task 6. The upper figure displays the linear perspective and proximity-luminance-covariance depth cues. The lower stereo pair displays the linear perspective and retinal disparity depth cues. To view the stereo pair, use the viewing glasses found in the sleeve on the last page of this document.*

Task 7 stimuli. The experimental stimuli for Task 7 was a single octahedron placed at random in the 3D cube (Figure 16). Over time, the octahedron appeared to move in depth; closer and farther and closer and farther in cyclical fashion. By referencing the 8-segment grids at the top and bottom of the 3D cube, the viewer's task was to judge the absolute magnitude of the depth movement.

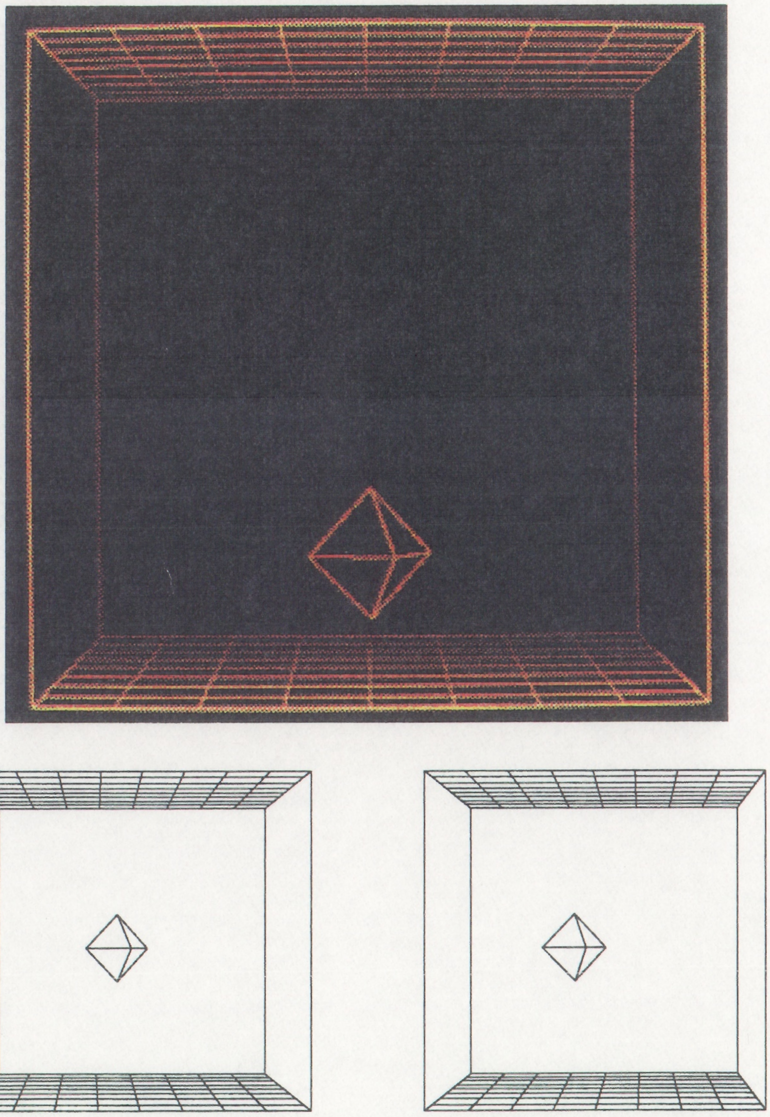


Figure 16. *Samples of experimental stimuli used in Task 7. The upper figure displays the linear perspective and proximity-luminance-covariance depth cues. The lower stereo pair displays the linear perspective and retinal disparity depth cues. To view the stereo pair, use the viewing glasses found in the sleeve on the last page of this document.*

Task 8 stimuli. In Task 8, participants viewed two octahedra that were positioned at equal "height" but different depths (Figure 17). The depth difference between the two octahedra ranged from 1 to 8 units, corresponding to the 8-segment grids at the top and bottom of the 3D cube. Using these 8-unit scales, the viewer's task was to judge the absolute depth between the objects on a scale of 1–8.

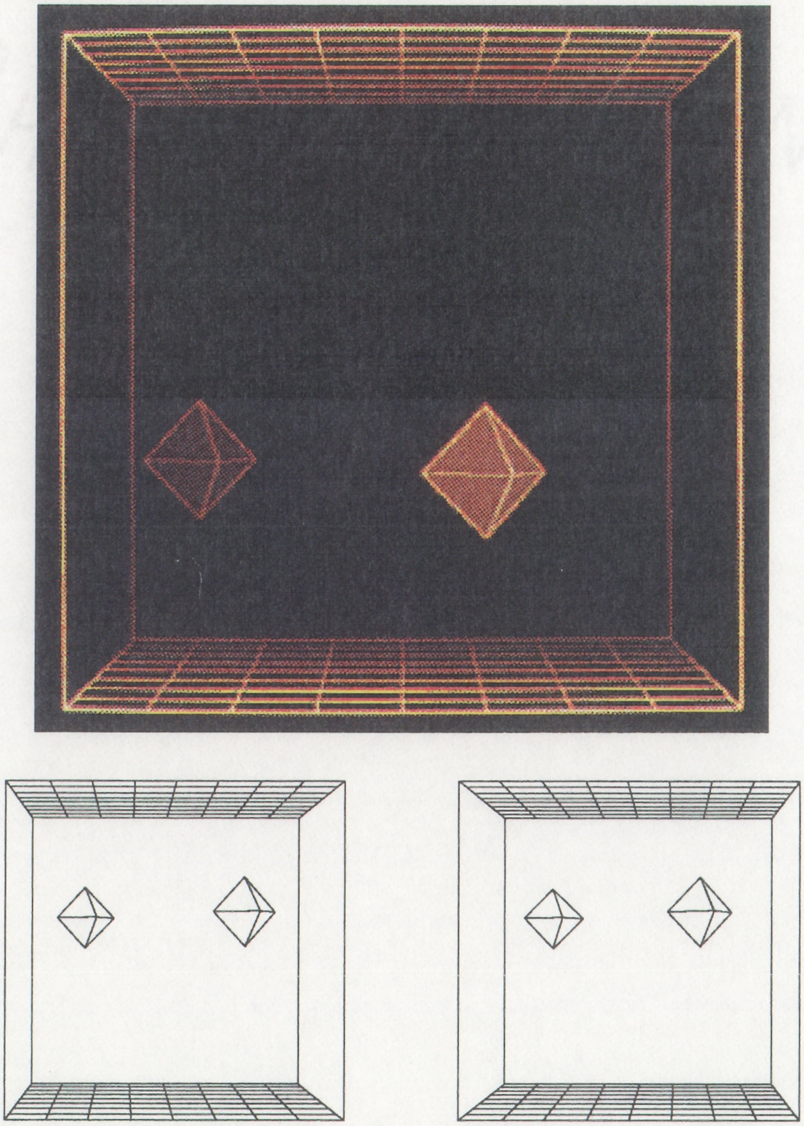


Figure 17. *Samples of experimental stimuli used in Task 8. The upper figure displays the linear perspective and proximity-luminance-covariance depth cues. The lower stereo pair displays the linear perspective and retinal disparity depth cues. To view the stereo pair, use the viewing glasses found in the sleeve on the last page of this document.*

Task 9 stimuli. In Task 9, participants viewed two octahedra that were positioned at equal “height” but different depths (Figure 18). The viewer’s task was to determine which of the octahedra appeared closer, the left or the right.

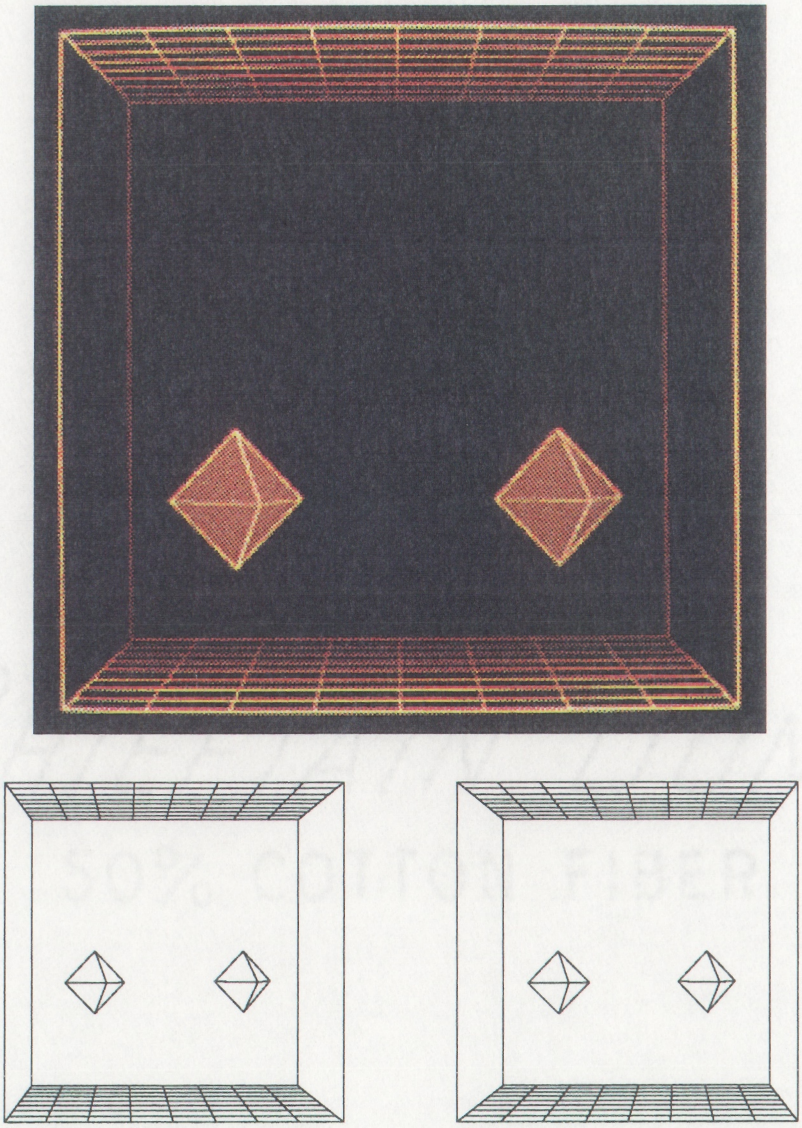


Figure 18. *Samples of experimental stimuli used in Task 9. The upper figure displays the linear perspective and proximity-luminance-covariance depth cues. The lower stereo pair displays the linear perspective and retinal disparity depth cues. To view the stereo pair, use the viewing glasses found in the sleeve on the last page of this document.*

Task 10 stimuli. In Task 10, participants viewed three octahedra that were positioned at equal "height" but different depths (Figure 19). Each of the two outermost octahedra were a different depth distance from the center octahedron. The viewer's task was to determine which of the outer two octahedra was closer to the center octahedron, the left or the right.

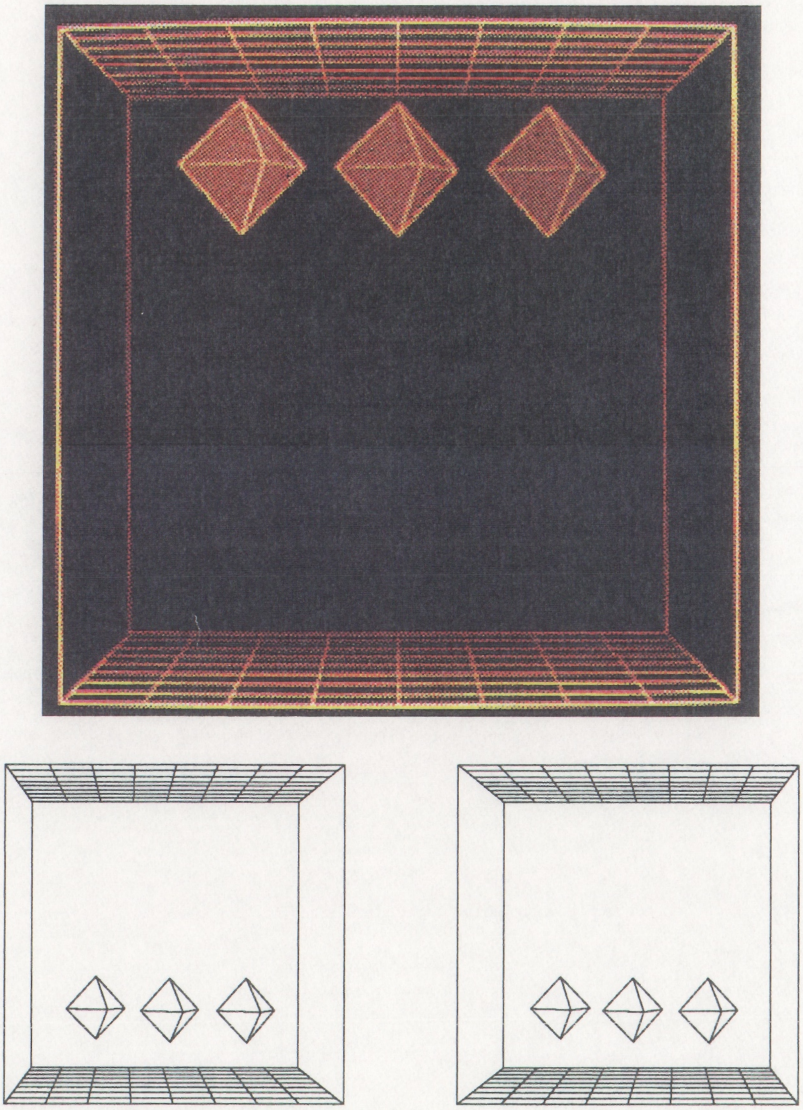


Figure 19. *Samples of experimental stimuli used in Task 10. The upper figure displays the linear perspective and proximity-luminance-covariance depth cues. The lower stereo pair displays the linear perspective and retinal disparity depth cues. To view the stereo pair, use the viewing glasses found in the sleeve on the last page of this document.*

Task 11 stimuli. In Task 11, participants viewed two stationary wire-frame objects placed side by side (Figure 20). Each of these objects resembled a contiguous grouping of cubes. These two objects appeared different from each other and there were two possible explanations for this difference. The objects were either: 1) the same shape but rotated (around the vertical axis) to appear different, or 2) two different shapes. The task was to determine whether the objects were identical except for a rotation.

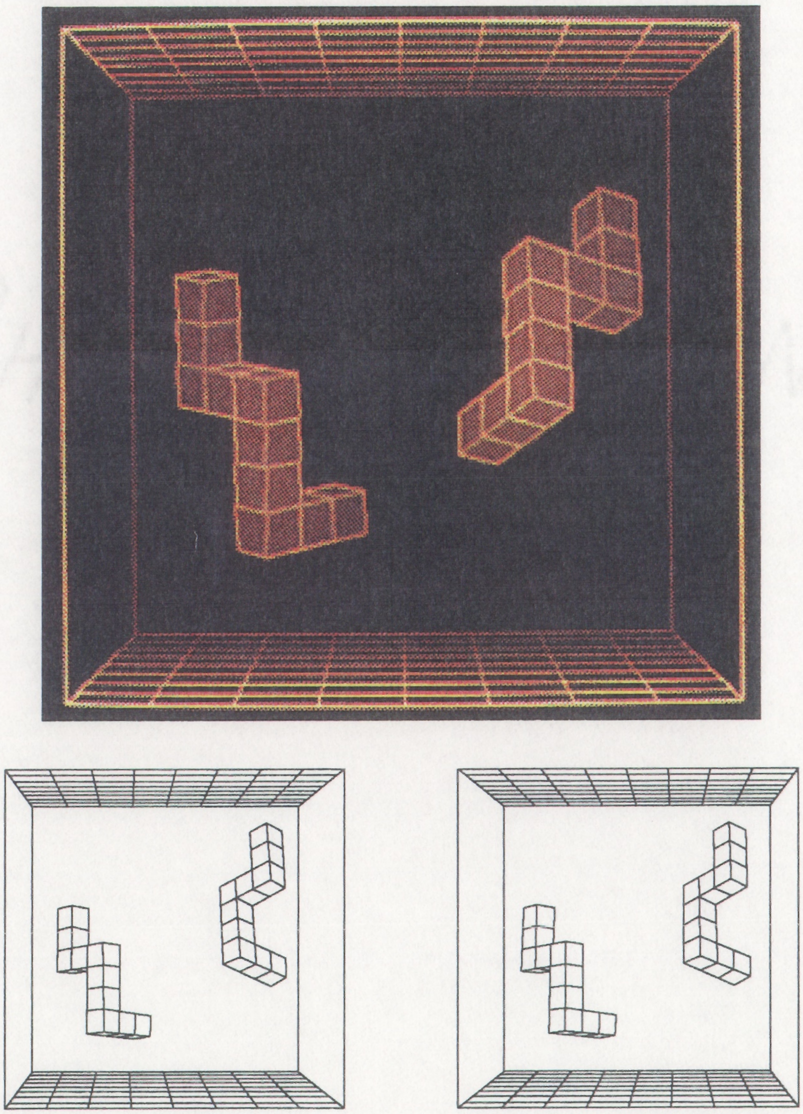


Figure 20. *Samples of experimental stimuli used in Task 11. The upper figure displays the linear perspective and proximity-luminance-covariance depth cues. The lower stereo pair displays the linear perspective and retinal disparity depth cues. To view the stereo pair, use the viewing glasses found in the sleeve on the last page of this document.*

Task 12 stimuli. In Task 12, participants viewed two point clouds, each composed of 40 dots (Figure 21). The two point clouds were viewed in sequence such that only one point cloud was seen at a time. The viewer pressed a key to switch between the two views. The pattern of the two point clouds was either identical or slightly different, and the viewer's task was to identify the similarity or dissimilarity between the two clouds.

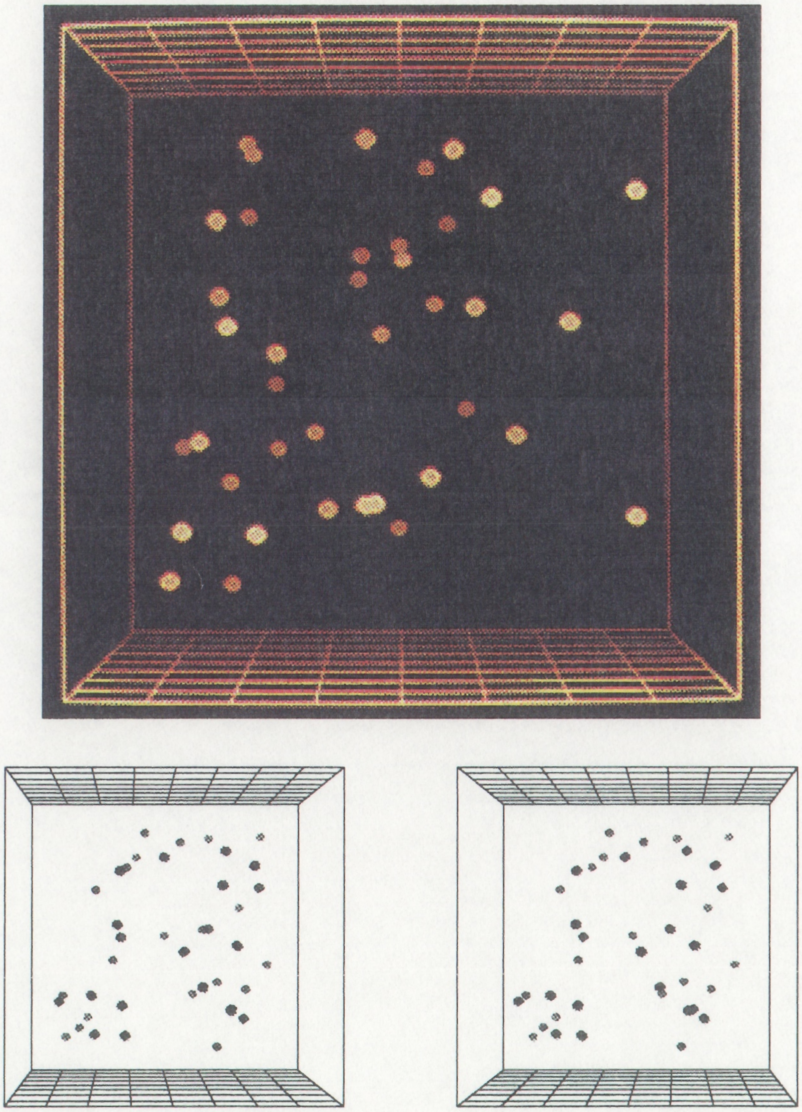


Figure 21. *Samples of experimental stimuli used in Task 12. The upper figure displays the linear perspective and proximity-luminance-covariance depth cues. The lower stereo pair displays the linear perspective and retinal disparity depth cues. To view the stereo pair, use the viewing glasses found in the sleeve on the last page of this document.*

Display luminance. The technique of proximity-luminance-covariance dictated that the luminance of all stimuli components vary as a function of depth within the virtual 3D space. Because of the limited transmissivity of the LC shutter and polarized glasses (13.5%), it was important to construct a luminance map that: 1) did not drop below the limits of mesopic viewing, and 2) maintained adequate increments to allow for depth discriminability. It was determined that a luminance increment of 14.4% percent was suitable.

The luminance output at the center of the SGS-620 is plotted as a function of the red gun value in Figure 22. The predictive equation in Figure 22 was inverted to calculate appropriate gun values that would yield equal 14.4% luminance increments over 20 depth intervals.

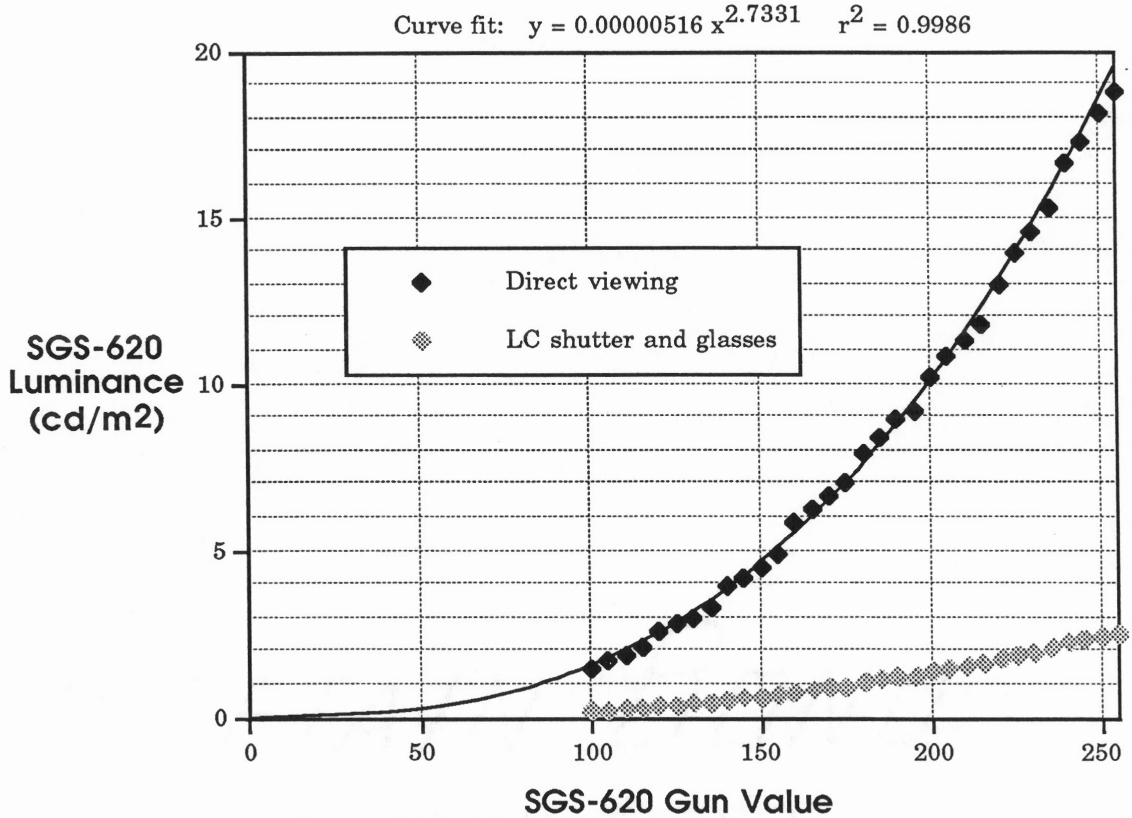


Figure 22. *SGS-620 CRT luminance as a function of red gun value. Values represent luminance at the center of the screen. Note the poor transmissivity of the LC shutter and polarized glasses (13.5%).*

The luminance values for object edges and details ranged from 1.50 to 19.33 cd/m^2. The luminance values for area fills ranged from 0.37 to 4.81 cd/m^2. After passing through the LC shutter and glasses, the resultant overall display luminance ranged from 0.05 to 2.54 cd/m^2. These levels are in the range of mesopic viewing.

Experimental Design

The experiment used a 2 x 12 within-subjects factorial design. Each participant experienced 24 unique conditions, with each experimental condition replicated four times. Replications differed only with respect to the relative position and orientation of scene elements.

The following independent variables were manipulated in this study:

Depth Cue (2 levels). Visual stimuli was presented in either a 2½D mode containing monocular cues only or a virtual 3D (stereoscopic) mode containing both monocular cues and retinal disparity cues. The following monocular cues were used in all scenes: linear perspective (relative size), proximity-luminance covariance, and interposition where appropriate.

Component Task (12 levels). Twelve tasks were used to represent the elemental components of performance in more complex 3D graphic tasks. There were two tasks for each of six task groups: (1) Object Proximity, (2) Course/Collision Prediction, (3) Dimensional Ratios, (4) Absolute Depth, (5) Relative Depth, and (6) Object Shape Recognition. In brief, these 12 tasks were as follows:

Object Proximity
1. From Reference
2. Between Pairs

Course/Collision Prediction
3. One moving object
4. Two moving objects

Dimensional Ratios
5. Axes
6. Solids

Absolute Depth
7. One moving object
8. Two stationary objects

Relative Depth
9. Between Pairs
10. From Reference

Object Shape Recognition
11. MRT
12. Point Clouds

Experimental trials were blocked hierarchically by the six task groups and the two tasks within each group. That is, each participant completed all trials for a given task group before proceeding to the next task group. Within each task group, participants completed all trials for a single task before proceeding onto the next task. This order of presentation was used because it reduced the amount of confusion from switching between tasks. The presentation order of task groups was partially counterbalanced in Latin Square fashion, using four Latin Squares of six participants each. Within this framework, the order of presentation of tasks and depth cue was fully counterbalanced.

The following three dependent variables were used to measure both objective and subjective task performance:

Viewing time. Viewing time was the time required to complete a task, measured from initial presentation of the stimulus to the time that the stimulus was removed from the participant's view. Not included in this measure was the time during which the participant entered his or her response. For each experimental cell, viewing time was recorded as the mean viewing time across the four replications.

Accuracy. Accuracy was recorded as the percent correct responses across the four replications within each experimental condition.

Cognitive workload. At the end of each experimental condition, participants reported the level of cognitive workload according to the following 9-point scale:

9 VERY HIGH Mental Effort
8
7 HIGH Mental Effort
6
5 AVERAGE Mental Effort
4
3 LOW Mental Effort
2
1 VERY LOW Mental Effort

Participants were exposed to the full range of experimental conditions during the training portion of the experiment. Based on that experience, participants were instructed to anchor the upper end of the scale (VERY HIGH Mental Effort) to the experimental condition perceived by them to be most difficult. Similarly, participants were instructed to anchor the lower end of the scale (VERY LOW Mental Effort) to the experimental condition perceived by them to be easiest. This way, all participants were expected to use the full range of the scale to describe the subjective workload associated with the various experimental conditions. A copy of the instructions regarding mental effort is located in Appendix A.

Procedures

Participant screening. The screening procedure took place on a volunteer basis and consisted of two parts: a test for acuity and phoria, and a test for stereo acuity/deficiency. Informed consent to the screening procedure and the actual experiment was required before participants were allowed to undergo screening.

All participants were screened with a Bausch & Lomb Master Ortho-Rater for near and far visual acuity as well as vertical and lateral phoria. The criteria for the visual acuity test was a minimum monocular Snellen acuity (both near and far) of 20/22 in each eye, and phoria scores approximately within the 70th percentile. Since the polarized lenses used in conjunction with the SGS-620 may be incorporated with other lenses, participants were allowed to wear corrective lenses, either contact lenses or spectacles.

Phoria scores were included in the screening procedure since phoria assesses the ability to fuse independent retinal images (an important component of viewing stereo images) and therefore phoria deficiencies are most likely correlated with stereo deficiencies.

A stereo screening procedure developed by Reinhart (1990) was used to test for stereoscopic acuity and stereoscopic deficiency. Both stages of this procedure were presented on the Tektronix SGS-620 display system using random dot stereograms (RDS). RDSs are useful in assessing stereoscopic acuity since the perception of depth information depends entirely on retinal disparity information. A stereo-blind viewer will not be able to identify the area in which disparity has been applied; the frame will appear as a random dot pattern and nothing more.

Two demonstration RDSs were used to familiarize participants with the screening technique. These RDSs presented planar squares at 10 arcminutes of disparity, one crossed and the other uncrossed. Participants who could not identify these squares were regarded as stereo-blind and, therefore, excused from further participation in the study.

In the second phase of stereo screening, participants viewed two sets of RDSs, each containing 12 images at various values of crossed or uncrossed disparities ranging from 2 to 20 arcminutes. Participants were required to identify the orientation of a virtual Landolt ring imbedded in the stereogram as well as the direction of disparity of that ring, that is, whether the ring appeared to be in front of or behind the rest of the field. If the participant was incorrect with regard to either the direction of disparity or the ring's orientation, then the entire response was considered incorrect. Participants were allowed two incorrect responses for each set of 12 images.

Training. Training consisted of an instructional period in which both experimenter and participant progressed through a series of written instructions (Appendix A). Participants were instructed on all relevant aspects of the experiment including use of the keyboard, the nature of the tasks, and what to expect as the experiment progressed. Once the participant confirmed understanding of the overall procedure, he or she progressed through the series of tasks, receiving task-specific instructions prior to each task (Appendix B). The participants were introduced to the experimental conditions in the same order as for the actual experiment. On-screen feedback was provided during the training exercises to promote rapid learning. The training trials were self-paced and timed automatically to avoid inconsistent interactions between the experimenter and the participant.

Experimental trials. Once training was completed, participants received a brief set of instructions (Appendix C), then proceeded immediately with the experimental trials. Unlike the training session, however, no feedback was provided during experimental trials.

Prior to trials within each experimental cell, participants were provided with on-screen information regarding the nature of the upcoming trials. Within each experimental cell, participants received one practice trial; the performance data from this practice trial were not recorded. The practice trial served to refresh each participant's memory, especially in later trials when sub-task objectives may have been forgotten. Subsequent to the practice trial, experimental trials proceeded and viewing time and accuracy measures were collected accordingly. Participants were instructed to proceed with the tasks as quickly as possible without sacrificing accuracy. Trials progressed in the order described above and, as mentioned earlier, four replications were completed for each experimental condition. Immediately subsequent to each experimental condition, participants reported their level of subjective mental workload.

Participants continued trials until all experimental cells were completed. The entire procedure lasted between 1½ and 2 hours.

Analysis

Within each experimental cell, all data were averaged across the four replications. Separate Analysis of Variance (ANOVA) procedures were then used to analyze the effects of Depth Cue and Component Task on the dependent measures of accuracy, viewing time, and cognitive workload. To compensate for biased F-values due to heterogeneity of co-variance, degrees of freedom were adjusted using the value ε, the value of which was calculated using the computer application SuperANOVA (Abacus Concepts, Inc.), applying equation 11 below (from Box, 1954). Simple-effects F-tests then were used to evaluate any significant interactions between Depth Cue and Component Task. Newman Keuls post-hoc tests were applied to evaluate any significant main effects of Component Task. The level of significance applied to all tests was 0.05.

$$\varepsilon = \frac{\left(\sum_{i=1}^{k} \lambda_i\right)^2}{(k-1)\left(\sum_{i=1}^{k} \lambda_i^2\right)} \tag{11}$$

where

k = the degrees of freedom of the hypothesis matrix, and

λ = eigenvalues of the correlation matrix.

RESULTS

The 12 component tasks developed for this experiment were designed to represent a breakdown of more complex 3D tasks. However, if the performance measures associated with these 12 tasks are restricted in some way, then the generalizability of the resulting data is likewise restricted. Therefore, it is useful to examine the distribution of performance measures. This analysis is not rigorous. Rather, it is provided only so that the reader may examine the distributions as necessary.

Distributions of Dependent Measures

Viewing time. The distribution of viewing times over all 12 component tasks is shown in Figure 23. The median viewing time was 4.4 seconds (standard deviation = 5.3). Fewer than 10 individual replications resulted in viewing times longer than 30 seconds.

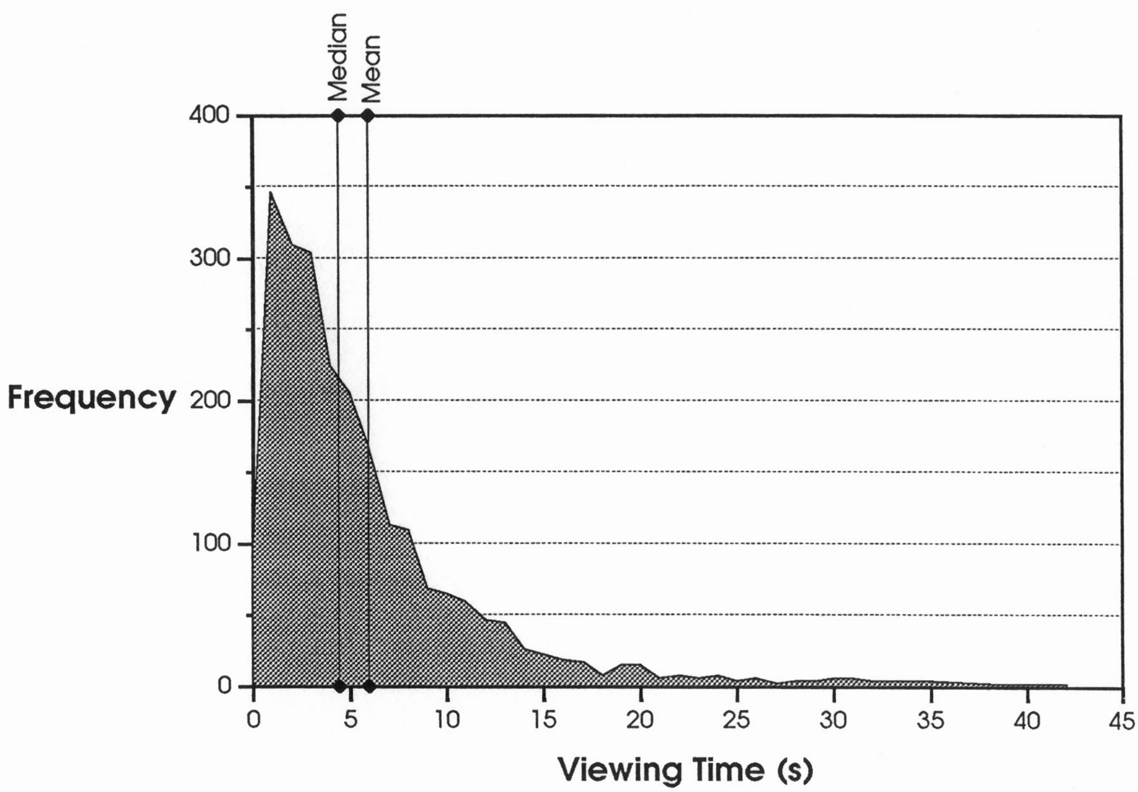


Figure 23. *Distribution of viewing times across all 12 component tasks.*

Accuracy. The distribution of accuracy scores is shown in Figure 24. The values in Figure 24 represent trial accuracy averaged over the four replications in each experimental cell, yielding the five possible categories (0%, 25%, 50%, 75%, and 100% correct). The mean accuracy was 84% (standard deviation = 21%). In only one experimental cell did a participant respond incorrectly on all four replications.

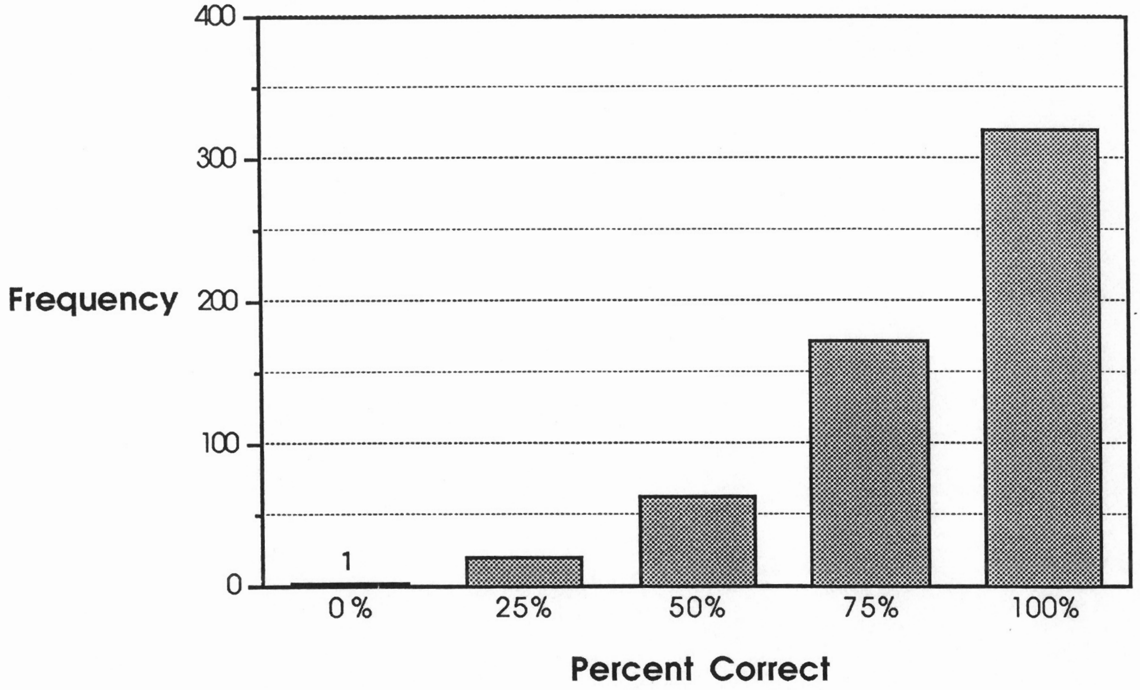


Figure 24. *Distribution of accuracy (percent correct) across all 12 component tasks.*

Mental effort rating. The distribution of mental effort ratings is shown in Figure 25. The values in Figure 25 represent mental effort ratings averaged over the four replications in each experimental cell. The overall mean mental effort rating was 4.58 (standard deviation = 1.74).

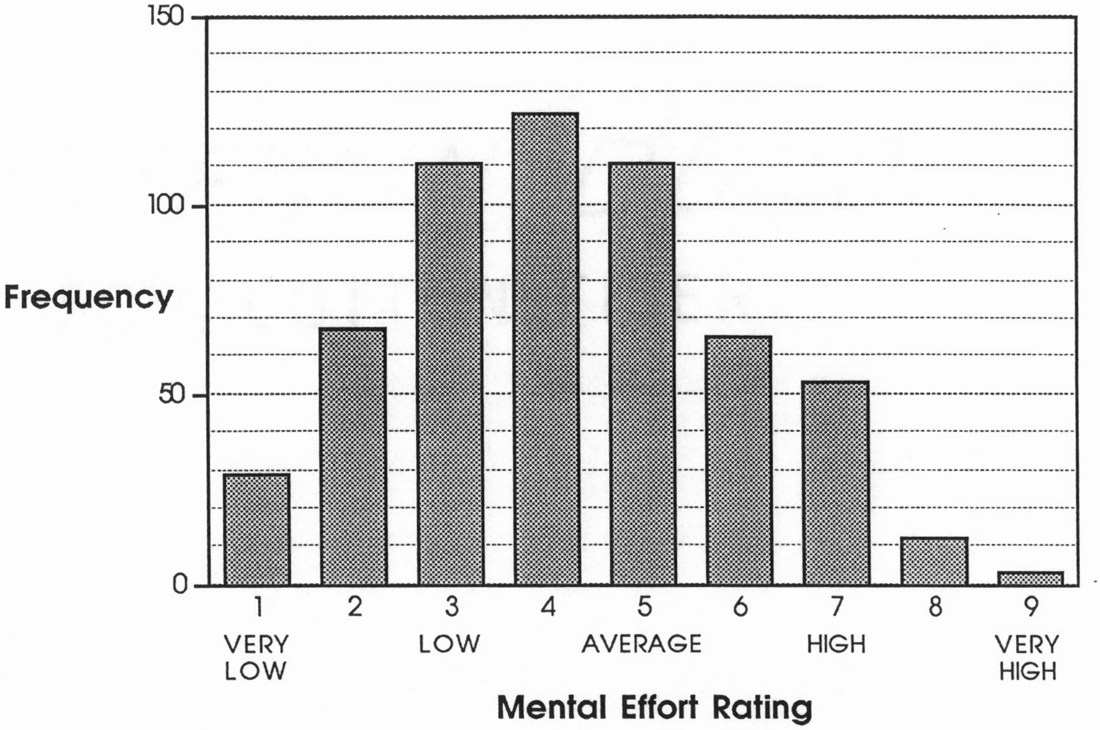


Figure 25. *Distribution of mental effort ratings across all 12 component tasks.*

Viewing Time

The ANOVA results for viewing time are presented in Table 1. The interaction of Depth Cue and Component Task is significant (p_{ε} = 0.0038). This interaction is presented graphically in Figure 26. The main effect of Component Task also is significant ($p_{\varepsilon} < 0.0001$). The mean viewing times associated with each of the 12 Component Tasks and the results of a Newman Keuls post-hoc analysis are presented in Figure 27. The main effect of Depth Cue is not significant.

TABLE 1

ANOVA Summary Table for Viewing Time

Source of Variation	*df*	MS	F	*p*	ε	p_{ε}*
Between-Subjects						
Subject	23	108.9384				
Within-Subjects						
Depth Cue	1	11.4963	2.4620	0.1303		
Depth Cue * Subject	23	4.6695				
Component Task	11	274.8719	22.0952		0.4315	<0.0001
Component Task * Subject	253	12.4403				
Depth Cue * Component Task	11	13.4306	3.3871		0.5468	0.0038
Depth Cue * Comp. Task * Subject	253	3.9652				
Total	575					

* Adjusted probability based on ε.

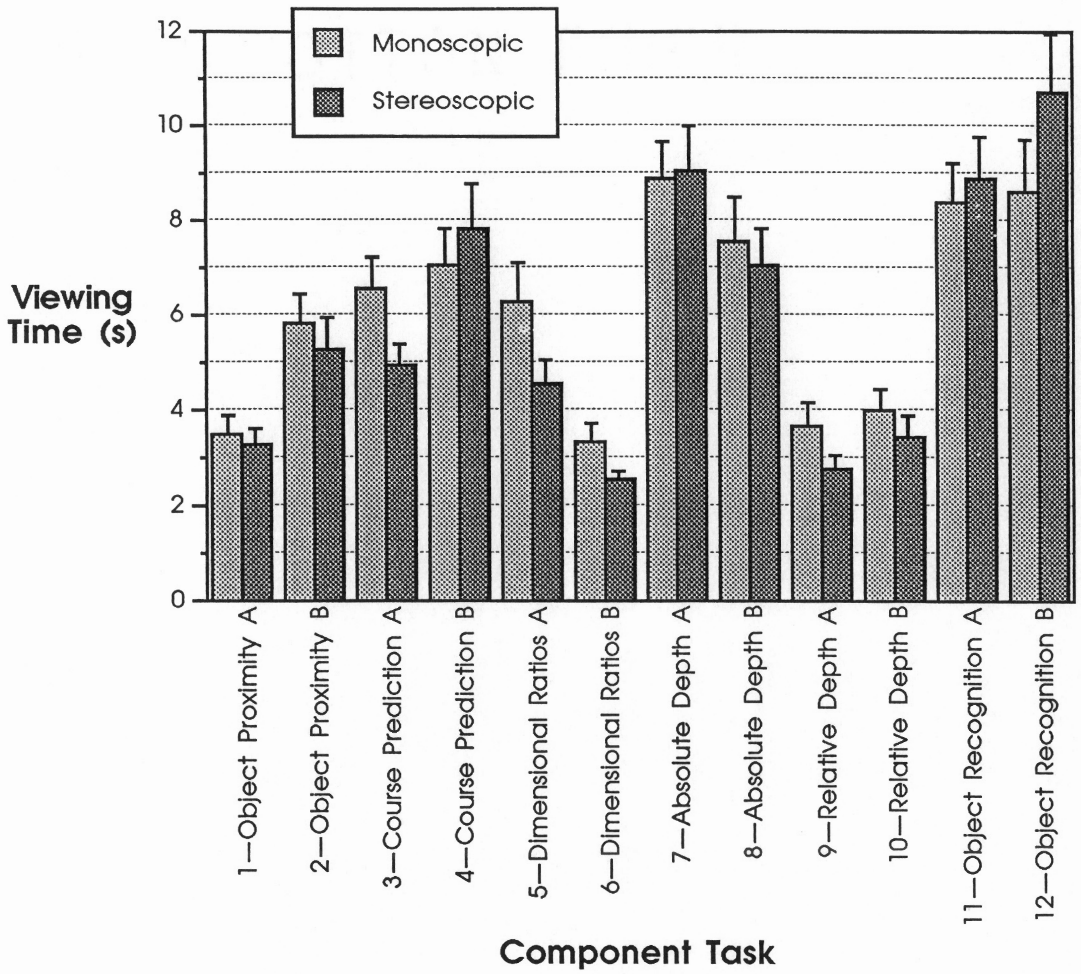


Figure 26. *Interactive effect of Depth Cue and Component Task on viewing time. Error bars represent +1 standard error of the mean.*

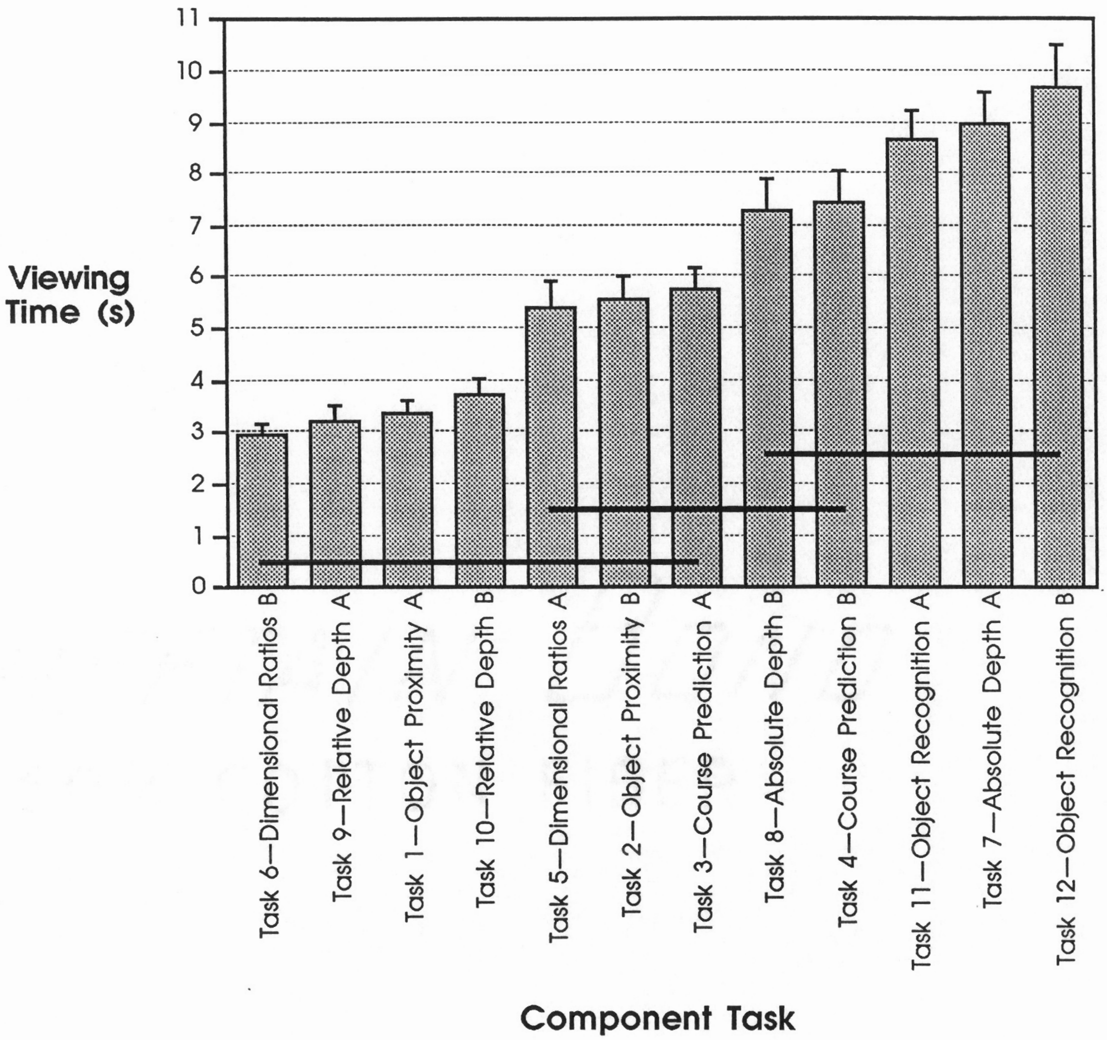


Figure 27. *Mean viewing times for each of the 12 component tasks. Error bars represent +1 standard error of the mean. The horizontal bars represent the results of a Newman-Keuls post-hoc analysis—means which share a horizontal bar are not significantly different at* $p \leq 0.05$.

The results of the F-test of the simple effect of Depth Cue on viewing time for each level of Component Task are presented in Table 2. Note that adjusted *p* values are not necessary as there are only two levels of Depth Cue.

As shown in Table 2, the simple effect of Depth Cue on viewing time is significant within the following three Component Tasks: Task 3—Course Prediction A, Task 5—Dimensional Ratios A, and Task 12—Object Recognition B. Viewing times were improved (decreased) in all but one of these three tasks, as Task 12—Object Recognition B yielded a significant *increase* in viewing time as a result of stereoscopic display. Of particular interest is the large number of scene objects displayed in Task 12. Whereas the other 11 tasks displayed 9 or fewer objects, Task 12 displayed 40 objects within the 3D workspace (Figure 21).

The results specific to Task 1 replicate the results of Miller (1991), who found that the addition of retinal disparity depth cues do not significantly affect viewing time for a similar object proximity task. However, the results specific to Task 3 are *contrary* to the findings of Miller. For a similar course prediction task, Miller found that viewing times do not change as a result of adding stereoscopic depth cues.

The results specific to Task 9 agree with the findings of Reinhart (1990), who found no stereoscopic viewing time advantage for a similar relative depth task.

TABLE 2

Simple Effect of Depth Cue on Viewing Time within each Level of Component Task.

Source of Variation	*df*	MS	F	*p*
Task 1—Object Proximity A				
Depth Cue	1	0.8328	0.2100	> 0.25
Depth Cue * Subject	23	3.9652		
Task 2—Object Proximity B				
Depth Cue	1	4.3185	1.0891	> 0.25
Depth Cue * Subject	23	3.9652		
Task 3—Course Prediction A				
Depth Cue	1	32.0501	8.0828	**< 0.01**
Depth Cue * Subject	23	3.9652		
Task 4—Course Prediction B				
Depth Cue	1	7.7261	1.9485	> 0.1
Depth Cue * Subject	23	3.9652		
Task 5—Dimensional Ratios A				
Depth Cue	1	33.1627	8.3634	**< 0.01**
Depth Cue * Subject	23	3.9652		
Task 6—Dimensional Ratios B				
Depth Cue	1	7.8328	1.9754	> 0.1
Depth Cue * Subject	23	3.9652		
Task 7—Absolute Depth A				
Depth Cue	1	0.2844	0.0717	> 0.25
Depth Cue * Subject	23	3.9652		
Task 8—Absolute Depth B				
Depth Cue	1	3.1982	0.8066	> 0.25
Depth Cue * Subject	23	3.9652		
Task 9—Relative Depth A				
Depth Cue	1	10.2328	2.5807	> 0.1
Depth Cue * Subject	23	3.9652		
Task 10—Relative Depth B				
Depth Cue	1	3.4080	0.8595	> 0.25
Depth Cue * Subject	23	3.9652		
Task 11—Object Recognition A				
Depth Cue	1	2.8191	0.7110	> 0.25
Depth Cue * Subject	23	3.9652		
Task 12—Object Recognition B				
Depth Cue	1	53.3672	13.4589	**< 0.01**
Depth Cue * Subject	23	3.9652		

Accuracy

The ANOVA results for accuracy are presented in Table 3. The interaction of Depth Cue and Component Task is significant ($p_\varepsilon = 0.0163$). This interaction is presented graphically in Figure 28. The main effect of Component Task is also significant ($p_\varepsilon < 0.0001$). The mean accuracy levels associated with each of the 12 Component Tasks and the results of a Newman-Keuls post-hoc analysis are presented in Figure 29. The main effect of Depth Cue is significant ($p = 0.0083$). The mean accuracy levels associated with the two levels of Depth Cue are presented in Figure 30.

TABLE 3

ANOVA Summary Table for Accuracy

Source of Variation	*df*	MS	F	*p*	ε	p_ε
Between-Subjects						
Subject	23	4.5743				
Within-Subjects						
Depth Cue	1	3906.2500	8.3333	0.0083		
Depth Cue * Subject	23	468.7500				
Component Task	11	4488.6364	13.3439		0.5772	<0.0001
Component Task * Subject	253	336.3801				
Depth Cue * Component Task	11	823.8636	2.7562		0.5204	0.0163
Depth Cue * Comp. Task * Subject	253	298.9130				
Total	575					

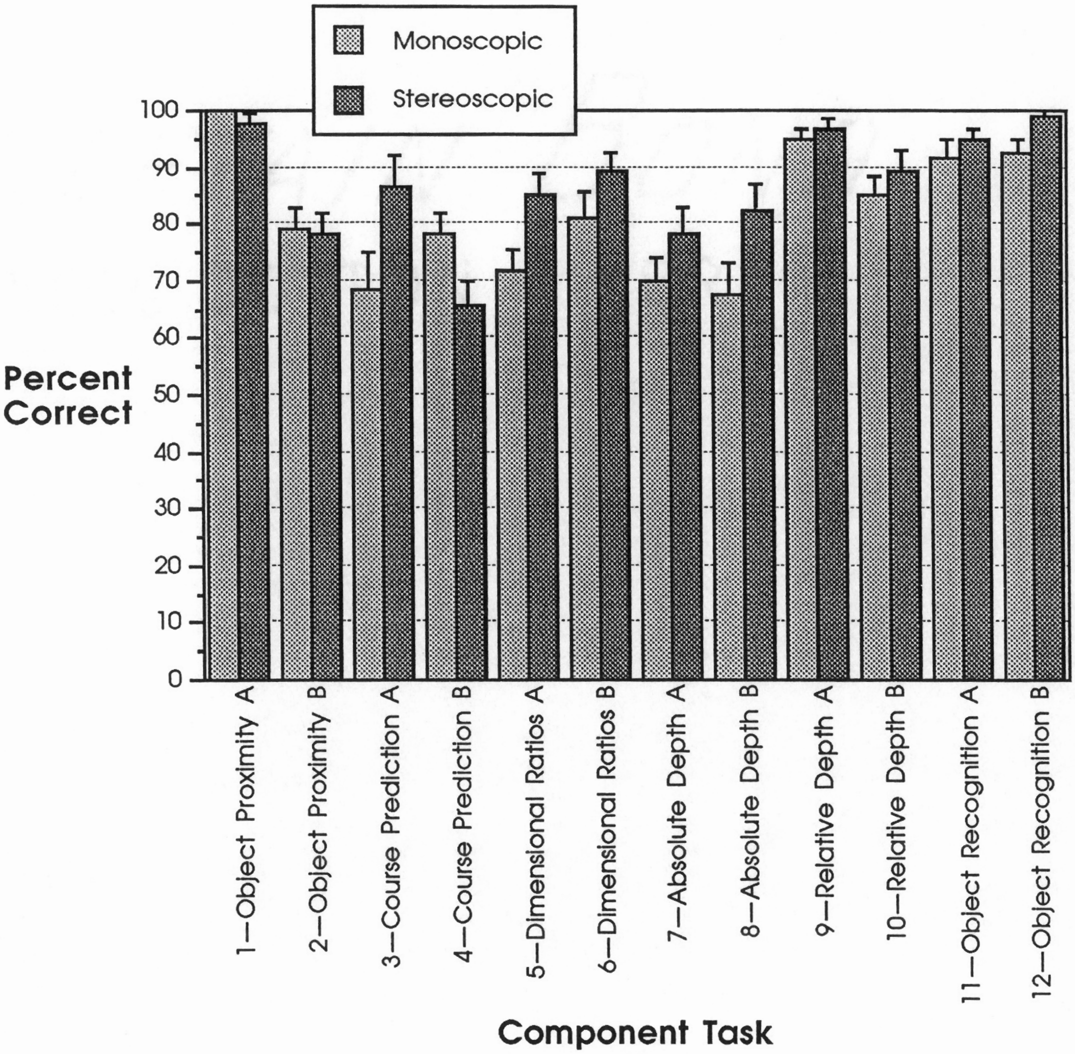


Figure 28. *Interactive effect of Depth Cue and Component Task on accuracy. Error bars represent +1 standard error of the mean.*

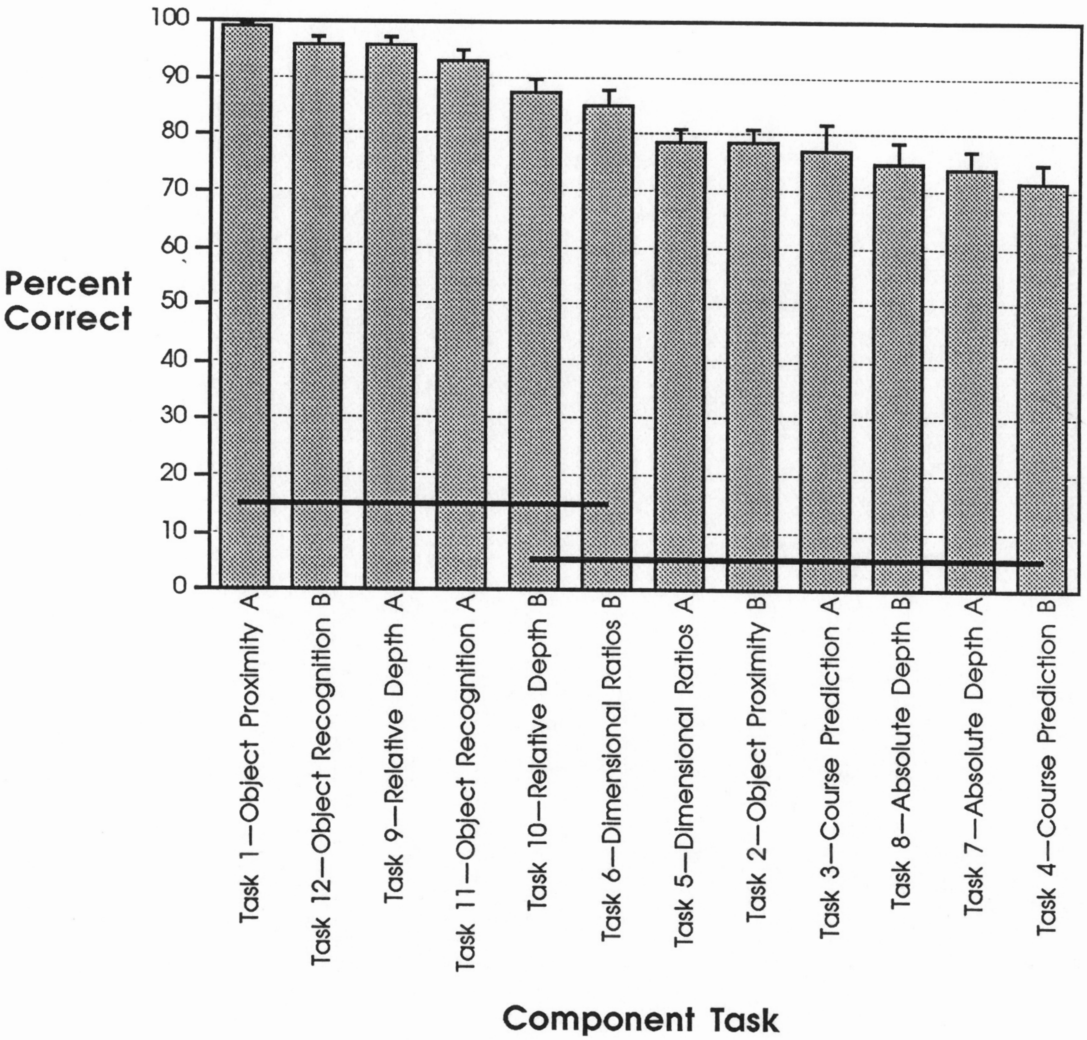


Figure 29. *Mean accuracy (percent correct) scores over the 12 component tasks. Error bars represent +1 standard error of the mean. The horizontal bars represent the results of a Newman-Keuls post-hoc analysis—means which share a horizontal bar are not significantly different at* $p \leq 0.05$.

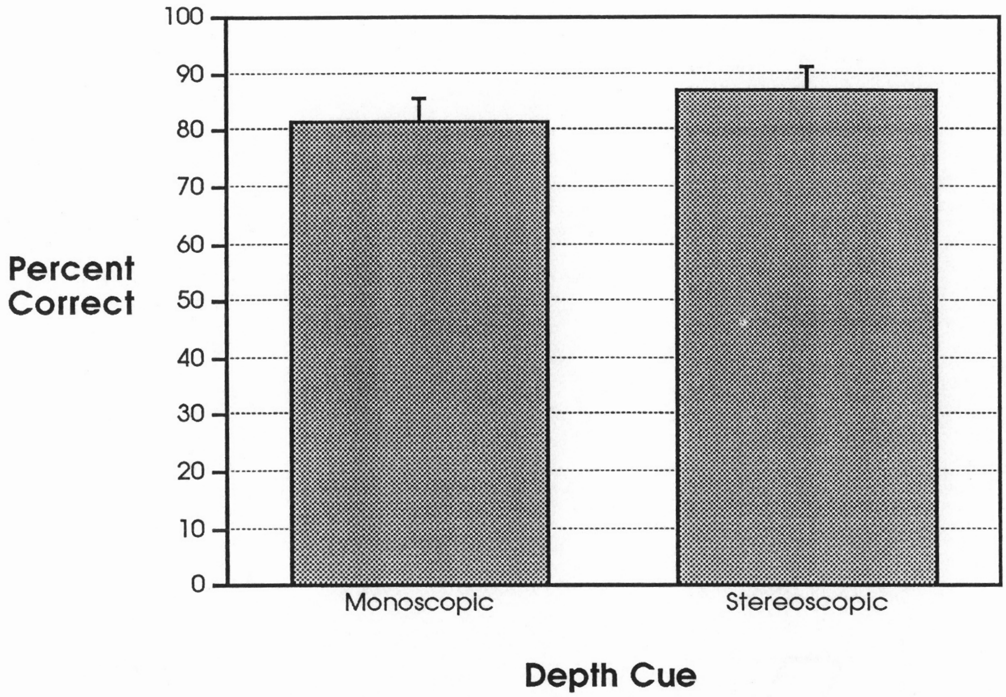


Figure 30. *Main effect of Depth Cue on accuracy (percent correct). Error bars represent +1 standard error of the mean.*

The results of the F-test of the simple effect of Depth Cue on accuracy for each level of Component Task are presented in Table 4. The simple effect of Depth Cue on accuracy is significant within the following four Component Tasks: Task 3—Course Prediction A, Task 4—Course Prediction B, Task 5—Dimensional Ratios A, and Task 8—Absolute Depth B. Accuracy was improved in all but one of these four tasks, as Task 4—Course Prediction B yielded a significant decrement in accuracy as a result of stereoscopic display. Task 4 is notable in that it is the only task which includes more than one moving object.

TABLE 4

Simple Effect of Depth Cue on Accuracy within each Level of Component Task.

Source of Variation	*df*	MS	F	*p*
Task 1—Object Proximity A				
Depth Cue	1	52.0833	0.1742	> 0.25
Depth Cue * Subject	23	298.9130		
Task 2—Object Proximity B				
Depth Cue	1	13.0208	0.0436	> 0.25
Depth Cue * Subject	23	298.9130		
Task 3—Course Prediction A				
Depth Cue	1	3763.0208	12.5890	**< 0.01**
Depth Cue * Subject	23	298.9130		
Task 4—Course Prediction B				
Depth Cue	1	1875.0000	6.2727	**< 0.025**
Depth Cue * Subject	23	298.9130		
Task 5—Dimensional Ratios A				
Depth Cue	1	2200.5208	7.3617	**< 0.025**
Depth Cue * Subject	23	298.9130		
Task 6—Dimensional Ratios B				
Depth Cue	1	833.3333	2.7879	> 0.1
Depth Cue * Subject	23	298.9130		
Task 7—Absolute Depth A				
Depth Cue	1	833.3333	2.7879	> 0.1
Depth Cue * Subject	23	298.9130		
Task 8—Absolute Depth B				
Depth Cue	1	2552.0833	8.5379	**< 0.01**
Depth Cue * Subject	23	298.9130		
Task 9—Relative Depth A				
Depth Cue	1	52.0833	0.1742	> 0.25
Depth Cue * Subject	23	298.9130		
Task 10—Relative Depth B				
Depth Cue	1	208.3333	0.6970	> 0.25
Depth Cue * Subject	23	298.9130		
Task 11—Object Recognition A				
Depth Cue	1	117.1875	0.3920	> 0.25
Depth Cue * Subject	23	298.9130		
Task 12—Object Recognition B				
Depth Cue	1	468.7500	1.5682	> 0.1
Depth Cue * Subject	23	298.9130		

The results specific to Task 1 replicate the results of Miller (1991), who found no significant stereoscopic accuracy advantage for a similar object proximity task. The results specific to Task 3 also are in agreement with the findings of Miller. For a similar course prediction task, Miller found that accuracy levels increase significantly as a result of adding stereoscopic depth cues.

The results specific to Task 9 are consistent with the findings of Reinhart (1990), who found no stereoscopic accuracy advantage for a similar relative depth task. Reinhart also observed a high overall accuracy rate for this task (98.5 % correct).

Mental Effort Rating

The ANOVA results for mental effort rating are presented in Table 5. The interaction of Depth Cue and Component Task is significant ($p_\varepsilon = 0.0119$). This interaction is presented graphically in Figure 31. The main effect of Component Task is also significant ($p_\varepsilon < 0.0001$). The mean mental effort ratings associated with each of the 12 Component Tasks and the results of a Newman-Keuls post-hoc analysis are presented in Figure 32. The main effect of Depth Cue is significant ($p = 0.0006$). The mean mental effort ratings associated with each Depth Cue are presented in Figure 33.

TABLE 5

ANOVA Summary Table for Mental Effort Rating

Source of Variation	*df*	MS	F	*p*	ε	p_ε
Between-Subjects						
Subject	23	14.9268				
Within-Subjects						
Depth Cue	1	7.7354	15.5672	0.0006		
Depth Cue * Subject	23	0.4969				
Component Task	11	39.4262	12.3489		0.4819	<0.0001
Component Task * Subject	253	3.1927				
Depth Cue * Component Task	11	1.3199	2.8578		0.5443	0.0119
Depth Cue * Comp. Task * Subject	253	0.4619				
Total	575					

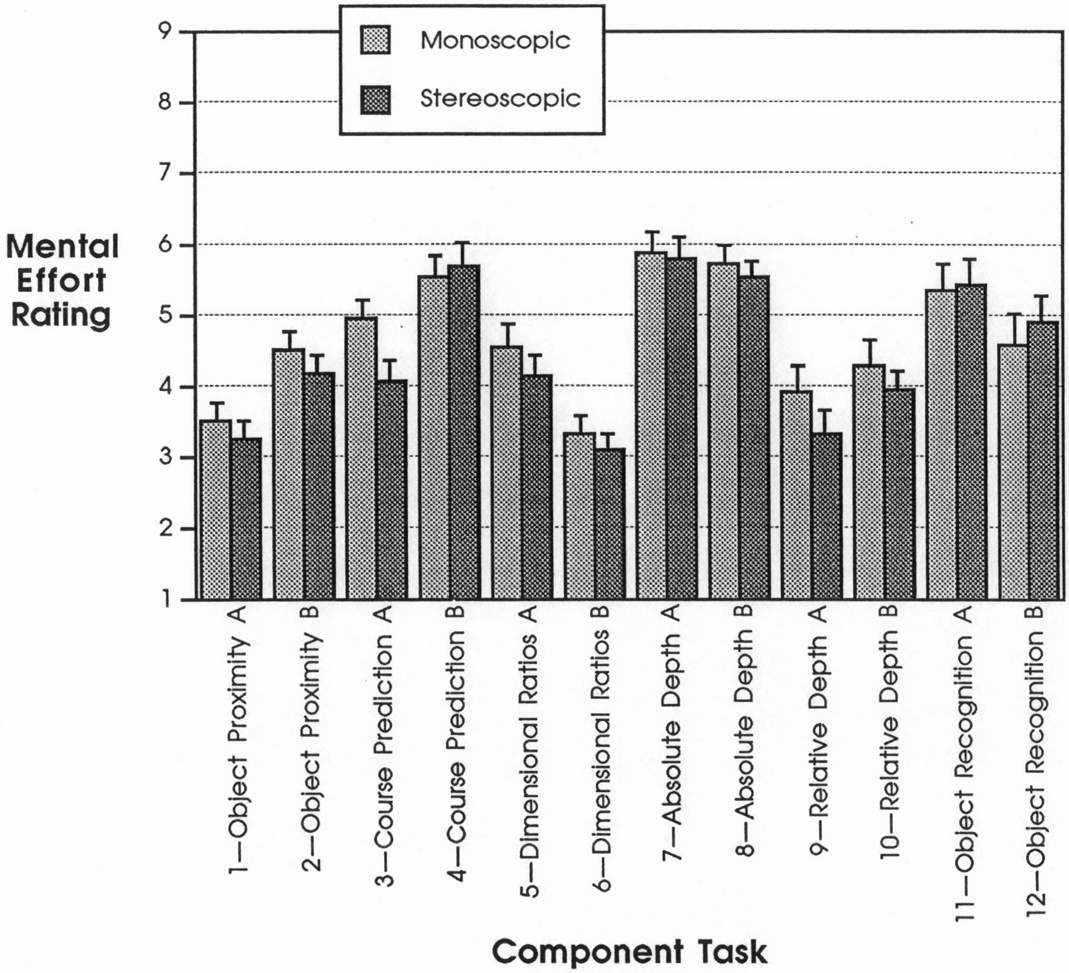


Figure 31. *Interactive effect of Depth Cue and Component Task on mental effort rating. Error bars represent +1 standard error of the mean.*

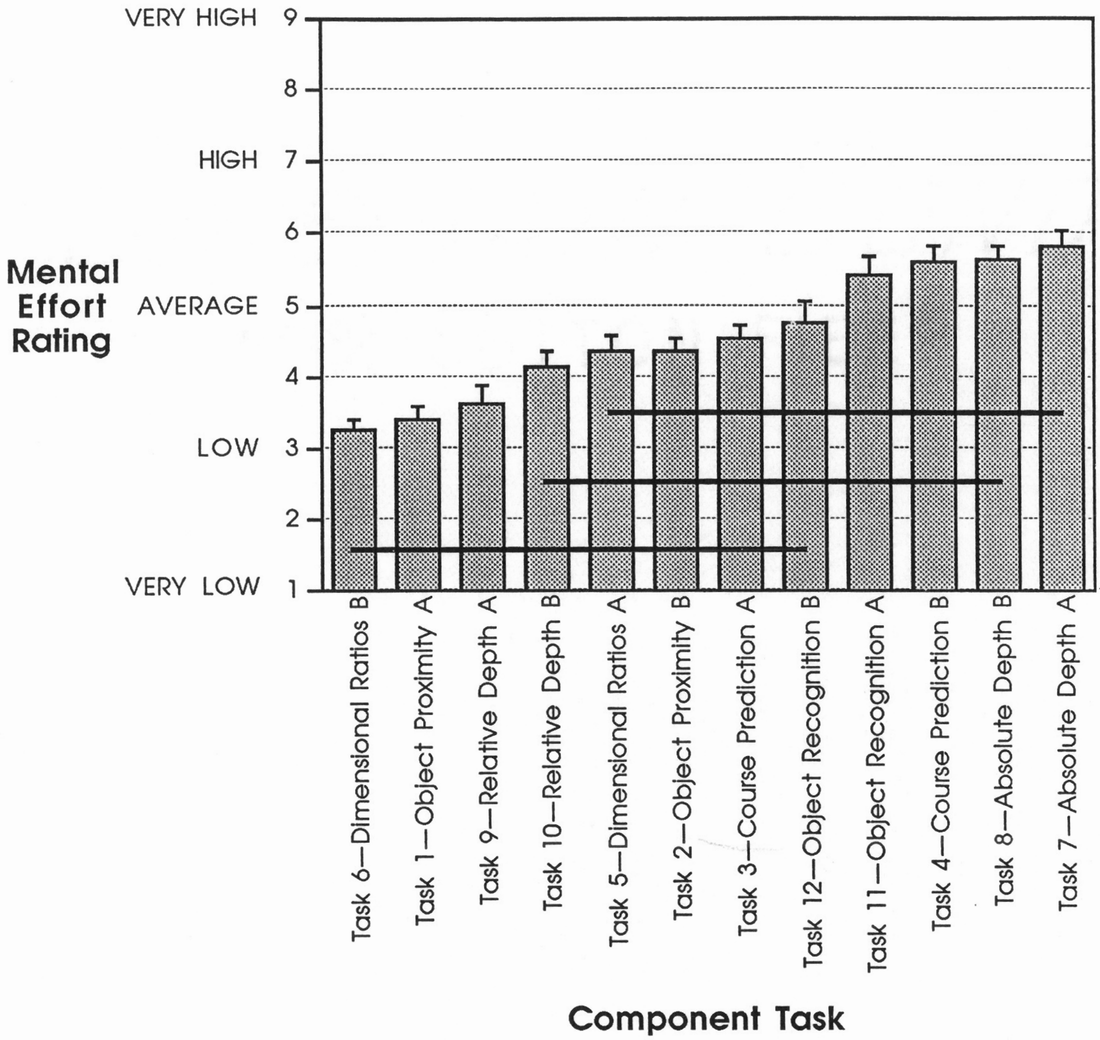


Figure 32. *Mean mental effort ratings for the 12 component tasks. Error bars represent +1 standard error of the mean. The horizontal bars represent the results of a Newman-Keuls post-hoc analysis—means which share a horizontal bar are not significantly different at $p \leq 0.05$.*

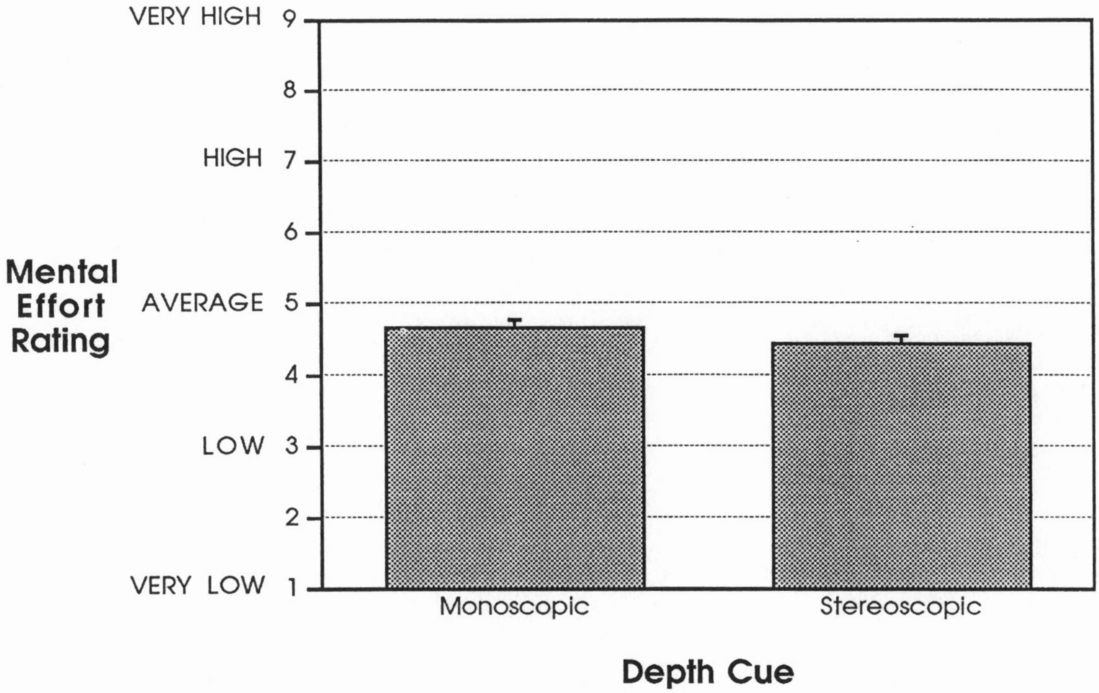


Figure 33. *Main effect of Depth Cue on mental effort rating. Error bars represent +1 standard error of the mean.*

The results of the F-test of the simple effect of Depth Cue on mental effort rating for each level of Component Task are presented in Table 6. The simple effect of Depth Cue is significant within three of the 12 Component Tasks: Task 3—Course Prediction A, Task 5—Dimensional Ratios A, and Task 9—Relative Depth A.

TABLE 6

Simple Effect of Depth Cue on Mental Effort Rating within each Level of Component Task. Note that individual error terms are used to calculate F-values.

Source of Variation	*df*	MS	F	*p*
Task 1—Object Proximity A				
Depth Cue	1	0.9492	2.0549	> 0.25
Depth Cue * Subject	23	0.4619		
Task 2—Object Proximity B				
Depth Cue	1	1.4180	3.0699	> 0.05
Depth Cue * Subject	23	0.4619		
Task 3—Course Prediction A				
Depth Cue	1	9.1875	19.8907	**< 0.001**
Depth Cue * Subject	23	0.4619		
Task 4—Course Prediction B				
Depth Cue	1	0.3333	0.7216	> 0.25
Depth Cue * Subject	23	0.4619		
Task 5—Dimensional Ratios A				
Depth Cue	1	1.9805	4.2877	**< 0.05**
Depth Cue * Subject	23	0.4619		
Task 6—Dimensional Ratios B				
Depth Cue	1	0.7500	1.6237	> 0.1
Depth Cue * Subject	23	0.4619		
Task 7—Absolute Depth A				
Depth Cue	1	0.0638	0.1381	> 0.25
Depth Cue * Subject	23	0.4619		
Task 8—Absolute Depth B				
Depth Cue	1	0.3763	0.8147	> 0.25
Depth Cue * Subject	23	0.4619		
Task 9—Relative Depth A				
Depth Cue	1	4.3802	9.4830	**< 0.01**
Depth Cue * Subject	23	0.4619		
Task 10—Relative Depth B				
Depth Cue	1	1.4180	3.0699	> 0.05
Depth Cue * Subject	23	0.4619		
Task 11—Object Recognition A				
Depth Cue	1	0.0638	0.1381	> 0.25
Depth Cue * Subject	23	0.4619		
Task 12—Object Recognition B				
Depth Cue	1	1.3333	2.8865	> 0.1
Depth Cue * Subject	23	0.4619		

Summary of Results

Figure 34 summarizes the performance differences associated with the addition of stereoscopic viewing. Although stereoscopic viewing in most tasks results in improved performance, only 10 of the 36 differences are statistically significant at $p \leq 0.05$. Note that almost all values in Figure 34 indicate a performance improvement. Only two tasks (4—Course Prediction B, and 12—Object Recognition B) show a significant performance decrement as a result of stereoscopic viewing.

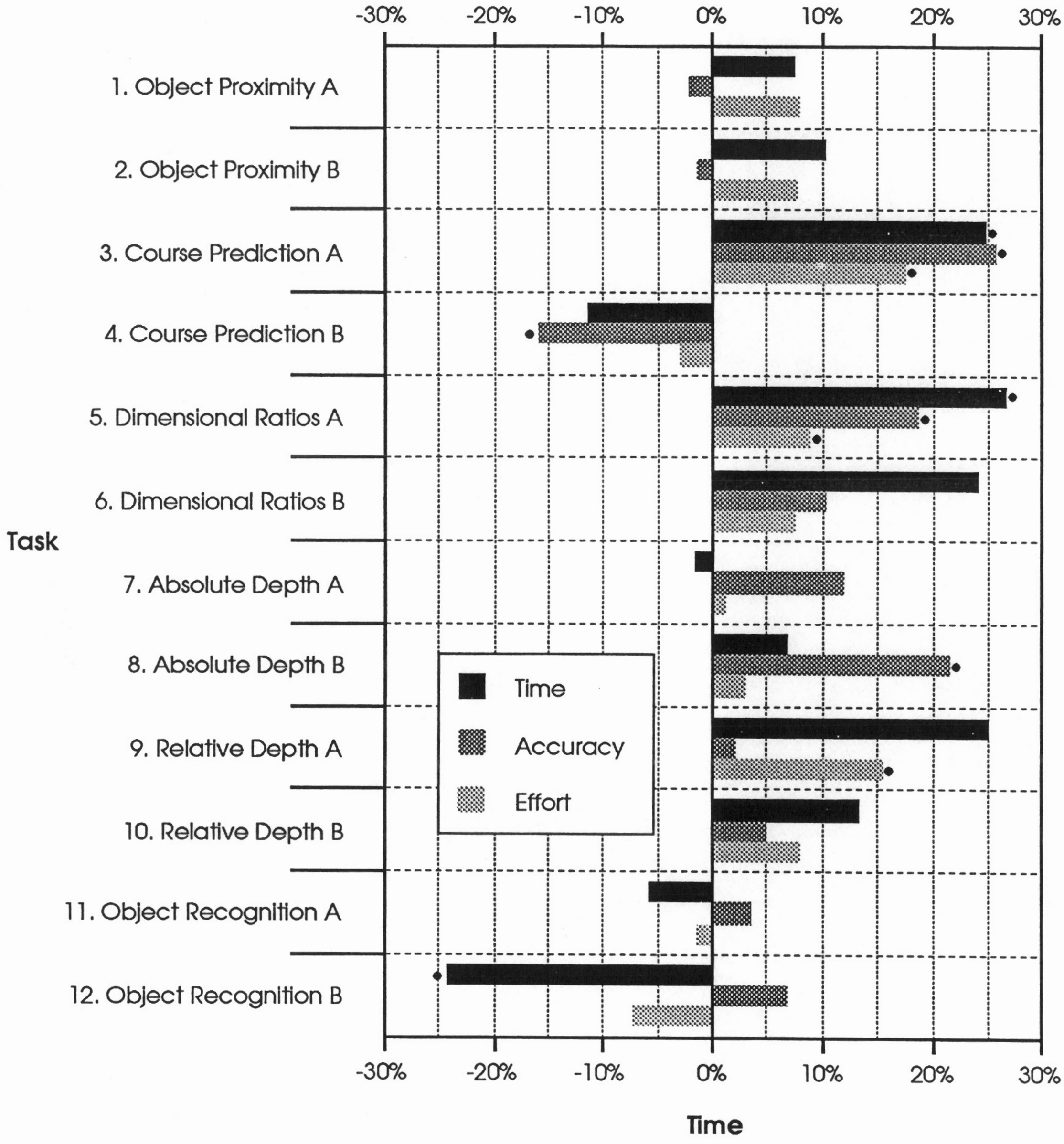


Figure 34. *Performance differences resulting from the addition of stereoscopic viewing. Values represent the percent improvement or decrement as viewing conditions change from monoscopic to stereoscopic. Values marked with • are significant at $p \leq 0.05$.*

DISCUSSION

Primarily, this research attempted to gain a better understanding of the stereoscopic advantage. Within this context, the apparent task-dependent nature of the stereoscopic advantage forms a significant issue. If it can be determined that the stereoscopic advantage is indeed task dependent, then it would be difficult to anticipate whether stereoscopic display would be appropriate for a given application. Accordingly, this work presents a component task method for handling such a dilemma. On the other hand, if it can be determined that the stereoscopic advantage is *not* dependent on the nature of the task, but instead dependent on more easily identifiable characteristics of the display system, then it would be easier to anticipate the appropriateness of stereoscopic displays for a given application.

Component Resolution

The observed performance differences across the 12 tasks demonstrate a task dependency. However, the 12 tasks used for this experiment were designed as representatives of six unique groups, with the underlying assumption that these six groups would provide adequate resolution for isolating the apparent task dependencies of the stereoscopic advantage. For some groups, this level of resolution was adequate. Tasks 1 and 2, for example, represented object proximity judgments. The observed stereoscopic advantages for each task were quite similar, implying that the level of component resolution for object proximity tasks is adequate and that we might successfully predict the stereoscopic advantage for other object proximity tasks. Tasks 3 and 4, on the other hand, displayed clearly different levels of stereoscopic advantage. The tasks both represented course prediction, but the observed performance differences make it difficult to generalize about the stereoscopic advantage for a generic course prediction task. Therefore, we must advance course prediction tasks to the next level of component resolution—a level at which Tasks 3 and 4 belong to different task groups. To do so, we must examine the qualitative differences between Tasks 3 and 4 that might have contributed to the performance differences. One prominent difference is that only one object moved in Task 3, but two objects were moving in task 4. A test of this difference would include two or more course prediction tasks with only one moving object, tested against two or more tasks in which two or more objects moved. If the stereoscopic advantage was similar within each of these groups but different between the groups, then the level of component resolution would be adequate for predicting the stereoscopic advantage. If not, then we might test another qualitative difference between Tasks 3 and 4—perhaps the direction of motion or distance traveled.

Because there are task groups which display significantly different levels of stereoscopic advantage, much of the following discussion will concentrate on identifying qualitative differences between the various tasks.

What other variables may have been responsible for significant performance effects? Likely candidates are: 1) the geometry of the elements that made up each scene, 2) the number of elements in each scene, and 3) whether the scene contained moving or stationary elements (or both). Each of these variables is directly related to display design, and each can affect viewer performance. Consider, for example, the possibility that the stereoscopic advantage is found to be present only within scenes that contain moving elements. This would suggest a link between stereopsis and dynamic stimuli, either leading to a conclusion about 3D display design (e.g., to promote or discourage dynamic stimuli in 3D displays) or suggesting appropriate future research. Either way, such a finding would weaken the claim that the stereoscopic advantage is, for the most part, dependent on the nature of the task.

Performance Decrements—Tasks 4 and 12

Given the question of whether to apply stereoscopic displays to an unspecified workplace, we should ask the question, "Are there any performance *decrements* associated with the introduction of stereoscopic displays?" If not, then we could apply 3D displays liberally, without concern for performance losses. However, there are to be at least two cases in which performance can suffer from stereoscopic display. Within Task 4, accuracy was degraded significantly (by approximately 17%) as a result of stereoscopic display. Withint Task 12, viewing time increased significantly (by approximately 24%) as a result of stereoscopic display. But were these decrements a result of a peculiarity of the paticular component task, or was it instead some design characteristic within the display for either task? Task 12 differed from the other 11 tasks in several known ways: 1) there were many scene elements (40 for Task 12 vs. 9 or fewer for Tasks 1–11), 2) the task contained two scenes displayed over time, and 3) the scene elements were disconnected and spread throughout the 3D space. Task 4 also differed significantly in that it was the only task which displayed more than one object in motion.

In Task 12, the 40 similar objects (dots) provided the eyes with multiple opportunities to misinterpret the 3D space. During stereoscopic presentation, Task 12 probably suffered from a "blizzard effect," in which multiple similar objects present multiple potentially corresponding stereoscopic pairs. In a scene like the one presented in Task 12, the eyes may fuse each laterally matched similar pair of dots. Each pairing presents a different retinal disparity and, consequently, a different depth cue. Because falsely corresponding pairs yield erroneous depth cues, the scene can be perceived falsely. Furthermore, until proper fusion is achieved, dots without a pair appear to only one eye and—due to the stereoscopic technology—appear half as bright as when paired correctly. Because brightness is a depth cue, these lone dots yield further erroneous depth cues. Eventually, the eyes will fuse on the veridical image in which all dots are matched with their correct pair. According to Julesz (1971), false correspondences are not a problem when the density of dots is low. However, "With increased dot densities, the visual system cannot find uniquely the corresponding points, and a new process has to be evoked which can resolve ambiguities by global considerations." In other words, the viewer must perceive the correct whole before attending to the correct parts. In Task 12, the task required the viewer to scan for displaced individual dots (parts). Needing to perceive the correct whole (to overcome the blizzard effect) would have interfered with the goal of perceiving the parts and thus may have interfered with the task itself. This interference may have been exaggerated as the participants "flipped" between the two scenes, as each scene flip would require additional fusion time.

In Task 12, no control was established to keep multiple dot pairs from lying along the same horizontal plane. Had there been such a control, the number of potential dot correspondences may have been reduced, and we might expect a decrease in the time required for processing. Indeed, a good guideline for 3D displays is to avoid placing similar objects along the same horizontal plane. If objects must lay along the same horizontal plane, then they should be made dissimilar (e.g., through shape or texture coding) to prevent falsely corresponding pairs. Furthermore, 3D displays should not contain patterns that repeat in horizontal fashion.

In summary, it appears that the Task 12 results are more due to the nature of the scene (a potentially ambiguous point cloud) than to the nature of the task (object recognition).

It was noted previously that the most unique feature of Task 4 was that it displayed more than one object in motion. Could this feature account for the decrement in accuracy when viewed stereoscopically? Or were there particular aspects of the stimulus design that contributed to the accuracy loss? It is known that motion parallax is a potent depth cue, often rivaling retinal disparity for importance. Certainly motion parallax was a useful depth cue in Task 4. However, under stereoscopic viewing conditions on displays such as the SGS-620, motion parallax can exist as an anti-cue. This is because when the viewer moves his head, he expects to see "around" to the sides of an object, yet because of the truly flat nature of the display medium, the object appears to "follow" the viewer from side to side. This illusion often interferes with proper presentation of motion parallax cues, and may have contributed to the accuracy decrement seen for Task 4 under stereoscopic viewing conditions. The motion parallax anti-cue would not have been present under monoscopic viewing conditions, and so the viewer would have been able to interpret the available motion parallax cues without interference.

Overall High Accuracy

For the stereoscopic advantage to exist, there must be room for improvement over monoscopic viewing performance. In other words, if a task is performed without error when using a monoscopic display, then there can be no error rate *improvements* by changing to a stereoscopic display. The accuracy rate for the 12 tasks employed in this experiment was quite high overall (Figure 28). All tasks displayed an average accuracy rate of at least 70%, and four tasks (1, 9, 11, and 12) displayed an average accuracy rate of greater than 90% *in both viewing conditions.* How does this affect the interpretation of the stereoscopic advantage for these tasks? At what accuracy rate can we expect no improvements, regardless of the superiority of the alternate display? Of the four tasks that displayed significant effects on accuracy, all were among the more difficult tasks as measured by accuracy (Figure 29). Yet the most difficult task (Task 4—Course Prediction B, at ~72% accuracy) displayed a significant accuracy *decrement* when viewed stereoscopically. Merritt (1983) commented that an adequate evaluation of 3D display benefits requires that (among other things) the study incorporate "time constraints, error penalties, and learning factors typical of the real world situation." While such a concept is contrary to the goal of this research (to be able to predict real-world effects without testing every real-world scenario), it is important to consider that the 3D visual task is likely to be influenced by other cognitive workloads. Because the tasks were simple and because the participants were not subject to the kind of constraints described by Merritt, the threat of ceiling effects is considerable. Useful future research might attempt to replicate the performance effects found in this study, but across a range of task difficulties, possibly by applying secondary tasks or time constraints.

Effort Rating Equal to Viewing Time

Note the similarities between viewing time and mental effort rating as shown in Figures 26 and 31. These similarities are displayed below in a scatterplot of mean viewing time versus mental effort rating (Figure 35). Note the deviation represented by Task 12, which has an average viewing time below the line of best fit. This deviation is not surprising considering the significant increase in viewing time in the stereoscopic condition. In fact, when the data point corresponding to Task 12 is removed from consideration, the correlation increases to 0.96 ($r^2 = 0.924$, $p < 0.001$). A second scatterplot (Figure 36) shows the relation between the stereoscopic advantage in viewing time versus the stereoscopic advantage in mental effort rating. Here, the correlation is a significant 0.92 ($p < 0.001$). Apparently, the mental effort ratings were little more than a subjective measure of viewing time, in spite of the fact that subjects were clearly instructed that mental effort referred to "the amount of attention or concentration that is required to perform a task" (Appendices A and C). It is not surprising that two (Tasks 3 and 5) of the three tasks that displayed a stereoscopic advantage for mental effort also displayed the same advantage for viewing time.

The mental effort ratings do reveal one result not observed with viewing time—a significant main effect of depth cue. Mental effort ratings were significantly lower in the stereoscopic condition. Although this difference is practically negligible (see Figure 33), it does suggest that viewers enjoy a modest reduction in mental effort when viewing stereoscopic images.

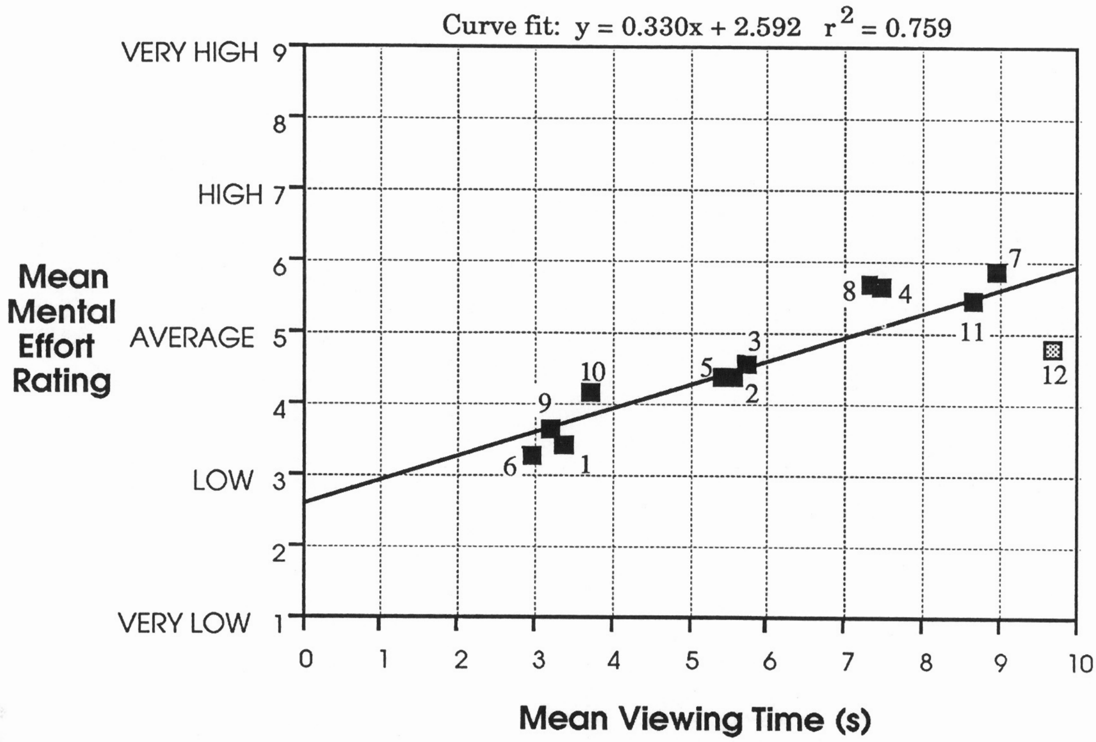


Figure 35. *Plot of mean mental effort ratings vs. mean viewing times. The high correlation suggests that mental effort ratings were little more than subjective judgments of viewing time. When Task 12 (data point at far right) is excluded, the r^2 value increases to 0.924.*

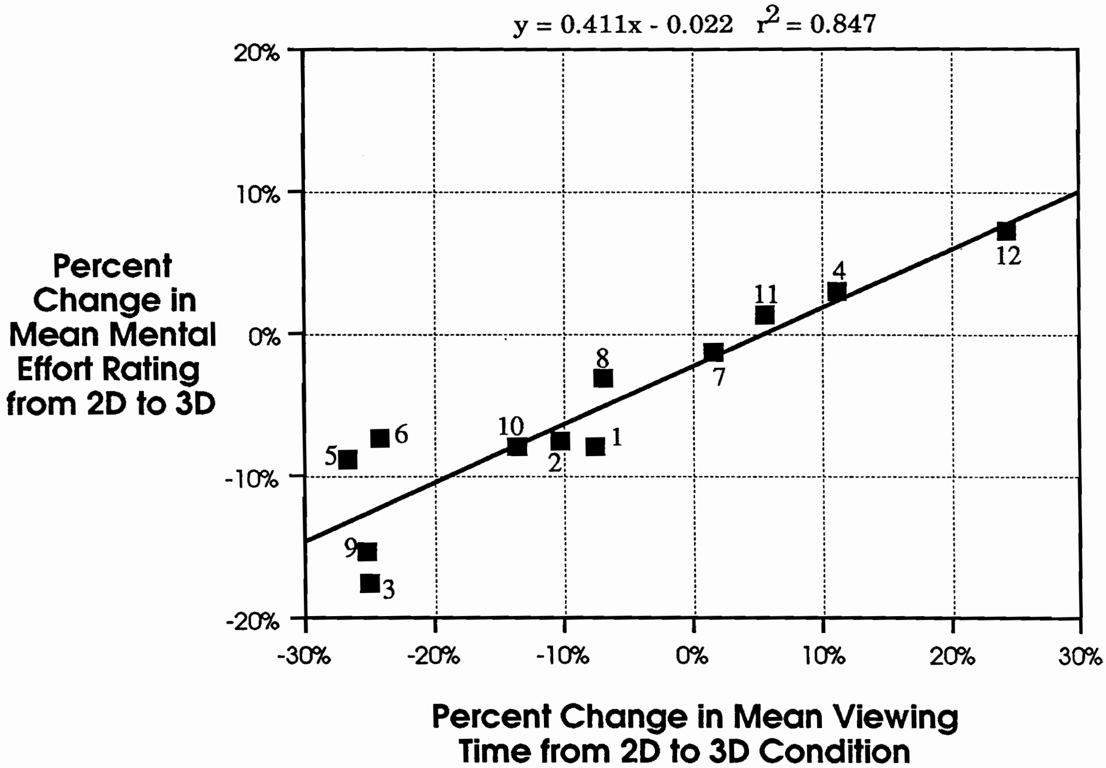


Figure 36. *Plot of 2D-to-3D percent changes in mean viewing time versus the corresponding changes in mean mental effort ratings.*

Viewing Time

For 3 of the 12 component tasks, stereoscopic display yielded a significant difference in viewing time. As discussed earlier, the effect within Task 12 was to *increase* viewing time. For the remaining two tasks (Tasks 3 and 5), however, the effect was to *decrease* viewing time. Aside from the inherent task differences, what characteristics among these tasks might have led to the significant decreases in viewing time? Tasks 3 and 5 are among the "fastest" tasks in this study (Figure 27), but Task 12 was among the "slowest" tasks, which implies a link between viewing time and the stereoscopic advantage: tasks which are performed more quickly are more likely to display viewing time improvments when viewed stereoscopically. A possible explanation is that, for some tasks, the minimum time required to perform the task itself (e.g., rotate an object, judge distance) may exceed the maximum time required to perceive the image, even when presented monoscopically. In such a case, stereoscopic presentation, even though it may yield faster interpretation of the 3D scene, will not improve the *overall* task time. Since we are unable to measure the time required to perform each separate mental process, this hypothesis cannot be tested. Regardless, the results suggest that 3D displays do not yield improvements in viewing time for tasks that, by their nature, require a relatively long viewing time.

Texture and Depth Cues

Julesz (1971) stated that "only textured objects can lead to unambiguous localization in depth. A stereogram of outlined circles could be the retinal projection of a disk, a sphere, or even a cone." While the scene elements in this experiment were not textured *per se*, all stationary objects were filled and hidden-line algorithms were applied. Truly textured surfaces (especially ones which followed proper spatial contours) were not achievable and so were not implemented. If Julesz's statement is true without exception, then the lack of textured surfaces may have biased the study. Although filled surfaces do reduce scene ambiguity, the depth information provided by a fill is not entirely accurate. In fact, the area within the fill provides *no* depth information up to the point at which the eye reaches the edge of the figure.

Because of the limitations of computing speed, the moving scene elements in Tasks 3, 4, and 7 were drawn without filled areas. Thus, these elements contained greater depth ambiguity than the filled elements. Does this ambiguity compromise the validity of results for these particular tasks? The mean accuracy for these tasks was low (ranked 4th, 1st , and 2nd, respectively—Figure 29), but because the filled versus non-filled division was confounded with the stationary versus dynamic division, it cannot be determined which feature—fill or motion—was responsible for the observed results. The object motion in Tasks 3, 4, and 7 did provide depth cues in the form of motion parallax. Because motion parallax is such a potent depth cue, it is expected that the wire-frame ambiguity did not adversely affect the overall performance in these tasks.

Application to a Complex 3D Task

An earlier section, "Describing a Complex 3D Task," described a method for applying the performance data gathered in this experiment toward prediction of the stereoscopic advantage for a more complex 3D task. As an example of this procedure, this section will describe application to a hypothetical 3D task. The example will be limited to effects on viewing time.

Example task. Consider an air traffic control task in which an operator must process the 3D position of numerous aircraft within a 3D space above an airport. To be able to apply the component task data, we must first determine which 3D component tasks are present in the air traffic control task, and weight them according to their importance. For the purposes of this example, we will treat all 12 component tasks as unique task groups. Hypothetically gained through interview of trained operators, the weighted importance ratings are shown in the first column of Figure 37. Because the operator's most important duty in performing this task is to prevent collisions between moving aircraft, Task 4 (Course Prediction B) accounts for the largest proportion of the control task (40%). Other tasks have less importance, and some tasks are irrelevant.

The observed performance effect for each of the 12 component tasks is shown in the second column of Figure 37. Significant performance effects are represented by the ratio of monoscopic to stereoscopic performance (see Figure 34 for an overview of these values).

The resulting stereoscopic advantage for viewing time for this hypothetical task is 1.048. This means that we could expect a 4.8% improvement in viewing time by changing from a monoscopic to a stereoscopic display system. While a 4.8% improvement may be useful, we see that the stereoscopic display system does not improve performance for the major component of the task—that which corresponds to Task 4 (Course Prediction B).

Task independence. This process is a linear additive model, and one of the assumptions underlying this type of model is that the components are independent. It was mentioned earlier that there was expected to be some overlap among the 12 component tasks, which implies a known overlap, thus violating the assumptions of the model. Because this is an exploratory concept, such violations are not of immediate concern. However, they should be of concern during any future efforts to improve the set of component tasks and their associated predictive ability. One known method for isolating independent groups of predictive factors is that of factor analysis. To ensure independence (as well as more valid task groupings), a factor analysis could be applied to a larger set of component tasks. Those independent tasks could be used to predict the stereoscopic advantage for a complex 3D task. Next, to validate the predictive ability of the tasks (and the component task method), actual operator performance would be measured using both monoscopic and stereoscopic display systems. The predictive stereoscopic advantage would then be compared to that measured under actual use.

Other Issues

The stereoscopic advantage may be task dependent, but surely this is not the only dependency. Cole and Parker (1989) mention the extra need for 3D viewing in an otherwise visually degraded environment. Cole and Parker tested remote manipulation in an environment that simulated the bright and diffuse lighting experienced while in orbit. A significant stereoscopic advantage was found. Buffett (1980) described stereoscopic viewing as a means for overcoming fidelity losses associated with various television systems. Merritt (1984) described several tasks "virtually impossible without 3-D." These tasks involve either: scenery which is degraded by camouflage, low-fidelity transmission, poor lighting, et cetera; or scenery which is unfamiliar. Julesz (1971), noted that there are certainly tasks for which stereoscopic viewing is a clear benefit or is sometimes required, such as defeating monocular camouflage.

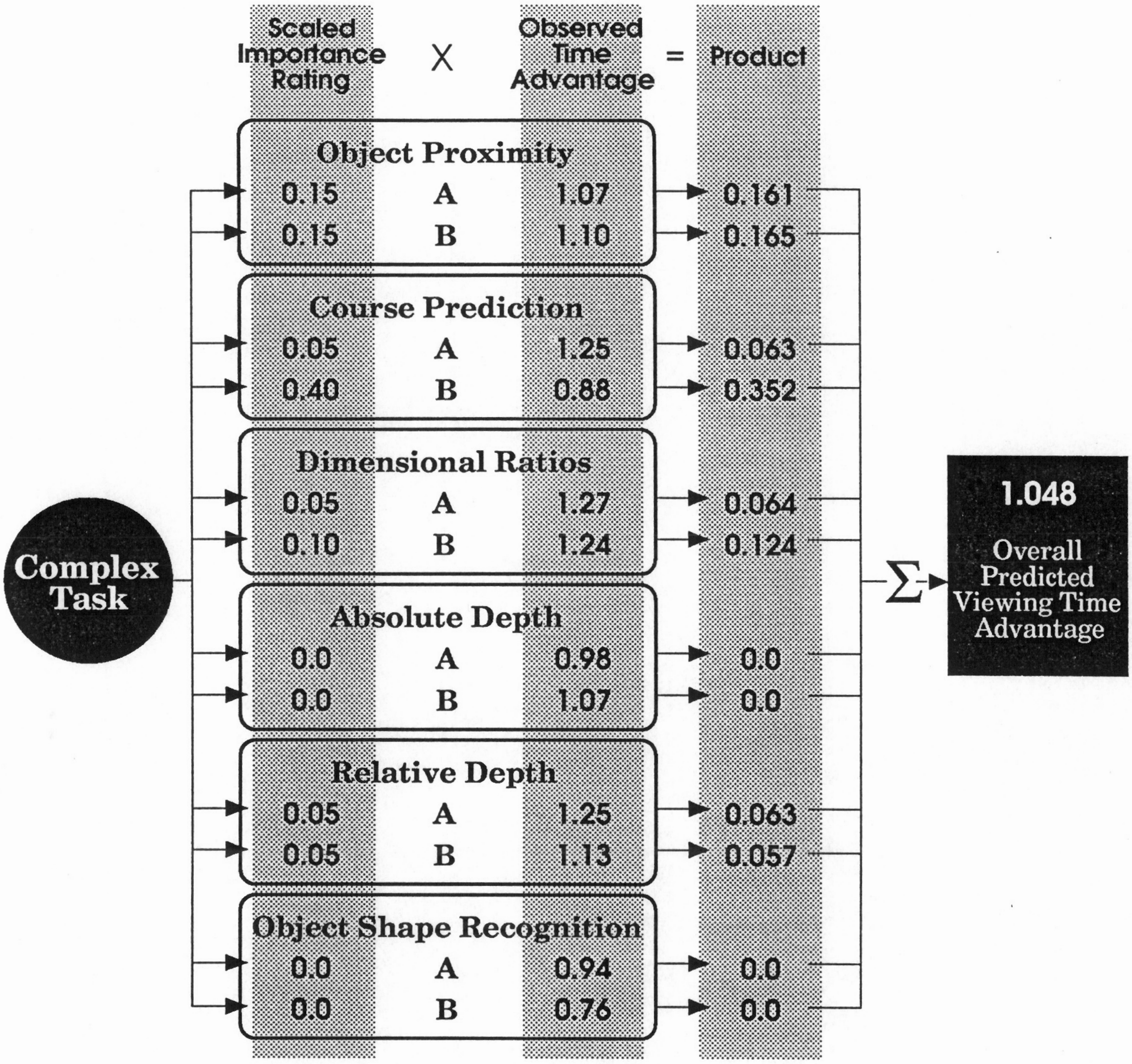


Figure 37. *Example of applying the component task performance data to a hypothetical 3D task. Here, the viewing time demands of an air traffic control task have been described by the level of importance of each component task (Column 1). These importance ratings are multiplied by the stereoscopic advantage factors (Column 2) and summed to yield an overall predictive factor. In this case, a stereoscopic display would improve viewing times by 4.8% over a monoscopic display.*

CONCLUSIONS

The larger goal of this research was to contribute to the growing body of knowledge about 3D display systems. The results are analyzed not in an attempt to better understand human visual processes, but more to determine how such factors impact the design of 3D displays. Accordingly, any conclusions drawn from this study are ones that might assist in the design of 3D displays. If such conclusions are couched in terms of 3D visual processing, that is only because 3D visual processing is the arena of interest, not the area of application.

3D versus 2D

When taken in the context of all relevant research, the results of this study suggest that when certain display guidelines are not violated, stereoscopic display is expected to improve or at least maintain the overall level of viewer performance for most tasks (as measured by viewing time and accuracy). However, this generalization must be tempered by the knowledge that the stereoscopic advantage, although quite real, is dependent on the task, and that this task dependency is not yet clearly understood.

As evidence for the above conclusion, consider the results of Task 12, which displayed a significant performance decrement in the presence of stereoscopic viewing. The particular flaws in the display for Task 12 probably had more to do with the performance loss than did the task of object recognition itself. One of these display flaws is an example of violating a 3D display guideline—to avoid placing similar objects so that they repeat or form a pattern in horizontal fashion. Although much of the 3D literature is devoted to such display guidelines, they are neither fully mature nor fully complete. Still, these guidelines represent the best known means for designing a useful and effective stereoscopic display system.

Task Dependency

This research indicates that the stereoscopic advantage is dependent on the nature of the visual task. However, the six component task groups developed for this study are inadequate to describe the precise nature of this task dependency. Within-group differences in the stereoscopic advantage suggest that a greater level of resolution is required for some task groups (e.g., Course Prediction). Even so, the component task method appears promising. Future work with factor analysis techniques may help to form more valid and independent component task groupings.

Although much of the relevant literature mentions tasks for which stereoscopic presentation has clear advantages, these tasks may be better defined by their visual characteristics than by the nature of the task. For example, the alleged observed benefits of stereoscopic viewing while performing remote operations in an undersea environment probably have more to do with the degraded viewing conditions than with the remote operation task itself.

Future Research

This research has created a foundation for a new method of analyzing and predicting the efficacy of stereoscopic display systems. As such, there are many shortcomings in the method, and future research is required to make improvements:

1) The component tasks require more accurate definitions so that they can be applied without ambiguity in the breakdown of complex tasks.

2) The set of six tasks may need to be enlarged to include new components.

3) This research identified several component task groups that require a greater level of resolution before any proper generalizations can be made toward complex tasks. Factor analysis may prove useful in developing predictive and independent component task groups.

4) Because this research was limited to graphical images, the scope is inherently limited and future research is required to determine whether the stereoscopic advantage differs in situations that include realistic scenery. The same limitation is true for sophisticated animation, which was not included here.

5) A more rigorous measure of mental effort ratings may reveal important results.

6) By selecting more difficult tasks, or by applying secondary tasks or time constraints, future research may identify stereoscopic advantages that were not observed in this study due to ceiling effects in accuracy.

7) This research suggests the possibility of a viewing time constraint on the presence or absence of the stereoscopic advantage. Future research could investigate whether this viewing time effect is a real phenomenon or is simply due to the constraints of this study.

REFERENCES

Beaton, R. J. (1989, August). *Photometric evaluation of liquid-crystal shutter devices for field-sequential stereoscopic display devices: Final report.* Blacksburg, VA: Displays and Controls Laboratory, Virginia Polytechnic Institute and State University.

Box, G. E. P. (1954). Some theorems on quadrative forms applied in the study of analysis of variance problems. *Annals of Mathematical Statistics, 25*, 290–302.

Buffett, A. R. (1980). The visual perception of depth using electro-optical display systems. *Displays, 2* (1), 39-45.

Carswell, M. C., and Wickens, C. D. (1987). Information integration and the object display: An interaction of task demands and display superiority. *Ergonomics, 30* (3), 511–527.

Cavanaugh, P., and LeClerc, Y. G. (1989). Shape from shadows. *Journal of Experimental Psychology: Human Perception and Performance, 15* (1), 3-27.

Cole, R. E., and Parker, D. L. (1989). Stereo TV improves manipulator performance. In W. E. Robbins and S. S. Fisher (Eds.), *Three-dimensional visualization and display technologies*, SPIE Vol. 1083. Bellingham, WA: Society of Photo-Optical Instrumentation Engineers.

Coules, J. (1955). Effect of photometric brightness on judgments of distance. *Journal of Experimental Psychology, 50*, 19-25.

Davies, T. N. (1985). An analysis of the depth cues in technical graphics. *Le Travail Humain, 48* (4), 321–330.

Dosher, B. A., Sperling, G., and Wurst, S. A. (1986). Tradeoffs between stereopsis and proximity luminance covariance as determinants of perceived 3D structure. *Vision Research, 26* (6), 973–990.

Gooding, L. W. (1991). *Effects of retinal disparity depth cues on cognitive workload in 3-D displays.* Unpublished doctoral dissertation. Virginia Polytechnic Institute and State University, Blacksburg, VA.

Gooding, L., Miller, M. E., Moore, J., and Kim, S. (1991). The effect of viewing distance and disparity on perceived depth. J. O. Merritt and S. S. Fisher (Eds.), *Stereoscopic displays and applications II*, SPIE Vol. 1457 (pp. 259–266). Bellingham, WA: Society of Photo-Optical Instrumentation Engineers.

Julesz, B. (1960). Binocular depth perception of computer-generated patterns. *Bell System Technical Journal, 39* (pp. 1125-1162).

Julesz, B. (1971). *Foundations of Cyclopean perception.* Chicago: University of Chicago Press.

Kim, W. S., Ellis, S. R., Tyler, M. E., Hannaford, B., and Stark, L. W. (1987). Quantitative evaluation of perspective and stereoscopic displays in three-axis manual tracking tasks. *IEEE Transactions on Systems, Man, and Cybernetics, SMC-17* (1), 61-72.

Lippert, T. M., and Benser, E. T. (1987). Photointerpreter evaluation of hyperstereographic forward looking infrared (FLIR) sensor imagery. In D. F. McAllister and W. E. Robbins (Eds.), *True three-dimensional imaging techniques,* SPIE Volume 761 (pp. 85-87). Bellingham, WA: Society of Photo-Optical Instrumentation Engineers.

Livingstone, M., and Hubel, D. (1988). Segregation of form, color, movement, and depth: Anatomy, physiology, and perception. *Science, 240,* 740–749.

McBurney, D. H., and Collings, V. B. (1984). *Introduction to sensation/perception* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Merritt, J. O. (1983). Common problems in the evaluation of 3D displays. *1983 SID International Symposium: Digest of Technical Papers*. Coral Gables, FL: Lewis Winner.

Merritt, J. O. (1984). Visual tasks requiring 3-D stereoscopic displays. C. Outwater (Ed.) *Optics in entertainment*, SPIE Vol. 462 (pp. 56–59). Bellingham, WA: Society of Photo-Optical Instrumentation Engineers.

Miller, R. H. (1991). *Efficacy of retinal disparity depth cues in three-dimensional visual displays.* Unpublished master's thesis. Virginia Polytechnic Institute and State University, Blacksburg, VA.

Miller, R. H., and Beaton, R. J. (1991). Some effects on depth-position and course-predictions judgments in 2-D and 3-D displays. J. O. Merritt and S. S. Fisher (Eds.), *Stereoscopic displays and applications II,* SPIE Vol. 1457 (pp. 248–258). Bellingham, WA: Society of Photo-Optical Instrumentation Engineers.

Mountford, S. J., and Somberg, B. (1981). Potential uses of two types of stereographic display systems in the airborne fire control environment. In *Proceedings of the Human Factors Society 25th Annual Meeting* (pp. 235-239). Santa Monica, CA: Human Factors Society.

Nataupsky, M., and Crittendon, L. (1988). Stereo 3-D and non-stereo presentations of a computer-generated pictorial primary flight display with pathway augmentation. *Proceedings of the AIAA/IEEE 8th Digital Avionics Systems Conference*: San Jose, CA.

Reinhart, W. F. (1990). *Effects of depth cues on depth judgments using a field-sequential stereoscopic CRT display*. Unpublished doctoral dissertation. Virginia Polytechnic Institute and State University, Blacksburg, VA.

Reinhart, W. F. (1991). Depth cueing for visual search and cursor positioning. J. O. Merritt and S. S. Fisher (Eds.), *Stereoscopic displays and applications II,* SPIE Vol. 1457 (pp. 221–232). Bellingham, WA: Society of Photo-Optical Instrumentation Engineers.

Shephard, R. N., and Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science, 171,* 701–703.

Smith, D. C., Cole, R. E., Merritt, J. O., and Pepper, R. L. (1979). *Remote operator performance comparing mono and stereo TV displays: The effects of visibility, learning, and task factors*. Naval Ocean Systems Center, San Diego, CA. NOSC-TR-380.

Spain, E. H., and Holzhausen, K. P. (1991). Stereoscopic versus orthogonal view displays for performance of a remote manipulation task. J. O. Merritt and S. S. Fisher (Eds.), *Stereoscopic displays and applications II*, SPIE Vol. 1457 (pp. 103–110). Bellingham, WA: Society of Photo-Optical Instrumentation Engineers.

Tye, M. (1991). *The imagery debate*. Cambridge, MA: MIT Press.

Uttal, W. R. (1983). *Visual form detection in 3-dimensional space*. Hillsdale, NJ: Lawrence Erlbaum.

Vandenberg, S. G., and Kuse, A. R. (1978). Mental rotations: A group test of three-dimensional spatial visualization. *Perceptual and Motor Skills, 47*, 599–604.

Way, T. C. (1988). Stereopsis in cockpit display—a part-task test. In *Proceedings of the Human Factors Society 32nd Annual Meeting* (pp. 58-62). Santa Monica, CA: Human Factors Society.

Way, T. C. (1989). 3-D in pictorial formats for aircraft cockpits. In *Proceedings of the Human Factors Society 33rd Annual Meeting* (pp. 23-27). Santa Monica, CA: Human Factors Society.

Way, T.C., Hobbs, R. E., Qualy-White, J., and Gilmour, J. D. (1989). *3D imagery cockpit display development*. Boeing Report WRDC-TR-89-XX, April.

Wickens, C. D. (1984). *Engineering psychology and human performance*. Columbus, OH: Charles E. Merrill.

Wickens, C. D., Todd, S., and Seidler, K. (1989). *Three-dimensional displays: Perception, implementation, and applications*. CSERIAC SOAR 89-001 Technical Report. Wright-Patterson AFB, OH: AAMRL.

Woodruff, R. R., Hubbard, D. C., and Shaw, A. (1986). Comparison of helmet-mounted visual displays for flight simulation. *Displays, 76*, (4), 179-185.

Woodson, W. E. (1981). *Human factors design handbook*. New York, McGraw-Hill.

Yeh, Y., and Silverstein, L. D. (1990). Visual performance with monoscopic and stereoscopic presentations of identical three-dimensional visual tasks. In *Society for Information Display Digest of Technical Papers* (Vol. (XXI). Playa del Rey, CA: Society for Information Display.

Yeh, Y., and Silverstein, L. D. (1992). Spatial judgments with monoscopic and stereoscopic presentation of perspective displays. *Human Factors, 34* (5), 583–600.

Zenyuh, J. P., Reising, J. M., and Walchli, S. (1988). A comparison of a stereographic 3-D display versus a 2-D display using an advanced air-to-air format. In *Proceedings of the Human Factors Society 32nd Annual Meeting* (pp. 53-57). Santa Monica, CA: Human Factors Society.

APPENDIX A—INSTRUCTION SET A

In this experiment, you will be trained to perform a battery of 12 simple tasks. Each of these tasks involves viewing a computer-generated image that depicts a three-dimensional space. Based on the task at hand and the content of each image, you will make decisions regarding the spatial relationships between various objects in each image.

Training vs. Performance

This experiment will be divided into two sessions: a training session and a performance session. Because the purpose of the training is to ensure your understanding of each of the 12 tasks, you are encouraged to ask questions during the training session. Although each of the tasks is quite simple, it is important that you clearly understand the nature and goal of each task. Therefore, do not hesitate to be inquisitive during this portion of the experiment. Once the training is completed, you will proceed immediately with the performance session.

During each session, you will perform each task 10 times before progressing on to the next task. As the *training* session progresses and each new task appears in turn, the experimenter will read aloud a short description of the goal for that particular task. Also during the training session, you will receive feedback as to the correctness of your answers. You should use this feedback to refine your performance. During the performance session, you will not receive feedback and so you must rely on your training to answer correctly.

How to Perform a Task

Before each task, the computer will prompt you with a statement in the form of a question that describes the decision required to complete the upcoming task. This prompt will change according to the goal of each task. Once you have read the question and are ready to view the task image, you will begin the task trial by pressing the space bar on the keyboard. At this time, the task image will appear on the computer display. After you have viewed the image long enough to answer the task question, you will again press the space bar and the image will disappear. At this time, you will view a screen that restates the task question at the top of the display, followed by a list of possible answers. Here you will use the arrow keys to move a cursor and highlight your choice of answers. Once you have done this, you will press the space bar to enter your answer and move on to the next trial.

At all times during the experiment, you will need to use only the space bar and arrow keys on the keyboard. In general, you should use the space bar to progress through the various phases of the task trials and to enter your choice of task answers or mental effort ratings. You should use the arrow keys to choose between task answers and mental effort ratings. One task in particular will require special use of the arrow keys—you will be instructed in this usage during the training for that particular task.

Measuring Your Performance

Your performance will be measured by the computer, which will record both viewing time and number of correct answers. The computer will consider viewing time as the time when the task image is present on the computer screen. This period begins when you press the space bar to first view an image, and ends when you press the space bar again to clear the image. A third measure used in this experiment deals with the level of Mental Effort required to perform each task. This measure is gained by simply asking you to rate your level of Mental Effort using a 9-point scale. The Mental Effort measure will only be recorded during the performance session.

Because Mental Effort is a special measure, let's discuss what exactly is meant by mental effort. Mental Effort refers to the amount of attention and/or concentration required to perform a task. Mental Effort could involve memorizing items, performing calculations, concentrating on listening to a speaker for important points, problem solving, or making difficult decisions. High levels of Mental Effort are required in situations that demand total concentration, whereas lower levels of Mental Effort are required when your mind wanders or your attention is distributed over more than one easy task component. In the experimental tasks that will follow, the tasks could be very difficult, requiring you to accurately judge an object's location based on limited spatial information. Conversely, the tasks could be very easy, with the answer to the question being immediately obvious. The tasks will require you to: judge an object's spatial location, predict an object's future position, mentally rotate objects, compare the positions of several objects, or perceive subtle features of objects. The difficulty of these tasks will change as you are provided with varying levels of different types of spatial information.

One way to measure Mental Effort is to measure the precision or speed with which an individual performs a task. Although the computer will be recording correct answers and viewing time, these measures cannot assess what you as a participant experience in the way of Mental Effort. Therefore, we will measure Mental Effort by asking you to use the following 9-point scale to rate your level of Mental Effort.

9 - VERY HIGH Mental Effort
8 -
7 - HIGH Mental Effort
6 -
5 - AVERAGE Mental Effort
4 -
3 - LOW Mental Effort
2 -
1 - VERY LOW Mental Effort

If we were to use this scale as an *absolute* measure, we might consider daydreaming as a "1" and intense concentration as "9." However, in this experiment, you should use the scale as a *relative* measure. Once you have completed your task training, you will have been exposed to every possible task condition in the experiment. Therefore, you will be acquainted with the full range of Mental Effort required to perform the tasks in this experiment. During the performance session when you will be asked to rate Mental Effort, you should do so by rating the task **as it compares with all the tasks in the experiment**. For example, if you felt a certain task trial was among the most difficult in this experiment, then you should rate that task trial as a "9." Conversely, if a certain task trial was among the easiest in this experiment, you should rate that task trial as a "1." Altogether, you should use the entire 9-point scale to rate Mental Effort over the 12 tasks.

Your Strategy

Throughout the training session, you should proceed with each trial taking as much time as is necessary to answer the task question. If you find a task easy or quickly become able to answer correctly, then you should feel free to proceed at a faster pace. However, it is important to remember that during the training session your level of performance is less important than your level of understanding.

During the performance session, you should proceed with each task trial by proceeding as quickly as possible *without sacrificing accuracy*. However, if you cannot decide on an answer after viewing the scene for a considerable period of time, you should simply make your best guess.

At all times, remember that your speed is measured only during that time when the task image is present on the computer screen. Therefore, you may proceed at a comfortable pace whenever the task image is not on the screen. If you need to pause or rest during the experiment, you may do so at any of these times. If you need a break or have to leave the room, please alert the experimenter and he will take appropriate measures to pause the experiment.

APPENDIX B—INSTRUCTION SET B

Prelude (read during dark adaptation)

When viewing each of the task images, you will see various objects that are located in what should appear to be a hollow cube. This cube, which defines the boundaries of our three-dimensional workspace, should appear to be as tall as it is wide as it is deep. In some task conditions, the appearance of this cube may differ slightly. However, you should always interpret the three-dimensional space as a true cube.

When analyzing objects within the 3D cube, remember that in order to properly determine an object's position, you must take into account all three dimensions: height, width, and depth. Whether an object is in motion or stationary, you must analyze all three dimensions to perform the tasks properly.

As I read the task instructions for each task, do not be concerned if you don't understand at first. Each task will become more clear once you have had a chance to view the task image. If after you have attempted the task once or twice and still do not understand, then you should ask any questions you might have. Although most of these tasks are quite simple, they may not become clear until you have tried them once or twice.

Before a new task begins, you will see a blinking green alert. Whenever you see this green alert during the training session, tell me that you are ready to receive instructions for the next task.

Task 1

Which object is closest to the center object X?

In this task, you will view five objects located within the 3D cube. One object (labeled X) will always be located in the center of the cube. The remaining four objects will be located at random positions, each a different distance from the center object. Your task is to determine which of the four outer objects is closest to the center object. Each of these four objects will be identified with a number, one through four.

As usual, remember to account for all three dimensions when analyzing object positions.

To begin this task, press the space bar. Once you have viewed the task image and determined which object is closest to the center object, press the space bar again. You may begin when ready.

Task 2

Which pair of objects is closest to each other?

Here you will view four objects located within the 3D cube. Your task is to determine which pair of objects has the smallest distance between them. The objects will be numbered one through four.

As usual, remember to account for all three dimensions when analyzing object positions.

To begin this task, press the space bar. Once you have viewed the task image and determined which pair of objects is closest, press the space bar again. You may begin when ready.

Task 3

Which object, if any, will be hit by the moving object?

This task image will contain nine objects. Eight of these objects will be stationary and located in the eight corners of the 3D cube. The ninth object will originate at the center of the cube and begin moving towards an edge or corner of the 3D cube. However, you will only be able to witness one half of the moving object's complete path. Based on this information, your task is to predict the moving object's final destination. The moving object will cycle through its path again and again, so if you are unable to determine its destination after the first cycle, keep watching and the object will retrace its path. Finally, remember that the moving object will not always hit one of the stationary objects.

As usual, remember to account for all three dimensions when analyzing object positions and movements.

To begin this task, press the space bar. Once you have viewed the task image and determined whether the moving object will hit one of the eight stationary objects, press the space bar again. You may begin when ready.

Task 4

Will the two objects collide?

This task image will contain two moving objects. As the objects move from their starting point, your task is to determine whether they will collide at some point in space. However, you will only be able to witness one half of the objects' complete paths. Based on this information, your task is to predict the final destinations of the two objects and, consequently, whether the two objects will collide. The moving objects will cycle through their paths again and again, so if you are unable to determine the final destinations after the first cycle, keep watching and the objects will retrace their paths.

As usual, remember to account for all three dimensions when analyzing object positions and movements.

To begin this task, press the space bar. Once you have viewed the task image and determined whether the two objects will collide, press the space bar again. You may begin when ready.

Task 5
Which block is a different length?

In this task, you will view three blocks located in the 3D cube. One of the three blocks will be longer than the other two. This difference in length may be in any direction in three-dimensional space. Your task is to determine which block is longer. The blocks will be labeled A, B, and C.

As usual, remember to account for all three dimensions when analyzing object positions and sizes.

To begin this task, press the space bar. Once you have viewed the task image and determined which block is longest, press the space bar again. You may begin when ready.

Task 6
Is the object a true cube? That is, are all dimensions of the object identical?

In this task, you will view a single object that may or may not be a true cube. A "true cube" is defined as an object that has six sides, all of which are perfect squares. If any of the sides is not a perfect square but instead just a rectangle, then the object is not a true cube. Thus, your task is to analyze all three dimensions of the object and determine whether they are identical.

As usual, remember to account for all three dimensions when analyzing object positions and sizes.

To begin this task, press the space bar. Once you have viewed the task image and determined whether the object is a true cube, press the space bar again. You may begin when ready.

Task 7
Using the grids at the top and bottom as a scale, how far in depth does the object move?

In this task, you will view a single object that will move in depth. This means that the object will appear to move closer or farther away from you. Sometimes the object will move a large distance (up to the full depth of the 3D cube) and other times it will appear to move only a very small distance. In either case, the object will move closer and farther and closer and farther, continuing in cyclical fashion. Your task is to determine the absolute value of this depth movement. By using the 8-segment grids at the top and bottom of the 3D cube, you should be able to assign a number (1 through 8) to the amount of depth over which the object moves.

To begin this task, press the space bar. Once you have viewed the task image and determined the extent of depth movement, press the space bar again. You may begin when ready.

Task 8

Using the grids at the top and bottom as a scale, how far apart in depth are the two objects?

In this task, you will view two stationary objects that are separated in depth by some distance. By depth, we mean that the objects may be closer or farther away from you. Your task is to determine the absolute depth distance between the two objects. By using the 8-segment grids at the top and bottom of the 3D cube, you should be able to assign a number (1 through 8) to the amount of depth that separates the two objects. Please remember that it is the difference in ***depth*** that you should evaluate. Although the two objects will be separated by some horizontal difference, you should ignore this horizontal separation when judging depth.

To begin this task, press the space bar. Once you have viewed the task image and determined the depth separation between the two objects, press the space bar again.

Task 9

Which object appears closest to you?

In this task, you will view two stationary objects that are separated in depth by some distance. By depth, we mean that the objects may be closer or farther away from you. Your task is to determine the relative depth of the two objects and report which appears closest to you, the left object or the right object. Please remember that it is the difference in depth that you should evaluate. Any vertical or horizontal differences should be ignored.

To begin this task, press the space bar. Once you have viewed the task image and determined which object is closest (left or right), press the space bar again.

Task 10

Which object is closest (in depth) to the center object?

In this task, you will view three stationary objects. One object will be in the center, and the outer two objects will lie to the left and right of the center object. Each of the two outer objects will be separated in depth from the center object. By depth, we mean distance as in farther or closer. In making depth judgments in this task, you should ignore any horizontal or vertical differences and focus only on depth. Your task is to determine which of the depth separations is smallest.

As usual, remember to account for all three dimensions when analyzing object positions.

To begin this task, press the space bar. Once you have viewed the task image and determined which object is closest in depth to the center object, press the space bar again. You may begin when ready.

Task 11

Except for a rotation, are the two objects the same or different?

In this task, you will see two stationary objects positioned side by side. Each object will resemble a group of blocks that form a particular shape. The two objects will appear different from each other and there are two possible explanations for this difference. The two objects may be the same shape but rotated to appear different or the two objects may be two different shapes. When analyzing the objects, you may need to rotate them in your own mind to satisfy yourself that they are the same object or two different objects.

As usual, remember to account for all three dimensions when analyzing object positions and shapes.

To begin this task, press the space bar. Once you have viewed the task image and determined whether the objects are the same or different, press the space bar. You may begin when ready.

Task 12

Are the two point clouds identical?

In this task, you will see a group of about 40 dots. We call this group of dots a point cloud, because the dots (or points) are grouped together at random in a cloud. This point cloud will have a pattern, and your task will be to compare the pattern of this point cloud to the pattern of a second point cloud. To view the second point cloud, you will use the arrow keys. Once you have viewed the first cloud, you should press any of the arrow keys to view the second point cloud. When you press an arrow key, the image will disappear for a moment, then the second point cloud will appear. If you press any of the arrow keys again, you will return to the first point cloud. You may use any of the arrow keys as many times as you want to go back and forth between the two point clouds. In doing so, you should be able to determine whether the patterns formed by the two point clouds are identical.

As usual, remember to account for all three dimensions when analyzing the positions of the dots that form the two point clouds.

To begin this task, press the space bar. Once you have viewed the task image and determined whether the two point clouds are the same or different, press the space bar. You may begin when ready.

APPENDIX C—INSTRUCTION SET C

Performance Session Instructions

Now that you have completed the training session, you are ready to proceed with the performance session. Because your training session has ended, the computer will no longer tell you whether you answered correctly or incorrectly. In the tasks which follow, you must rely on your training to answer correctly. Also, you do not need to await instructions before each new task. Simply proceed with each new task when you are ready.

After each task, you will be asked to rate your level of Mental Effort required to perform the previous task. Remember that Mental Effort means the amount of attention or concentration that is required to perform a task. In other words, if a task seems difficult, then a high Mental Effort rating is appropriate. If the task seems easy, then a low Mental Effort rating is appropriate. Also remember that you should base your rating on the total range of task difficulties over the 12 tasks. Therefore, when making a Mental Effort rating, try to recall the task difficulties you experienced during your training, then make your rating based on that range.

From now on, please do not ask questions **while you are viewing a task image**. It is OK to ask questions before or after you view a task image.

If you need a break, please tell me at this time. Otherwise, you may begin the first task.

VITA

Robert Howard Miller, the son of Mr. and Mrs. Robert Charles Miller, was born October 20, 1966, in Chicago Heights, Illinois. From 1984 to 1988, Robert attended the University of Illinois in Urbana, Illinois, where he received the Bachelor of Science degree in Liberal Arts and Sciences. From 1988 to 1994, Robert attended the Virginia Polytechnic Institute and State University, where he received the Master of Science and Doctor of Philosophy degrees, both in Industrial and Systems Engineering. Both Robert's thesis and dissertation efforts focused on human factors issues regarding three-dimensional visual display systems and their application.

Robert currently is employed as a human factors engineer with Thomson Consumer Electronics in Indianapolis, Indiana. There he is involved with the research, design, and usability testing of display and control systems for various consumer electronics products.