The Effectiveness of Alternative 2D Projections of 3D Forms

Eric N. Wiebe
North Carolina State University

Presented at the Mid-Year Meeting of the Engineering Design Graphics Division of the American Association for Engineering Education
Oct 30-Nov 2, 1993
Ohio University, Athens, Ohio

Abstract

Presentation styles of 2D projections of 3D forms on a computer display is explored. The two independent variables explored are: the rate in which a series of orthographic projections are displayed and whether they were depicted as line drawings or shaded. Subjects saw 40 objects in one of six possible display designs and asked to make a forced choice decision based on the display of each object. The error rate and reaction time were used as an indicator of effectiveness of the display design to form a mental image of the object. Scoring on the Vandenberg Mental Rotations test given at the beginning of the experiment was used as a covariate in the statistical analysis examining the display designs.

Introduction

Computer graphics continues to make inroads into the classroom at all educational levels. Increasingly, the graphics employ not only static graphic images but also dynamic sequences of images driven by animation software. Animation tools have reached the point of maturity that standards have been established for the format image sequences. Animated sequences can be developed, exported and integrated in a wide range of instructional software including most popular CAD software packages.

Though it is inevitable computers and computer generated animations are going to be a part of instruction in Engineering Design Graphics, little is known about the effectiveness of various forms of graphic presentation. Of particular concern is the use of computers for the effective development of spatial visualization abilities in students (Wiley, 1990). Current research from members of our division has looked at a number of issues pertaining to the use of computers in the classroom. Sexton (1992) compared 3D wireframe modeling to traditional non-computer methods of teaching in an introductory engineering graphics course. Miller (1992) looked at the effectiveness of 3D solid models on the computer and real solid models as supplements to a traditional engineering graphics course.

One of the commonalties in the two cited research studies is the use of paper-based tests as instruments in the experimental studies. Central to the study of spatial visualization is the identification of reliable measures. Both studies used more than one instrument with varying success. Others, such as Ross & Aukstakalnis (1993), question the usefulness of any paper-based vehicle and advocate the use of emerging technologies such as Virtual Reality for the development of new instruments.

Another approach to the study of the effectiveness of presentation techniques is to apply them in an operational setting and evaluate performance on a specific task. McWhorter, Rodriguez, & Hodges (1992) evaluated the use of stereoscopic viewing technology and the angle of view of a 3D scene on the performance of a simulated crane operation task. Here the dependent measure was not a score on a test, but the time taken to complete the task.

Confounding any of these approaches is the lack of agreement among researchers as to a theoretical underpinning to the concept of spatial visualization. Researchers in a number of fields — including perception, cognition, human-computer interfaces, and artificial intelligence — have proposed often divergent theories (Wiebe, 1993). One of the themes that does emerge from much of this research is the concept of the *canonical form*. Through the evaluation of edges, corners, and other primitive features of an object, an internal mental representation of the object is created independent of any specific viewpoint (Biederman, 1987; Hochberg, 1964).

If this internal representation of an object is built up from the evaluation of lower level features of the object, what is the best way of presenting representations of these features to the viewer? Traditional engineering graphics uses, among other techniques, a multiview drawing with three views of the object mutually perpendicular to each other. Edges are represented as contrasting lines on a background with faces and vertices represented by various configurations of these edges. This could be called a line representation of the object. Figure 1. Computer graphics has allowed a number of variations on this representation. Using a 3D computer database, projected line representations of the object can be generated from any location about the object. Animation techniques use this 3D database to produce a series of projections and display them fast enough to give the perception of apparent motion.

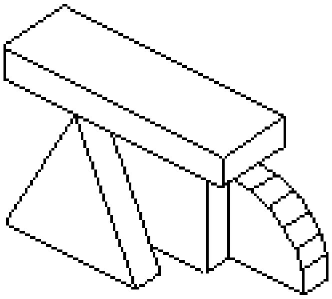


Figure 1.

Another variation on representing these lower level features is to alter the method of defining edges between faces. Figure 2. Rather than using a line representation, shading methods can be used instead. With shading methods, the effect of a light source on the value of the surface color is simulated in the representation. Areas of constant or uniformly shifting gradients of value represent a single face, whereas boundaries defined by an abrupt shift in value represent an edge between two faces.

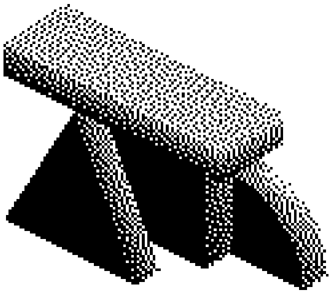


Figure 2.

As did McCuistion (1991), this study looks at the difference in static and dynamic presentations of projections of 3D forms on the development of visualization ability. Changes in the projection of an object alters the spatial configuration of its features in the image. A static presentation of multiple projections allows a holistic synthesis of changes in the pattern the edges make. There is the freedom to compare which ever projection you would like. Dynamic presentation of the projections structure the comparisons by only allowing them to be viewed in a sequential fashion. With dynamic presentation the projections are overlaid rather than distributed across the computer screen. This means the changes in location of the projections are not artificially induced in order to present them all simultaneously in an image.

Research Goal

The goal of this study is to explore the use of feature-based analysis as a theoretical underpinning for spatial visualization. Static and dynamic presentations of projections of 3D forms offers very different methods for the viewer to compare and contrast the low level features which comprise the object. Shaded versus line representation of edges (and therefore faces and vertices) offers alternate coding methods for these features. Though some researchers have found no improvement in the interpretation of 3-D forms displayed with shading over those presented as line drawings (Barfield, Lim, & Rosenberg, 1990), the popularity of computer rendering tools makes it worth confirming these results. A secondary goal for the study is to look at alternative methods for evaluating spatial visualization performance, avoiding some of the pitfalls encountered by other researchers.

Two different display techniques are used as independent variables in the experiment. The first display technique is parallel versus serial presentation of the same group of orthographic projections of an object. The fundamental difference in these two types of presentations is that the parallel format presents all of the projections on the display at the same time whereas the serial format presents the projections one at a time. Within the serial format there is also the variable of speed of presentation of the projections. One of the two serial sequences is presented at a fast enough rate to allow for the perception of apparent motion. The second display technique is the application of shading to faces of the object's projections. Of interest is whether either display technique assists in the evaluation of the low level features of the object and whether there is a significant interaction between the two display techniques.

Method

Subjects

72 subjects were used from a pool of students taking the introductory psychology course, PSY 200, at NC State University. The subjects were screened for normal visual acuity for near distances (such as a computer screen) and accepted into the experiment if they achieve at least 20/30 vision with or without lens correction. Subjects were also screened for previous training in technical/engineering graphics and rejected if they had taken any courses in these disciplines. Even though students at colleges and universities taking introductory engineering/technical graphics courses come from diverse backgrounds, they still are more highly pre-selected for developed visualization abilities than the student population at large. By drawing from the PSY 200 pool and specifically screening out subjects with previous graphics training, the goal was to get a more heterogeneous group of 'novice' subjects.

Experimental Stimuli

In all of the computer displays, the subjects will look at a three-dimensional object described through a series of orthographic projections. Orthographic projection is the historical method for depicting objects in technical and engineering graphics. Convergence, which adds a degree of realism in perspective projection, is not present in orthographics. Besides allowing for more simplified computation for generating the displays and ease in direct measurement of the object, both Braunstein (1986) and Ullman (1979) argue that convergence is not a significant factor in perceiving three-dimensional form.

Though researchers have shown that the detection of three-dimensional structure is possible in as few as two (Lappin, Doner, & Kottas, 1980) or four (Ullman, 1979) projections, a total of thirteen different projections are used. The thirteen projections will represent the object undergoing rotation in 15 degree increments about both the vertical and horizontal axes. This rotational sequence allows for the display of all three primary orthographic views of the object (front, top, and right side) along with intermediate views at 15 degree intervals. Figure 3. Though the rotational axes were primarily chosen for their significance to technical and engineering graphics, they also have perceptual significance. Green (1961) found that rotation about the vertical axis allowed superior accuracy of perception of three-dimension forms than any other rotational axes. All other axes, including skew axes, gave the same performance.

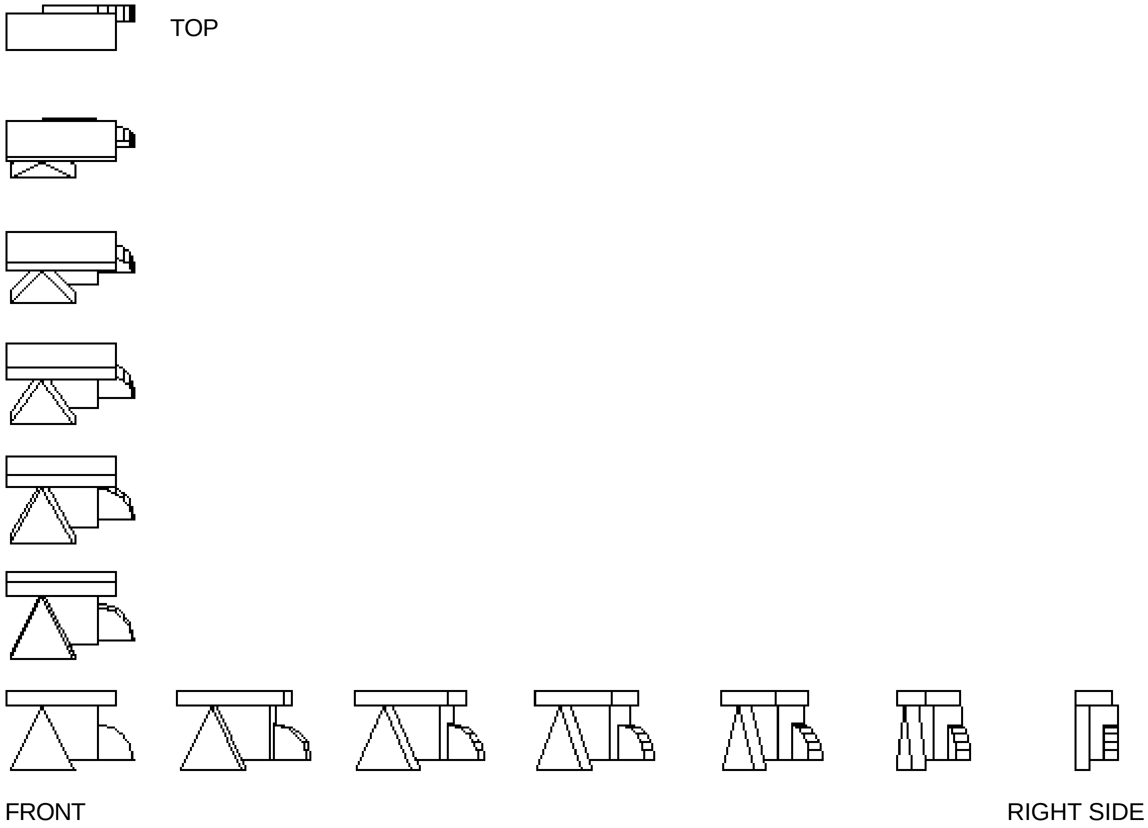
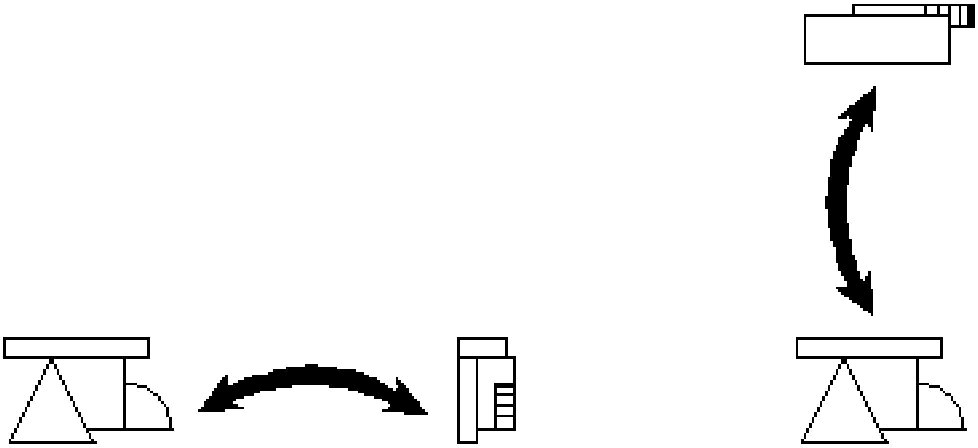


Figure 3 (principle views labeled for illustration only)

Whether these thirteen projections are shown in parallel — all displayed on the screen at the same time — or serially constitutes the Format technique. The thirteen projections shown in parallel will be considered the Static display; the content of the display will not change during the entire exposure interval. The organization of the Static display will be the same as is shown in Figure 3 and allows for uninterrupted scanning of the projections as the object goes through a 90 degree rotation about either axis.

The second variation of the Format technique (called Sequential) is to show the very same projections in a serial manner: only one projection of the object is shown on the computer display at a time. The projections are shown at a slow enough rate (1 frame/sec) as so to preclude apparent movement in depth as it undergoes rotation. Petersik (1980) concluded that temporal factors are much more critical than spatial factors in inducing the perception of apparent motion from a sequence of projections. Objects could be rotating as much as 180 degrees between projections as long as the time between projections is kept under 300msec. The third variation on this technique (called Dynamic) is a serial presentation where the projections are shown at a fast enough rate (4 frames/sec, 250msec between frames) as so to induce apparent motion. This presentation is akin to animations shown on computers or film.

Since all three variations on this display technique are to be shown on the display for the same length of time, the Sequential variation becomes the format determining the minimum display time. At 1 frame/sec for thirteen projections, 13 sec is needed to see all of the projections. These projections will be presented so the object undergoes rotation from the right side view to front view and then to top view. The Dynamic variation, rotating at 4 times the rate of the Sequential, will undergo a different sequence during the 13 sec. Presenting the projections at a rate that induces apparent motion, the object will rotate back and forth between the right and front views three times producing a "rocking" motion. Figure 4a. It then proceeds to rock back and forth between the front and top views three times. Figure 4b. The Static display will simply stay on the screen for the 13 sec.



a. b.

Figure 4.

The second display technique varies the representation of a projection of the object. In the variation called Line, the object's edges are represented by black lines against a white background (see Figure 1). Any edges that would normally be obscured from view would not be shown. The other variation on this technique, called Shaded, applies a gray shading to the faces on the object (see Figure 2). This variation recreates the effects of a single infinite light source projected onto the surface of the object.

The experimental stimuli will be grouped into two independent variable dimensions representing Format (Static/Sequential/Dynamic) and Rendering (Line/Shaded). Combined they represent six possible computer display designs. Table 1.

	Format		
Rendering	Static Line	Sequential Line	Dynamic Line
	Static Shade	Sequential Shade	Dynamic Shade

Table 1.

Task and procedure

The subject will view an object represented by one of the six possible display designs for the specified time of 13 sec. After a screen blanking interval of 1 sec, a pair of different objects in the same Rendering format is displayed side-by-side from the same projection. Figure 5. The subject is asked to chose whether the right or left object is the same as the one previously viewed. Both the response time and accuracy is measured. There is a practice block of stimuli (10 objects) at the beginning of the experiment after which the subject is shown their average response time and error rate before continuing with the experiment. The wrong object will vary from the correct one in topology and/or geometry. That is, both the number and types of faces on the object may be different along with the size and shape of them. The projection, an isometric pictorial, used to display this pair of objects will be different than any of the projections seen in the experimental display(s). This is done in order to confound direct visual comparison without interpreting the three-dimensional form.

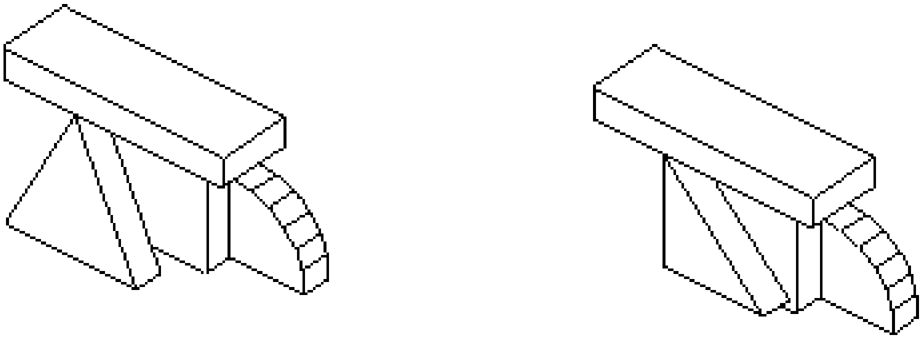


Figure 5.

Experimental design

The experiment will employ a split-block 3 x 2 design with three types of Format (Dynamic/Sequential/Static) and two types of Rendering techniques (Line/Shaded). The subjects are assigned to one of six groups (12 per group) counterbalancing for gender. Previous studies by Vandenberg & Kuse (1978) and others have shown there is a small but constistent difference in the performance between males and females on spatial visualization tasks. Each group will participate in one training block (10 objects) and one experimental block (30 objects). Both the training and experimental blocks use one of the six possible combinations of Format and Rendering shown in Table 1. Whether the correct object is displayed in the right or left isometric pictorial in the second display is randomized. Each group would see the same 40 objects in the same sequence.

The Error Rate and Response Time data collected in this experiment will be subjected to a two-way analysis of variance (ANOVA). Besides comparisons of individual groups, interaction between Rendering and Format will be analyzed along with (if appropriate) the main effects of these two groups. Prior to viewing the 40 objects, the subjects take a paper test designed to measure one's spatial visualization ability. The Mental Rotations test (Vandenberg & Kuse, 1978) is a 20 question paper test based on an experimental technique developed by Shepard & Metzler (1971). This test measures the subject's ability to mentally rotate three-dimensional objects and, it is hypothesized, a partial measure of the ability to do the type of task required in this experiment. The subject's score on the Mental Rotations test will be used as a covariate in the ANOVA.

Apparatus

An Apple Macintosh IIcx with a 16 inch Viewsonic monitor was used to display the experimental stimuli. The monitor was set to only display gray scale colors. Subjects sat 45cm from the monitor with there eyes approximately level with the center of the screen. This seating arrangement gave a visual angle of 3.6 degrees for a single projection in an experimental display, 29 degrees for the full Static display, and 9.5 degrees for a single projection in the forced choice pair.

A standard Extended Macintosh keyboard was used with the computer. The subjects responded to the forced choice pair by pressing either the right or left arrow key. They then gave the confidence level in their choice (high or low) by pressing either the up or down arrow keys. The subjects were instructed to adjust the keyboard on the desktop to a comfortable position an keep their fingers positioned over the arrow keys.

The initial subject information, the instructions to the responding to the stimuli, and the stimuli themselves were displayed and controlled through a script written in HyperCard 2.0. Timing for the stimuli was controlled by the script as was the collection of response selection and reaction time. This data along with the initial subject information was written to a text file after completing the viewing of the 40 stimuli. An external program, QuickTime, was used in conjunction with HyperCard to drive the image sequences seen in the Dynamic and Sequential formats.

Results

Response Time

The ANOVA performed on Response Time indicated there was no significant interaction between the two main effects of Format and Render or with the covariant, Mental Rotations test score (Test). Since the interaction terms appeared to be a minimal contribution to the model, an F-test was performed on the complete model and a model with no interaction terms. The null hypothesis was there is no difference in the Error Sum of Squares of the complete and reduced model. The null hypothesis was not rejected (F(7,60)=.275, p>>.05), justifying the removal of the interaction terms. A comparison of the R^2 terms ($R^2_{complete}$=.3046, $R^2_{reduced}$=..2823) also shows a minimal contribution by the interaction terms.

With interaction removed from the model, there was a significant difference in Response Time and the main effect of Format (F(2,67)=7.26, $p < .002$). Table 2. Given a Total Sum of Squares for the whole model of 44.70, the Format variable accounts for 15.5% of the variability in Response Time.

There was also a significant relationship between Response Time and the covariant Test (F(1,67)=10.46, $p < .002$). The Mental Rotations test score accounted for 11.2% of the variability in the Response Time. The removal of Test Score as a covariant in the model showed a considerable drop in its predictive capability ($R^2_{complete}$=.2823, $R^2_{reduced}$=.1703).

Source	**Df**	**Sum of Sq**	**Mean Square**	**F Value**	**Pr > F**
Format	2	6.94870805	3.47435402	7.26	0.0014
Render	1	0.66210219	0.66210219	1.38	0.2438
Test	1	5.00659562	5.00659562	10.46	0.0019

Table 2.

A Student-Newman-Keuls means test was performed *ad hoc* on the main effect of Format. It indicated that at the alpha level of .05, the significant difference was between the Dynamic format and the two other formats of Sequential and Static. Table 3.

SNK Grouping	**Mean** (Sec)	**N**	**Format**
A	2.662	24	Sequential
A	2.332	24	Static
B	1.903	24	Dynamic

Table 3.

Error Rate

The ANOVA performed on Error Rate indicated there was no significant interaction between the two main effects of Format and Render or with the covariant, Test. The interaction of Render and Test could possibly be considered marginally significant (F(1, 60)=2.85, $p < .10$). Just as with the dependent variable Response Time, an F-test was performed on the complete model and a model with no interaction terms. The null hypothesis was not rejected (F(7,60)=.634, p>>.05), justifying the removal of the interaction terms. A comparison of the R^2 terms ($R^2_{complete}$=.3319, $R^2_{reduced}$=.2825) shows more of a contribution by the interaction terms than with Response Time, but still not enough to be significant.

With interaction removed from the model, there was a significant difference in Error Rate and the main effect of Render (F(1,67)=15.12, $p < .001$). Table 4. Given a Total Sum of Squares for the model of 249.27, the Render variable accounts for 15.1% of the variability in Error Rate.

As with Response Time, there was a significant relationship between Error Rate and the covariant Test (F(1,67)=8.80, $p < .005$). The Mental Rotations test score accounted for 9.4% of the variability in the Error Rate. The removal of Test Score as a covariant in the model showed a considerable drop in its predictive capability ($R^2_{complete}$=.2825, $R^2_{reduced}$=.1882)

Source	**Df**	**Sum of Sq**	**Mean Square**	**F Value**	**Pr > F**
Format	2	9.36111111	4.68055556	1.75	0.1811
Render	1	37.55555556	37.55555556	14.07	0.0004
Test	1	23.49838267	23.49838267	8.80	0.0042

Table 4.

Though there is a significant difference between the two levels of the Render variable, note that the mean score difference is only slightly more than a single incorrect response. Table 5.

RENDER	**Mean** (Incorrect Responses)	**N**
Line	1.472	36
Shade	2.917	36

Table 5.

Discussion

As seen in the results of the ANOVA, both the Format and Render variables have a significant effect on different aspects of task performance. The Format variable had a significant effect on the amount of time it took the subjects to choose one of two objects from the forced choice pair. Those subjects seeing the Dynamic format responded significantly faster than those seeing either the Sequential or Static format. This result seems to concur with McCuistion's (1991) findings on the effects of dymanic imaging on mental rotation test scores. Though the difference was not significant, in was interesting to note that those subjects seeing the Static format responded faster than those seeing the Sequential format. This result seems to rule out the inherent superiority of a serial presentation for this particular task. If this were the case, then the Static format should have had to slowest Response Time.

By decreasing the time between projections, the Dynamic format differentiated itself from the Sequential format in two ways. First, the decreased time allowed for the perception of apparent motion. Second, because each format was shown on the screen for the same length of time, the faster presentation of projections in the Dynamic format allowed for multiple presentations of the same projections within the same time period.

Unlike the Sequential format, the dynamic format did not show the projections for uniform length of time. Viewing an animated sequence induces a "momentum", that is, an expectancy of what the next projection will look like. Since the direction of rotation seen in the projection was reversing during the rocking back and forth, the animation "paused" at the three primary projections (right, front, and top) before reversing directions to allow the subject to perceptually prepare for the change. The rocking motion — first between the right and front and then the front and top — coupled with the pausing created differences in both the number and length of time a projection was seen. Figure 6.

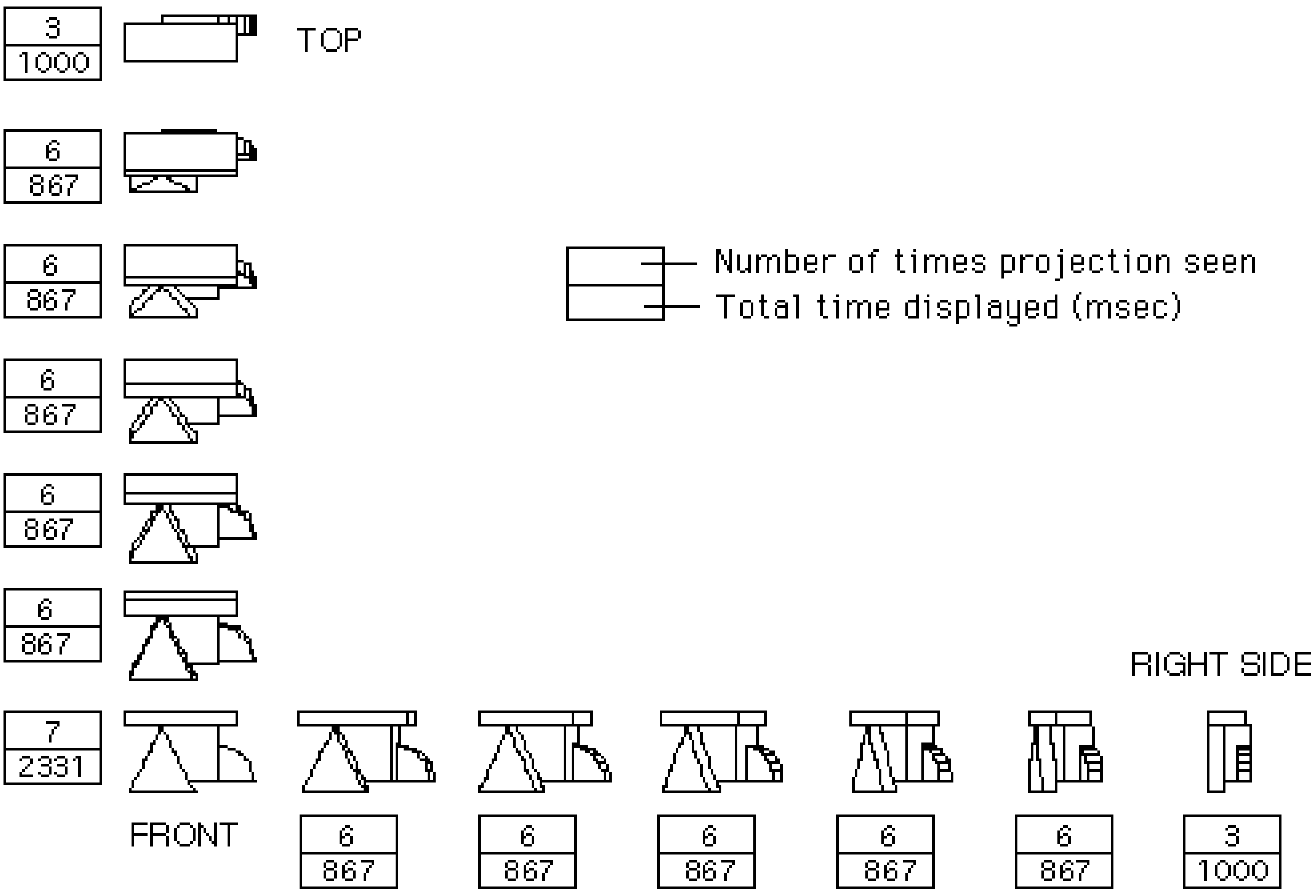


Figure 6.

The number and length of time each projection was seen in the Dynamic format may be closer to an optimal distribution of favored views for forming the mental representation of the object. With the Static format, the subject has the privilege of focusing attention on any projection, allowing a completely personalized strategy of viewing. With the complete Static display contained within a view angle of approximately 41 degrees and individual projections within approximately 7 degrees, focus on either single projections or small groups of projections is possible. What is not possible with the Static projection, is the apparent motion present in the Dynamic format. With the Sequential format, no one projection is favored, as each is seen only once for 1 sec. This format has neither the flexibility of viewing of the Static format nor the apparent motion of the Dynamic format.

The Render variable had a significant effect on the number of times the subjects chose the wrong object of the pair but not on the Response Time. Subjects viewing the Shade rendering had a higher Error Rate than those viewing the Line rendering. These results differed from the results from studies employing similar tasks. Barfield, et al. (1990) found no significant difference between line and shade renderings in a task that involved judging the location of two objects in 3D space. Sanford et al. (1987) found shade rendering was superior to line rendering in a mental rotation task modeled after Shepard & Metzler (1971). One possible explanation would be the Shade renderings were more likely to lead to incorrect interpretations in the object. The problems in interpretation could arise either during the initial formation of the mental image of the object or in analyzing the forced choice pair.

Both the Line and Shade renderings of the objects represent edges where two surfaces of an object join. The difference is in their method. The Shade rendering attempts to more closely represent how these edges are perceived on real objects. Light falling on a real object is perceived as having different value based in part on the orientation of the surface to the light source. The edge between surfaces is represented by a discontinuity in the gradient of value level. This line of discontinuity represents the juncture of two surfaces having different orientations to the light source and, therefore, to each other. The Line rendering abstracts this point of discontinuity by representing the edge as a line in a contrasting color to the surfaces. In the case of this experiment, the Shade rendering represented surfaces in gradients of gray whereas the Line rendering depicted the edges as black and the surfaces as white. It could be that though the Line rendering is more abstract than the Shade rendering, it allows a more direct recognition of edges and, therefore, less prone to erroneous interpretation. The edges, in turn, are used in the perception of the whole object (Biederman, 1987).

The high correlation of the Mental Rotations test score and performance on the task confirms the initial assumption that this task employed mental imaging aptitudes measured by this instrument. At the same time, the small shift in the F values of the Render and Format values when Test was removed as a covariate in the ANOVA seems to indicate that it was not the dominant factor affecting performance by the subjects. Analysis of the sum of squares indicate that score on the Mental Rotations test accounted for only 15.61% of the variability in the Response Time and 13.14% of the variability in the Error Rate. The results of this study is a reaffirmation of the complexity of the perceptual and cognitive processes employed in the task simulated by this experiment. Neither the Render nor the Format variable significantly affected both the measured elements of the task, Error Rate and Response Time. Even where these variables had a significant effect on performance on the task, neither accounted for the majority of the variability in performance. Format accounted for 15.1% of the variability of the Response Time and Render accounted for 22.6%.

Future Directions

The amount of variability accounted for by the two variables explored in this experiment is large enough to make either one worth further study. One possible route would be to examine more carefully the sequence of projections displayed to the viewer. With proper equipment, viewers scanning the Static format could be analyzed for which projections were attended to the most and in what sequence. Another approach related to this would be to use a serial presentation, but allow the user to interactively control which projection is viewed. The strategy employed by the viewer; including which projections, in what order, and for how long, could be analyzed. The use of interactive control of the projections could also be compared to the Dynamic presentation used in this experiment. A possible hypothesis is that active control of the projections displayed allows for a more efficient construction of a mental image of the object. This could be thought of as a comparison of passive and active viewing. Even restricting the format to a passive presentation, other formats could be explored. Having the object revolve a complete 360 degrees about each axis could be compared to the rocking motion shown in the Dynamic format.

The Render variable could also be explored in more depth. There are many computer algorithms used for shading computer models. Many of them create different gradients of value across the surfaces of the object and could be compared to the algorithm used in this experiment. Another approach would be to take a closer look at Line rendering. In addition to edges and surfaces, another perceived feature of 3D form is the termination of edges at vertices. With the current Line rendering, edges are distinguished from surfaces by a strong value shift. Vertices could also be coded by altering the color of the ends of the edges where they terminate. Other perceptual coding techniques such as thickening the ends of the edges might also be explored as a way of coding vertices.

The development of three-dimensional visualization skills contributes to the ability of students to both synthesize and manipulate three-dimensional forms from two-dimensional pictures on a computer display. These skills are important not only to the student while in school but in a broad range of engineering and technical professions. This line of research begins to look at a few aspects of the problem of how best to represent a three dimensional form on a computer display. For a given situation, what is the optimum temporal rate, format, and rendering for displaying a group of projections representing the object?

Bibliography

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. Psychological Review, 94, 115-147.

Braunstein, M. L. (1986). Perception of rotation in depth: the pyschophysical evidence. In N. I. Badler & J. K. Tsotsos (Eds.), Motion: representation and perception (pp. 242-247). Elsevier Science.

Green, B. F., Jr. (1961). Figure coherence in the kinetic depth effect. J. of Experimental Psych., 62, 272-282.

Hochberg, J. E. (1964). Perception. Englewood Cliffs, NJ: Prentice-Hall.

Lappin, J. S., Doner, J. F., & Kottas, B. L. (1980). Minimal conditions for the visual detection of structure and motion in three dimensions. Science, 209, 717-719.

McCuistion, P. J. (1991). Static vs. dynamic visuals in computer-assisted instruction. Engineering Design Graphics Journal, 55(2).

McWhorter, S., Rodriguez, W., & Hodges, L. F. (1992). The performance of graphical user interfaces in 3-D interactive animation (virtual) environments. Engineering Design Graphics Journal, 56(3), 12-17.

Miller, C. L. (1992). Enhancing visual literacy of engineering students through the use of real and computer generated models. Engineering Design Graphics Journal, 56(1), 27-38.

Petersik, J. T. (1980). The effects of spatial and temporal factors on the perception of stroboscopic rotation simulations. Perception, 9, 271-283.

Ross, W., & Aukstakalnis, S. (1993). Virtual Reality: Implications for Research in Engineering Design Graphics. In ASEE-Engineering Design Graphics Division Mid-Year Meeting, (pp. 1-8). San Francisco, CA:

Sanford, J., Barfield, W., & Foley, J. (1987). Empirical studies of interactive computer graphics: perceptual and cognitive issues. In Proceedings of the Human Factors Society - 31st Annual Meeting, (pp. 519-523).

Sexton, T. (1992). Effect on spatial visualization: Introducing basic engineering graphics concepts using 3D CAD technology. Engineering Design Graphics Journal, 56(3), 36-43.

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. Science, 171(Feb.), 701-703.

Ullman, S. (1979). The Interpretation of Visual Motion. Cambridge, MA: MIT Press.

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

Wiebe, E. N. (1993). Visualization of three-dimensional form: A discussion of theoretical models of internal representation. Engineering Design Graphics Journal, 57(1), 18-28.

Wiley, S. E. (1990). Computer graphics and the development of visual perception in engineering graphics curricula. Engineering Design Graphics Journal, 54(2), 39-43.