Effects of Animation, User-Controlled Interactions, and multiple static views in Understanding 3D Structures

Taylor Sando[*]
University of Manitoba

Melanie Tory[†]
University of Victoria

Pourang Irani[‡]
University of Manitoba

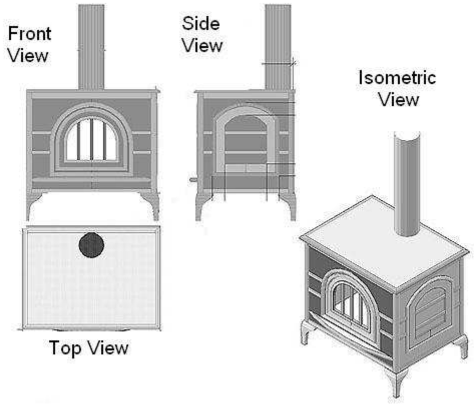

Figure 1: *Multiple 2D views are often used to discern 3D structure. Alternatives include user defined rotations or animated rotations of the 3D structure. Adapated from http://www2.arts.ubc.ca/TheatreDesign/crslib/drft_3/vw0.htm*

Abstract

Visualizations of 3D spatial structures use various techniques such as user controlled interactions or 2D projection views to convey the structure to users. Researchers have shown that motion cues can help assimilate the structure of 3D spatial data, particularly for discerning occluded parts of the objects. However, motion cues or smooth animations also have costs - they increase the viewing time. What remains unclear is whether any one particular viewing modality allows users to understand and operate on the 3D structure as effectively as a combination of 2D and 3D static views. To assess the effectiveness of understanding 3D structures, we carried out three experiments. In all three experiments we evaluated the effectiveness of perceiving 3D structures with either self controlled interactions, animated transitions, and 2D+3D static views. In the first experiment, subjects were given a task to estimate the relative distances of objects in a 3D scene. In the second experiment, subjects made judgements to discern and identify the existence of differences between 3D objects. In the third experiment, participants were required to reconstruct a 3D spatial structure based on the 3D models presented to them. Results of the three experiments reveal that participants were more accurate and performed the spatial tasks faster with smooth animations and self-controlled interactions than with 2D+3D static views. Our results overall suggest that the costs involved in interacting or animating a 3D spatial structure are significantly outweighed by the perceptual benefits derived from viewing and interacting in these modes of presentation.

Keywords: spatial visualization, 2D, 3D, projection, motion, interaction, rotation

1 Introduction

Visualization experts have long argued that users should be allowed to interactively rotate three-dimensional (3D) spatial scenes, and this is now common practice in spatial visualization tools. This interactivity serves many purposes: it enables the user to obtain a different view suitable to the task, reveals structures that may otherwise be occluded, and allows better perception of rigid 3D structure, through structure-from-motion, also called the kinetic depth effect [Wallach 1953].

For some tasks, it is clear that 3D motion is important. For example, understanding the shape of an unusual surface would be difficult from a static scene, even with shading cues. When shading cues are unavailable, motion cues are even more important [Ware 2006]. But despite its well-known advantages, interactive rotation also takes time. When users need to rotate the scene to find a good view, overall performance could be slower. Perhaps a combination of static views, such as multi-view projections common in computer aided design, or a fixed (non-interactive) rotation may lead to faster performance. In addition, some researchers have argued that two-dimensional (2D) views [St. John et al. 2001] or combinations of 2D and 3D views [Tory et al. 2006] may be better than 3D displays for precise relative positioning, although the majority of these experiments used static scenes. Thus it is not clear whether an interactive 3D view is necessarily best for all tasks, especially when precision is important.

We designed three experiments to compare a static 2D/3D combination display to 3D displays with motion, for tasks related to understanding the 3D nature of a scene. We particularly wanted to investigate users' abilities to judge precise relative distances, which can be difficult from static 3D scenes and have been shown to benefit

---

[*] e-mail: taylor.sando@gmail.com
[†] e-mail:mtory@cs.uvic.ca
[‡] e-mail:irani@cs.umanitoba.ca

APGV 2009, Chania, Crete, Greece, September 30 – October 02, 2009.

from combination displays. We believed that motion might enable precise distance judgments without requiring multiple views. We also compared fixed rotation to interactive (user-controlled) rotation to examine whether any difference in speed and accuracy might result from interactivity. Our results reveal that animated transitions and interactive rotations are far more effective than 2D/3D static views alone. These results were upheld for simple tasks suchs as relative distance judgement and also for complex tasks, such 3D structure formation. Contrary to the belief that animations may slow down performance, we found that in this mode of operation participants were faster at creating a mental structure of the 3D scene in comparison to a 2D/3D static view only.

2 Related Work

We briefly review work related to viewing 3D structures and on the interrelated benefits of animated displays.

2.1 Viewing 3D Structures

Numerous experiments have compared various 3D cues such as shading, binocular disparity (from stereo vision), and motion cues to determine their relative importance in understanding 3D scenes. For example, Ware [2006] found that motion, stereo viewing, and shading cues all helped users to understand orientation of 3D streamlines. The largest improvement was caused by rendering the streamlines as tubes to provide shading information, but motion led to a larger improvement than stereo. Earlier studies comparing motion and stereo cues found similar results: the best performance was seen with motion and stereo together, but motion contributed more than stereo [e.g., Ware et al. 1993, Sollenberger and Milgram 1993]. Results tended to be similar whether the object was rotated (causing the kinetic depth effect) or the user's head moved (causing motion parallax).

Another method for viewing 3D structure that has shown to be effective is to combine 2D and 3D views of the same scene into a single display [Tory 2003]. One such method, called ExoVis, uses three 2D views, each oriented along the major axes of the one 3D view. ExoVis has shown to be effective against other combination techniques [Tory and Swindells 2003]. For some tasks, particularly those requiring precise relative positioning in 3D space, combinations of 2D projections or slices and 3D views have been shown to benefit performance compared to static 3D views alone [e.g., Tory et al. 2004; Tory et al. 2006].

Such combination displays are becoming fairly common in practice [e.g., Tresens and Kuester 2003; Brooks and Whalley 2005]. However, Velez [2005] showed that a user's ability to understand 2D projections of a 3D scene depends on spatial ability, suggesting that combined 2D/3D views may not be suitable for all users. We wished to examine whether 3D rotation may achieve the same precision benefit as 2D/3D combination displays. The answer to this question is not obvious. While it is clear that rotation is an important cue to qualitatively understanding 3D scenes, it is not obvious that they improve measurable task performance. For instance, Liter et al. [1993] found that rotations (consisting of 30 static views) did not improve performance at a structure recovery task compared to two-view displays, even though users reported that it was difficult to perceive depth with only two views.

2.2 Benefits of Animated Displays

Perceptual constancy in terms of animated displays, which was originally defined in Robertson et al. [1993], defines a perceptual phenomenon that preserves the identity of an object, even when it is seen under varying conditions such as perspective, distance or lighting. One of the goals of animation is to facilitate a persons' understanding of a given object‘s true structure, and its relation to its surrounding, while promoting perceptual constancy. For example, Bederson et al. [1996] found that zoomable user interfaces allow users to view different levels of detail by zooming in or out. This process was facilitated through animation, which gave users a better conceptual model of the topological details, allowing them to interact more effectively with the display.

In addition to providing information about structure in the world, animation can also facilitate the understanding of some difficult to visualize conceptual phenomenon. For example, Sanger et al. [2000] found that students had a better conceptual understanding of the behaviour of gas particles during the crushing of a can. Previously, the students held various misconceptions about the process, but these were dispelled after the animation demonstration. Singer et al. [2001] also found that animation helped students to understand the processes of osmosis and diffusion. McClean et al. [2005] found that students retention of various biological phenomenon, such as the electron transport chain, RNA/DNA transcription and translation, was increased by animation. Burd et al. [2002] showed that animations improved the comphrension of UML sequence diagrams. Animations have also been shown to give students some improvement in learning computer alogirhtms[e.g. Bryne et al. 1996; Kehoe et al. 2001; Stasko et al. 1993].

In a review article, Tversky and Morrison [2002] suggest that interacting with graphics may improve student learning, but that the benefits of motion are less clear. They report that motion and interactivity were rarely considered as separate experimental factors in the studies they reviewed. Ware and Franck [1996] found that motion was helpful but the method for producing the motion (automatic or user controlled) had little effect. However, a recent study by Keehner et. al. [2008] found that animation was more important than interaction. In that study, participants were asked to draw cross sections of 3D structures. They found that participants who viewed a visualization that had been manipulated in an effective manner performed as well as participants that had the ability to interact with the animation.

One of the goals of this paper is to attempt to demarcate the benefits between interaction and animation. The paper will expand on what Keehner et. al. [2008] did by examining other tasks related to spatial reasoning. The second goal is to compare these benefits to a multiple static 2D/3D views. In total, there were three experiments, each of which had three display conditions. The display conditions used were a static animation condition, a user controlled animation and an ExoVis or combined 2D/3D static view. The static animation condition meant that users could pause, but could not otherwise alter the rate, direction or axis of rotation. In the user controlled animation conditions, the user could alter the direction, rate and axis of rotation. In the ExoVis view, the user could not alter the display. Experiment 1 involved the judgement of distance between three objects in a scene. Experiment 2 examined peoples' ability to compare two similar scenes. Experiment 3 was a block assembly task, wich required users to understand the scene using the display, and to reproduce the scene in the real world using blocks. These tasks were intentionally chosen to range from low-level perceptual operations to higher-level activities.

3 Experiment 1: Relative Judgement of Distance

One of the most important abilities related to using 3D models is the ability to estimate the distance between objects in the scene. The first experiment was designed to assess this skill. The objective

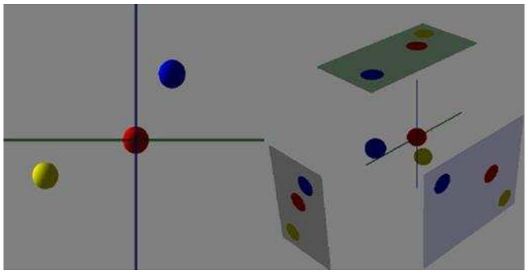

Figure 2: *The UCA and SA condition is shown on the left. The ExoVis condition is shown on the right. Note, The ExoVis condition has been scaled down slightly to fit.*

was to determine which presentation method resulted in the most effective performance.

3.1 Methods

3.1.1 Participants

Twenty-two participants, 15 male and 7 female from a university population volunteered to participate. Participants received 1% towards their final grade in an introductory computer science course in exchange for their participation.

3.1.2 Task

Three spherical balls of equal volume were presented on the computer screen, each with its own color (red, yellow and blue). Each ball had a center point defined in three dimensions. Each ball had a diameter of approximately 1.5-2cm in each of the presenatation conditions. The exact size of the ball on screen was variable as the camera was set to a perspective projection.

The red ball represented the target at the origin. It was the participant's task to identify which ball was closer to the target ball: the yellow ball, or the blue ball. The distance between the target ball and the other two balls, was between 1 and 3 cm. The distance between the yellow and blue balls was not equal. One of the balls was closer to the target ball, while the other ball acted as a distractor. Figure 2 shows an example. The yellow ball was closer to the red ball in only half the trials.

3.1.3 Design and Procedure

We used a 3x4 factorial design. The first factor was the method of presentation. There were three different presentation methods: the ExoVis (EV) method [Tory and Swindells 2003], a static animation (SA), and a user controlled animation (UCA). In the SA condition, the camera rotated at a fixed angle around the Y-Axis. The user was able to pause the animation using the space key on the keyboard, in order to select one of the targets. The rate of rotation was 5 degrees every 55 milliseconds. In the UCA condition, the user could control the axis of rotation. The participant could rotate around the X- or Y-axis using the up/down arrows for the X-Axis, or the left/right arrows for the Y-Axis, on a standard keyboard. Rotation occurred in a step-wise fashion when the buttons were pressed, but was continuous if the button was held down. In all conditions, the participant attempted to select the closest ball using the mouse.

The second factor used in the experiment was the difficulty level. The distance between the closest ball and the distractor was used to define the level of difficulty. There were four individual difficulty conditions (D1, D2, D3 and D4). D1 was considered the easiest condition, while D4 was considered the hardest condition. The distances used were 1.5cm for D1, 0.75cm in D2, 0.50cm in D3 and 0.25cm in D4. These distances were determined during pilot studies and were absolute distances.

A total of 40 ball configurations were used in the study, with 10 configurations for each difficulty level. The configurations were presented in random order until all 40 had been shown for each presentation condition. Participants were divided into three groups to control for any possible ordering effects. The first group completed the EV, SA and UCA conditions, first, second and last, respectively. The second group completed UCA, EV and then SA. The third group completed SA, UCA and EV. In total, each participant completed 120 trials.

Each participant was presented with a screen that prompted them to enter a user id. After entering this number, and using the mouse to click on the start button, participants were immediately presented with practice trials in one of the three presentation conditions. The configurations of the balls in the practice trials were random.

Participants used the mouse to select the ball that they felt was closest to the target red ball. Participants were asked to answer as correctly as possible. After they had selected the ball they thought was closest, they were informed on-screen, whether they had been successful or not and then the next trial started. Participants were allowed to complete as many practice trials as they felt was necessary to learn how to interact with the system. Generally, this was between two and five trials. When they felt they were ready, participants pressed the ‘s’ key, and the real experimental trials began. During the experiment, participants were still informed whether they were correct or not.

When a participant had finished all the trials for their first condition, they were then presented with the practice trials for the second presentation condition. Again, participants completed as many practice trials as they felt was necessary, and pressed the ‘s’ key to begin the real trials. When they completed the real trials from the second presentation condition, they were presented with the practice trials for the third condition. The participants completed the practice and then the real trials.

3.1.4 Measures

We measured two dependent variables. The first variable, success rate, was simply the percent of correct responses, while the second variable, elapsed time, was the time to complete the trial. The trial time started immediately when participants were shown a given scene, and ended when they made a selection.

3.2 Results

Data was found to be normally distributed and not in violation of the sphericity assumption. Separate two-way ANOVA tests were used to assess the effects that the presentation and difficulty conditions had on the elapsed time, and the success rate of the tasks. P-values less than 0.05 were considered to be significant. The margin of error is reported as +/- 2 SE.

3.2.1 Elapsed Time

Figure 3 summarizes the data. There was a significant main effect for the presentation conditions, $F(2, 42) = 20.083$, $p < .001$. A Bonferroni correction was used to assess the differences between presentation conditions. SA was significantly faster than both EV, $p < 0.001$ and UCA, $p < 0.001$. UCA was significantly faster than EV, $p < 0.001$. The average completion time for SA was 5.72 secs +/- .4 sec, for UCA was 8.92 secs +/- 1.1 sec, and for EV was 9.85

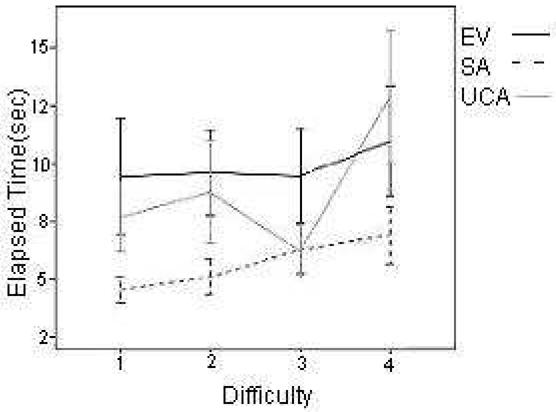

Figure 3: *Mean elapsed time that participants took to complete the relative distance task using the presentation methods and the difficulty conditions. Margin of error is +/- 2 SE.*

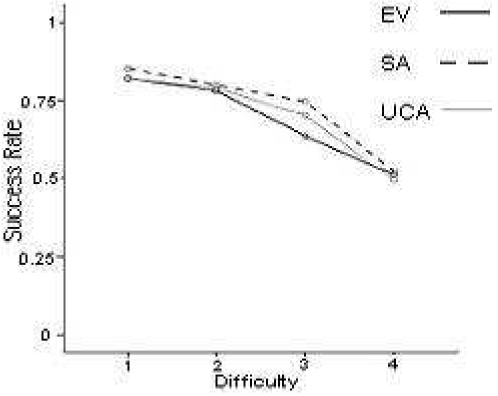

Figure 4: *Mean success rate of participants for the relative distance task using the presentation methods and the difficulty conditions. Margin of error is +/- 2 SE.*

secs +/- 1.1 sec. On average participants were 1.7 times faster with SA than with UCA or than EV for this task. There was also a significant main effect for the difficulty conditions, F(3, 63) = 6.227, $p < 0.01$. A Bonferroni correction was to assess differences between difficulty conditions. It was shown that the most difficult condition, D4, took significantly longer (10 secs versus 7 secs than all the other conditions), $p < 0.01$. D1, D2 and D3 did not significantly differ from each other. There was no evidence of an interaction effect between presentation and difficulty, F(6, 126) = 1.370, p = .227.

3.2.2 Success Rate

Figure 4 summarizes the data. There was no significant difference between the presentation conditions, F(2, 42) = 2.086, = .124. There was a significant difference between the difficulty conditions, F(3, 63) = 65.763, $p < .001$. A Bonferroni correction was used to assess the differences between the difficulty conditions. D1 and D2 had similar accuracy values and were not significantly different from each other, p = .637. Participants were significantly less accurate with D3 when compared to the first two difficulty levels, $p < 0.001$. Participants were significantly less accurate with D4 when compared with D3. There was no interaction effect between presentation and difficulty, F(6, 126) = .562, p = .761.

4 Experiment 2: Structure Comparison

Being able to quickly and accurately assess what, if any, differences exist between 3D structures is an important ability. It is common for medical professionals to look at two different medical images and attempt to determine if anything has changed. For example, whether a cancer has grown, shrunk or stayed the same. Therefore, the second experiment was a task designed to test a participants abilities to discern the difference between two similar but slightly differing 3D models.

4.1 Methods

4.1.1 Participants

Fifteen participants, 9 male and 6 female from a university population volunteered to participate in the study. Participants received credit towards their final grade in an introductory Psychology course in exchange for their participation.

4.1.2 Task

Participants were presented with two different 3D configurations on a screen. Each 3D configuration was made up of 10 blocks that were randomly assembled. Blocks varied in both their size and their color. The colors used were red, blue or yellow.

Both configurations were presented simultaneously with the same display type: EV, SA or UCA. The left side acted as the source, while the right side acted as the target. The target configuration differed from the source in the number of blocks that changed between the two. There were two sources of difference. The first is that a target block could have a different position. The second is that a target block could have a different color. Figure 5 shows an example of the SA/UCA conditions and the EV conditions. The target configuration either had 0,1,2, or 3 differences when compared to the source configuration. When the target configuration had 0 differences, it meant that the target and source configurations were actually the same.

In the UCA condition, participants were allowed to rotate the camera left or right using the left and right arrow keys of the keyboard. When rotating the camera, both configurations were rotated by the same degree. In the SA condition, both models rotated automatically at the same degree. The space bar acted as a toggle for rotation, either pausing rotation, or resuming it. There was no rotation in the EV condition.

4.1.3 Design and Procedures

Participants had a total of 96 trials, with 32 trials for each presentation condition. In those 32 trials, each participant was exposed to 8 trials where the target and the source were the same, 8 trials where the target differed from the source in one way, 8 trials where the target differed from the source in two ways and 8 trials where the target differed from the source in three ways. The order that trials were presented was random.

Participants were split into three groups to control for ordering effects. The first group completed the EV, SA and UCA conditions, first, second and last, respectively. The second group completed UCA, EV and then SA. The third group completed SA, UCA and then EV.

It was the participants task to identify whether the target configuration differed from the source configuration. If the target was different from the source, then the participant had to identify the differences by clicking on the target blocks that were different than

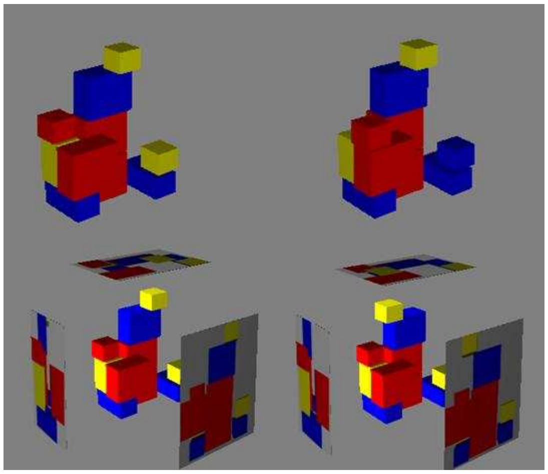

Figure 5: *At the top is the SA/UCA condition. The target side differs from the source in two ways. The yellow block in the lower right is now blue. The red block in the upper left has been moved forward. At the bottom is the EV condition. On the right, the yellow box at the top has been moved to the left.*

the source blocks. Clicking was done with the mouse. A trial ended when the participant pressed the enter button. The next trial began immediately.

Participants completed up to five practice trials before completing the 32 trials for each of the presentation condition. After completing the first presentation condition, participants then moved onto the second presentation condition, followed by the third presentation condition.

4.1.4 Measures

Participants were measured on two dependent measures. The first measure was elapsed time, which was the time from a participant beginning a trial, to the time the participant pressed the enter button to indicate that they were done. The second measure was the number of errors made during the trial. An error could occur in two ways. The first was that a user clicked on a target block that was not actually different from the source block. The second way was that a user failed to click on a target block that was different from the source block. The number of errors that each participant made was divided by the total number of trials per presentation condition to give the error rate.

4.2 Results

Data was found to be normally distributed and not in violation of the sphericity assumption. Separate One-Way ANOVA tests were used to assess the effects that the presentation had on elapsed time and error rate. P-values less than 0.05 were considered to be significant. Margin of errors were reported as +/- 2 SE.

4.2.1 Elapsed Time

Figure 6 summarizes the data. There was a significant main effect for the presentation conditions, F(2, 14) = 5.25, $p < .05$. A bonferroni correction was used to assess the differences between presentation conditions. SA and UCA were significantly faster than EV $p < 0.05$. SA and UCA were not significantly different from each other. Mean elapsed times for SA was 23.1 secs +/- 2.8 sec. UCA was 23.9 secs +/- 3.6, and EV was 29.9 secs +/- 2.1 sec.

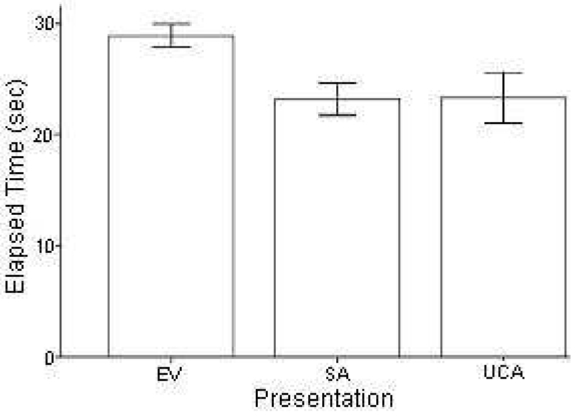

Figure 6: *Mean elapsed time participants took to complete the contrast task using the different presentation methods. Margin of error is +/- 2 SE.*

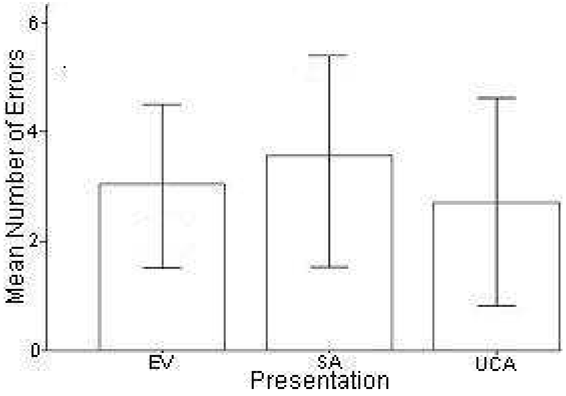

Figure 7: *Mean number of errors made by the participants in the contrast task for each presentation method. Margin of error is +/- 2 SE.*

4.2.2 Error Rates

Figure 7 summarizes the data. Error rate did not significantly differ between the presentation conditions.

5 Experiment 3: Reproducing 3D Visual Structure

The ability to view and manipulate 3D models that represent the real world is only useful if you can apply the knowledge gained to the real world. One simple task to test this ability is to see how well people can reproduce a 3D model in the real world. Therefore, the third experiment was another task to assess user's spatial reasoning abilities. This time participants viewed 3D models of mega blocks on a computer screen and attempted to reconstruct them using actual blocks.

5.1 Methods

5.1.1 Participants

Twelve participants, 7 male and 5 female from a university population volunteered to participate in the study. Participants received

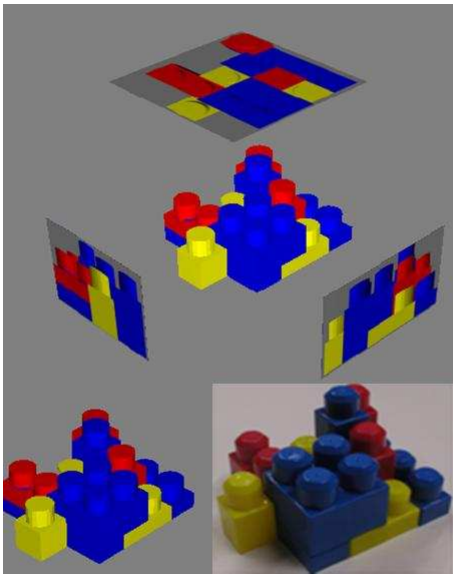

Figure 8: *At the top, the ExoVis condition is shown. The bottom left is the SA and UCA condition. The bottom right is the actual model that participants made.*

1% towards their final grade in an introductory Computer Science course in exchange for their participation.

5.1.2 Task

In this experiment, participants were presented with a configuration of blocks on a screen. The on-screen configuration acted as the instructions for assembling blocks in the real world. As with experiment one, the method of presentation acted as a factor in the experiment. The presentation methods were again EV, SA and UCA. There were six different block configurations, and participants completed 6 trials, one for each configuration. There were two block configurations per presentation method. The assignment of block configuration to presentation method was random. In order to counteract any kind of learning or ordering effect, participants were split into three groups. The first group was presented with EV, SA then UCA. The second group was presented with UCA, EV, SA. The third group was presented with SA, UCA and EV.

5.1.3 Design and Procedures

Participants were presented with the 10 necessary blocks on their desk. When participants were ready to begin the trial, they selected an on-screen start button. They were then presented with the block configuration on screen. Participants then began assembling the blocks. When they felt like they had accurately assembled the blocks, they selected the stop button on screen. The experimenter then rated the participant's accuracy. The elapsed time was the time from when the participant selected start, to the time they selected stop. When the trial was completed, the blocks were disassembled, and the next trial could begin. When the participant was ready for the next trial, they selected the start button. Participants continued in this manner until they had completed all six trials.

5.1.4 Measures

Participants were measured on two dependent variables: elapsed time and accuracy. Accuracy was defined on an ordinal scale and was assessed by the experimenter after the participant had completed the trial. A four represented perfect accuracy, meaning the participant had assembled the blocks exactly as shown on screen. A three represented good accuracy, meaning that at most, one or two blocks were out of place. A two represented poor accuracy, meaning that between two and four pieces were out of place. A one represented a failure, meaning that five or more pieces were out of place.

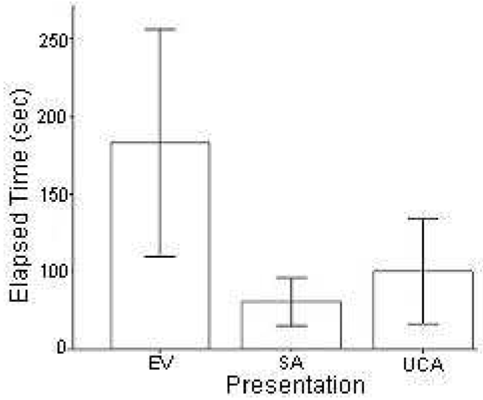

Figure 9: *Mean elapsed time participants took to complete the block assembly task using the presentation methods and the difficulty conditions. Margin of error is +/- 2 SE.*

5.2 Results

A Brown-Forsythe test was used to assess the effect of presentation on elapsed time. A Kruskal-Wallis H-test was used to assess the effect that presentation had on accuracy. P-Values less than 0.05 were considered to be significant. Margins of errors were +/- 2 SE.

5.2.1 Elapsed Time

Figure 9 summarizes the data. There was a significant difference between the presentation conditions, $F(2, 69) = 5.553$, $p < 0.01$. A Dunnet's T3 correction was used to assess the differences between the presentation conditions. SA and UCA were not significantly different, but both were significantly faster than EV, $p < 0.05$. The mean elapsed times for SA was 80 sec +/- 14 sec, for UCA was 100 sec +/- 32 sec. For EV was 180 +/- 72 sec.

5.2.2 Accuracy

Figure 10 summarizes the data. There was a significant difference between the presentation conditions, $X^2(2) = 8.56$, $p < 0.05$. Separate Mann-Whitney U tests were used as post hoc tests to assess the differences between the groups. SA had significantly higher accuracy than EV, $p < 0.01$. UCA was not significantly different than EV, but suggested a trend, $p = 0.058$. SA and UCA were not significantly different from each other, $p = 0.301$.

6 Discussion

6.1 Relative Distance

Time to complete the relative distance task was fastest with the SA condition, followed by UCA and then ExoVis. This suggested that animation provided 3D information to participants in a shorter amount of time than a combination 2D/3D display.

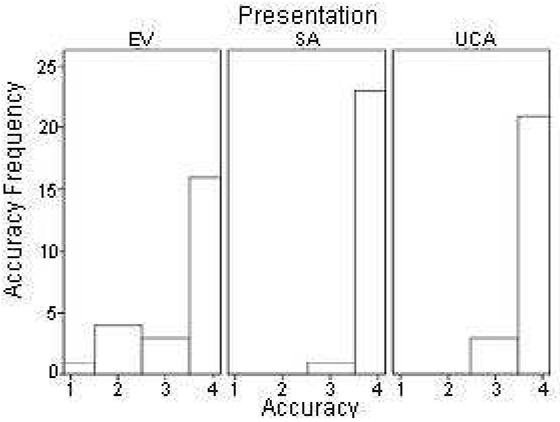

Figure 10: *Histogram of the accuracy in the block assembly task. The x-axis represents the accuracy rating, and the y-Axis represents the frequency of that accuracy rating.*

It is not entirely clear why the SA condition was faster than UCA for this task. Both conditions provide information through structure from motion. One possible explanation could be in the rate of rotation. In the SA condition, the rate of rotation was fixed at about 90 degrees a second. However, in the UCA condition, the user had more control over the rate of rotation. A single press of the rotate arrow key rotated the scene 10 degrees. The speed of rotation was dependent on how often the user pressed the rotation button. It could reach up to 180 degrees a second if the user held the key down. Participants could rotate along both the X- and Y-Axis, and in both directions on these axes. When participants encountered the UCA condition first, they tended to tap the rotational keys, making slight rotations, rather than holding the key down, which mimicked the SA condition. However, when users were presented with the SA condition before they were confronted with the UCA condition, they tended to mimic the SA condition by holding down the rotational keys. This suggests that the rate of rotation, and therefore, animation, may play a role in the ability to spatially reason.

The method of presentation did not seem to affect the success rate, as all conditions showed the same level of accuracy. Accuracy was primarily influenced by the difficulty level. At the easiest level, success rate was just over 75%. At the hardest level, success rate dropped to 50%, or no better than chance. Participants were slowest with the hardest difficulty level, but showed similar times for the other three levels. The range of success rates demonstrates that our experiment captured difficult or precise positioning tasks (not just coarse positioning).

There was no interaction effect between presentation method and difficulty, suggesting that the rate of success was mediated by difficulty level and that the response time was mediated by the presentation condition. Since the animation conditions showed the same accuracy as the 2D/3D display, but required less time to solve the task, our results suggest that tasks requiring distance judgements would benefit from animation.

6.2 Contrast

Participants took the least amount of time to identify differences between two 3D models when they used the SA and UCA methods. The EV condition was slightly slower, by about 7 seconds, or about 29% slower. Participants showed similar error rates (3-4%) for all the presentation conditions.

6.3 Block Assembly

Results were similar to the previous experiments. Participants assembled the blocks quicker when they were presented with the animated conditions, rather than the 2D/3D displays. There was no significant difference between the UCA and SA conditions, but the UCA condition showed a greater amount of variation in accuracy when compared to the SA condition. The reason for this is unknown, but could be attributed to the rate of rotation, for reasoning similar to that of the relative distance task.

In nearly all cases, participants assembled the blocks in the least amount of time, and most accurately, when they were presented with the SA conditions. However, it is interesting to note that participants showed very large variance in the amount of time it took to complete the building task with ExoVis. This meant that some people completed the task as fast as the animated conditions, but others took nearly four times as long to do so. One possible explanation for this observation is that the users differed in spatial ability, since it is known that ability to interpret 2D projections of block-shapes is correlated with results on spatial ability tests [Velez 2005]. We cannot be sure of this since our participants did not complete any tests on spatial ability. However, it is interesting that spatial ability differences may still strongly impact interpretation of ExoVis 2D / 3D displays, even though they were designed to minimize the difficulty of mentally registering multiple views.

6.4 Overall

Our results at least partially counter previous claims that 2D or 2D / 3D displays are better than 3D alone for precise relative positioning. It seems that 3D displays are sufficient as long as the scene can rotate, either with static animation or user control. Note that 2D displays may still be better when the relative position of objects can be captured accurately in only one 2D view. In this case, the user would only need to look at one static view. However, in many practical applications, objects are not aligned with a standard view direction (e.g., structures in the human body rarely differ in position only by one medical imaging axis). To obtain a useful 2D view in this case, the user would first have to position a view plane or slicing plane that intersected or ran parallel to the axis between the objects. Since orienting this plane would be quite time-consuming, the 3D view with animation may be a better overall solution.

We should note that our results only apply when the user can understand a scene through object surfaces and does not need to look inside objects. Viewing interior structures requires other techniques such as 2D slicing planes. Also, axis-aligned views (like the 2D views in our experiments) may be important for understanding whether or not objects align with the axes. Rotatable 3D displays could be augmented with shortcut keys to access these pre-defined viewpoints.

Based on the results of our three experiments, we offer the following recommendations: (1) static or interactive rotation should be an allowable operation in 3D systems since it increases the speed at which a 3D scene can be understood, (2) for some time-critical tasks, fixed animation may be a better choice than user-controlled animation, (3) multiple static views may not be very effective for understanding 3D objects, unless the interior of objects needs to be shown, and (4) if animation is not feasible for some reason, then a static ExoVis display can be used; users can effectively understand 3D ExoVis scenes, with increases in performance time but little change in accuracy.

7 Conclusion and Future Work

We presented three experiments comparing 3D views with static or interactive rotation to static 2D / 3D combination displays. Our results indicate that spatial positioning tasks can be accomplished faster and with equal or greater accuracy using rotatable 3D displays as compared to static 2D / 3D displays. The method of rotation (static vs. interactive) was less important, though in some cases static animation was faster. These results suggest that 2D / 3D combination displays may not be needed for precise relative positioning tasks, except when there are other motivating factors (e.g. occlusion).

Future work could examine how spatial ability and training (such as CAD training) correlate with the ability to interpret rotatable 3D displays and 2D / 3D combination displays. We also suggest conducting similar experiments with 2D slice views since we considered only 2D projection views. Slice views are more applicable when the scene has a lot of occlusion, as in many medical imaging data sets. Finally, our work has examined the low-level mental operation of relative positioning. It will be important to determine the relative strengths and weaknesses of 2D, 3D, and combination display designs in more complex higher-level tasks.

References

BEDERSON, B. B., D., H. J., PERLIN, K., MEYER, J., BACON, D., AND FURNAS, G. W. 1996. Pad++: A zoomable graphical sketchpad for exploring alternate interface physics. *Journal of Visual Languages and Computing 7*, 1, 3–32.

BEDERSON, B. B., HOLLAN, J. D., PERLIN, K., MEYER, J., BACON, D., AND FURNAS, G. W. 2001. Rethinking the evaluation of algorithm animations as learning aids: an observational study. *International Journal of Human Computer Studies 54*, 265–284.

BROOKS, S., AND WHALLEY, J. 2005. Pad++: A zoomable graphical sketchpad for exploring alternate interface physics. In *Proceedings of the 3rd international conference on Computer graphics and interactive techniques in Australasia and South East Asia*, 323–330.

BURD, E., OVERY, D., AND WHEETMAN, A. 2002. Evaluating using animation to improve understanding of sequence diagrams. In *Program Comprehension, 2002. Proceedings. 10th International Workshop*, 107–113.

BYRNE, M. D., CATRABONE, R., AND STASKO, J. T. 1996. *Do Algorithm Animations Aid Learning?* GVU Technical Report.

KEEHNER, M. AND HEGARTY, M. AND COHEN, C. AND KHOOSHABEH, P. AND MONTELLO, D. R. 2008. Spatial reasoning with external visualizations: What matters is what you see, not whether you interact. *Cognitive Sciences 32*, 1099–1132.

LITER, J. C., BRAUNSTEIN, M. L., AND HOFFMAN, D. D. 1993. Inferring structure from motion in two-view and multiview displays. *Perception 22*, 12, 1441–1465.

M., S., BRECHEISEN, D. M., AND HYNEK, B. M. 1993. Can computer animations affect college biology students' conceptions about diffusion and osmosis? *The American Biology Teacher 63*, 2, 104–109.

MCCLEAN, P., JOHNSON, C., ROGERS, R., DANIELS, L., REBER, J., SLATOR, B. M., TERPSTRA, J., AND WHITE, A. 2005. Molecular and cellular biology animations: development and impact on student learning. *Cell Biology Education 4*, 2, 169–179.

ROBERTSON, G. G., CARD, S. K., AND MACKINLAY, J. D. 1993. Information visualization using 3D interactive animation. *Communications of the ACM 36*, 4, 56–71.

SANGER, M. J., PHELPS, A. J., AND FIENHOLD, J. 2000. Using a computer animation to improve students' conceptual understanding of a can-crushing demonstration. *Journal of Chemical Education 77*, 11, 1517.

SOLLENBERGER, R. L., AND MILGRAM, P. 1993. The effects of stereoscopic and rotational displays in a three-dimensional path-tracing task. *Human Factors 35*, 3, 483–500.

ST. JOHN, M., COWEN, M. B., SMALLMAN, H. S., AND OONK, H. M. 2001. The use of 2D and 3D displays for shape-understanding versus relative-position task. *Human Factors 43*, 1, 79–98.

STASKO, J., BADRE, A., AND LEWIS, C. 1993. Do algorithm animations assist learning?: An empirical study and analysis. In *Proceedings of the INTERACT '93 and CHI '93 conference on Human factors in computing systems*, 61–66.

TORY, M., AND SWINDELLS, C. 2003. Comparing exovis, orientation icon, and in-place 3D visualization techniques. In *Graphics Inteferface Proceedings*, 57–64.

TORY, M., ATKINS, M. S., MOELLER, T., AND KIRKPATRICK, A. E. 2004. Combining 2D and 3D views for orientation and relative position tasks. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 73–80.

TORY, M., KIRKPATRICK, A. E., ATKINS, M. S., AND MOELLER, T. 2006. Visualization task performance with 2D, 3D, and combination displays. In *IEEE Transactions on Visualization and Computer Graphics*, vol. 12, 2–13.

TORY, M. 2003. Mental registration of 2D and 3D visualizations (an empirical study). In *Proceedings of IEEE Visualization*, 371–378.

TRESENS, M. A., AND KUESTER, F. 2003. Hybrid-reality: A collaborative environment for biomedical data. In *NSF Lake Tahoe Workshop on Collaborative Virtual Reality and Visualization*.

TVERSKY, B., MORRISON, J. B., AND BETRANCOURT, M. 2002. Animation: can it facilitate? *International Journal of Human-Computer Studies 57*, 247–262.

VELEZ, M., SILVER, D., AND TREMAINE, M. 2005. Understanding visualization through spatial ability differences. In *Proceedings of IEEE Visualization*.

WALLACH, H., AND O'CONNE U, D. N. 1953. The kinetic depth effect. *Journal of Experimental Psychology 45*, 205–217.

WARE, C., AND FRANCK, G. 1996. Evaluating stereo and motion cues for visualizing information nets in three dimensions. In *ACM Transactions on Graphics*, 121–140.

WARE, C., ARTHUR, K., AND BOOTH, K. S. 1993. Fish tank virtual reality. In *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems*, 37–42.

WARE, C. 2006. 3D contour perception for flow visualization. In *Applied Perception in Graphics and Visualization*, 101–106.