# Impact of Group Size on Spatial Structure Understanding Tasks

Taylor Sando[1], Melanie Tory[2], and Pourang Irani[1]
[1]University of Manitoba, [2]University of Victoria

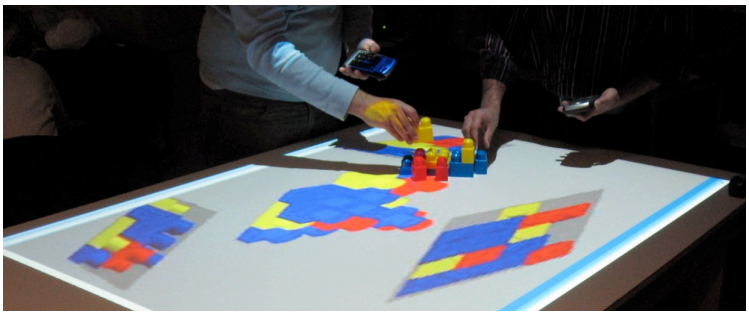

Figure 1. Example display and experimental setup used in our construction task (Experiment 3). When a visual representation must be related to the physical world, multiple 2D views of the same 3D structure can facilitate a group understanding of the 3D spatial structure. Group size affects the performance of the collective unit for two different presentation styles.

**ABSTRACT**

Co-located collaborative tasks allow teams to leverage the skills of each individual member. While numerous guidelines exist to develop visualizations for individuals working on desktops, very little is known about how groups of individuals interpret and comprehend diverse types of visual constructs on larger displays. To study whether group size impacts the collective understanding of relationships in three-dimensional (3D) spatial structures when using different types of presentation, we carried out three experiments. We compared individual performance at structure understanding tasks to performance of groups containing two or four members. We consider two alternate visualization techniques for extracting 3D structure information: a 3D view with animated rotations and a combination of one static 3D plus three static two-dimensional (2D) projection views. In general our studies suggest that as group size increases, so does accuracy but with a cost in efficiency. Our results also suggest that beyond a threshold limit in group size, performance on certain tasks begins to degrade. Regardless of group size, participants performed better when the display was presented in the animation condition instead of the multiple static views, except when large groups needed to relate the visualization to a physical counterpart. We summarize our results in terms of Steiner's model for explaining the effects of group size and task characteristics on group performance.

**KEYWORDS:** group size, collaboration, spatial structure, 2D and 3D views, animation, rotation, horizontal display.

**INDEX TERMS:** H.5.2 [User Interfaces]: screen design, evaluation /methodology; H.5.3 [Group and Organization Interfaces]: computer-supported cooperative work.

Email: taylor.sando@gmail.com, mtory@cs.uvic.ca, irani@cs.umanitoba.ca

## 1 INTRODUCTION

Do four eyes see better than two? Mark et al. [12][13] demonstrated that yes, two users can produce more accurate results than one user at both focused and open-ended visualization tasks involving tabular data. But do these results extend to other types of visualizations, such as 3D spatial structures?

Visually comprehending and assessing relationships in 3D spatial structures is critical for tasks such as scanning medical images for anomalies, inspecting drawings of buildings or visually exploring scientific visualizations. Various interaction and visualization techniques (e.g., animated rotations or multiple 2D perspective views) assist in carrying out such tasks, but have generally been designed for a single user. In reality, 3D spatial structures are often examined by multiple users simultaneously. For example, a team of surgeons may need to inspect tumour growth and discuss surgical plans, architects and designers brainstorm novel features with 3D mock-ups, and scientists with diverse skill sets may investigate the causes of natural disasters presented visually.

Although perception of 3D spatial visualizations has been studied extensively, it has examined only single users. When groups of users perform 3D spatial tasks, performance may be affected by factors such as team dynamics, the amount of discussion that takes place and the skills each bring to the group. Little is known about how such elements affect performance, particularly when it concerns discerning structures from 3D spatial representations. We therefore designed three studies to examine group ability to understand 3D structures. These studies focus on the primary experimental factor of group size, which has been found to influence group performance in other applications and might therefore play a significant role. We focus on groups working together around a horizontal display, since it is well established that large screen displays support collaborative work. Horizontal displays in particular are thought to facilitate collaboration and discussion among small teams by encouraging face-to-face contact and equitable interaction with the software.

This paper makes two key contributions. First, it examines how having more than one person working cooperatively on common 3D spatial tasks affects accuracy and efficiency. Based on previous analyses of group work (e.g., [8][12][13]), we hypothesized that groups would take longer, but would have a

higher success rate than individuals. Second, we investigate whether earlier results supporting the benefits of animated displays for 3D structure understanding [17] still hold in collaborative work scenarios. We also hypothesized that there could be an interaction between visualization type and group size, as found previously [13]. For example, it was possible that animation would be better than static views with two people, but not with four. In summary, our goals were to determine if there are benefits to additional people, whether there are limits to those benefits (e.g., hindered performance because of coordination overhead among participants), and how the effects of group size might vary depending on the visual representation.

Results of our first two experiments generally suggest that regardless of group size, participants are more accurate and efficient with the animated 3D display for making relative distance judgments and finding differences in 3D structures. Interestingly, we also find threshold limits depending on the type of task. For making judgments on distances, groups of two were more accurate than either one individual alone or groups of four. However for a more complex task, four participants were more accurate than two or just one alone. Surprisingly, in a final experiment that involved physical construction of a 3D object, groups of four were more efficient with a combination of static views than with the animated 3D view. We hypothesize some potential reasons for this finding in the discussion.

## 2 RELATED WORK

### 2.1 Visualization for Collaborative Use

Collaborative visualization has recently engaged a great deal of research interest. We focus on co-located collaboration, where users work together in the same place at the same time. Much of the work so far in this domain has developed interfaces and interaction techniques to enable smooth mechanics of collaborative work. Particular focus has been placed on large interactive surfaces such as wall and tabletop displays since they are known to facilitate effective collaboration. Example visualization tools designed for interactive surfaces include [7] and [21]. An analysis of requirements for collaborative visualization may be found in [5].

A few previous studies have compared group and individual performance at visualization tasks, typically finding that groups obtain more accurate results than individuals. For example, [12], [13], and [8] reported that groups were more accurate than individuals in studies involving questions about tabular data, but were often slower [12]. Additionally, these studies identified a set of activities involved in the collaborative analysis process. Similarly, [1] found that groups committed fewer errors than individuals for a security screening visual search task, but were slower under some conditions. One possible explanation for this trend is that groups spend longer validating their findings than individuals [8]. The finding that groups outperform individuals is also consistent with studies of computer-based tasks beyond the field of visualization (e.g., [6][16]).

Our own experiments are most similar to [1] since they involve spatial data. We extend the previous work by examining a different type of task (understanding relative positioning of objects) and exploring 3D data rather than 2D. We also examine whether the effect of group size is consistent across visualizations that are known to require different levels of cognitive effort.

### 2.2 Effect of Group Size on Performance

Psychologists have investigated the influence of group size on group dynamics and task performance for over a century (e.g., [18]). While the classic finding is that larger groups produce better results, this view is simplistic; in reality, group performance depends on a variety of factors. A detailed review of this literature is beyond the scope of this paper; however, a widely-cited theoretical model is Steiner's model of group productivity [20]. It describes a group's actual productivity as its potential productivity minus losses due to coordination overhead and decreased motivation. Both potential productivity and process losses tend to increase with increasing group size.

Potential productivity depends on the nature of the task. Steiner's model predicts that productivity for additive tasks (in which individual contributions can be summed to make a group output, such as picking up litter) should increase linearly with group size. By contrast, productivity at disjunctive tasks (which require the group to select one solution among alternatives) should increase due to more alternative suggestions, but with diminishing returns for larger and larger groups. This was verified experimentally [11]. Furthermore, if a task is disjunctive and there are high coordination costs, performance should increase with group size up to some optimum, but then decrease for larger groups due to coordination challenges. This was found to be true if one considered functional group size rather than actual group size [11] , where functional group size was defined as the number of members who actively coordinated with others.

### 2.3 3D Structure Visualization

We focus on visualizations of 3D structures. Previous work in 3D structure visualization has examined factors contributing to effective perception of depth and shape, and benefits of 2D and 3D views as well as their combination. In contrast, our study focuses on comparing group versus individual performance.

Numerous studies have compared 2D to 3D views of 3D scenes as well as their combination. For some tasks, particularly those requiring precise relative positioning, combinations of 2D projections or slices and 3D views have been shown to benefit performance compared to static 3D views alone (e.g., [24]). Understanding the relationships between the various views is easier when mental rotation is not required to align them [22]; hence, the ExoVis [23] technique (where 2D projections are shown at an angle) was chosen in our experiments over standard multi-view projections (where 2D projections are shown flat on the screen). It is also known that a user's ability to understand 2D views of a 3D scene depends on spatial ability [10][25], suggesting that combined 2D/3D views may not be suitable for all users. Moreover, [17] revealed that a rotatable 3D scene could enable precise positioning faster and with equal or greater accuracy than a combination of static 2D and 3D scenes. User control over the rotation was unimportant; both fixed rotation and user-controlled rotation were effective, echoing earlier results [9][26]. It seems that having the appropriate view for a task is more important than physically interacting with the scene [9].

The knowledge that rotation facilitates 3D structure understanding is not new. A variety of 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. Typically, these studies have found that both motion and stereo cues are very helpful for conveying 3D structure, but motion contributes more than stereo [19][27][28]. Results tend to be similar whether the object is rotated (causing the kinetic depth effect) or the user's head moves (causing motion parallax). Perceptual constancy, a phenomenon originally defined by [15], explains that the identity of an object is preserved during rotation as well as following other view changes such as perspective, distance, or lighting.

Our own experiments included both a 3D animated view as well as a combination of static 2D and 3D views. The static combination was included to examine how group performance is affected by a cognitively demanding representation.

### 2.4 Considerations for 3D Visualizations on Horizontal Displays

Wigdor et al. [29] have shown that comparison of some graphical primitives can be less accurate on horizontal displays compared to vertical displays. In addition, researchers have recently taken interest in studying the readability of different projections of a 3D model on a 2D surface [3][14][29]. Hancock et al. [3] proposed various alternatives to represent 3D models on a horizontal surface to minimize some of the distortion effects that result from the projection. A more user-centric approach was proposed to correct perspective of the display by tracking the user's viewpoint and position around the shared surface [14]. Hancock et al. recently showed that errors in judging object orientation increased as the center of projection diverged from the observer's viewpoint and were considerably lower when the center of projection was directly above the table [4]. For this reason, we used a center of projection directly above the table in our experiments. A review and taxonomy of 3D representations on tabletop displays is provided by Grossman and Wigdor [2].

## 3 Overview of Experiments

Our goal was to examine the effects of group size, and secondarily presentation style, on 3D spatial understanding tasks. In the first experiment, participants made a judgment of the relative distances between objects in 3D. Distance judgment is a common task that underlies 3D structure understanding and is necessary for many applications (e.g. to understand whether there is sufficient space between a planned surgical intervention and a brain structure responsible for critical functionality such as speech). In a second experiment, users were required to find differences between two 3D structures. This is common in medical applications, for instance where physicians have to find changes between images taken at different times, or in architectural applications where a user might need to determine whether a colleague has made changes to the positions of doors, closets, or other features. Finally, in the third experiment, the team needed to build a 3D spatial structure with MegaBloks™. This required participants to relate their perspectives of the 3D scene to an equivalent object in the real world, to see if they could translate their understanding of a virtual object to a physical reality.

All experiments required participants to coordinate their efforts to provide a response. Our hypotheses are summarized as follows:

**H1**: responses would be significantly slower as group size increases, consistent with previous studies in other areas [1][12].

**H2**: teams would be significantly more accurate with a larger number of members, consistent with previous studies in other areas [1][6][8][12][13][16].

**H3**: group size and display type would interact. Performance with the two display types (animation and static views) would not change in the same way as group size was increased.

### 3.1 Details Common to all Experiments

We used a large horizontal display measuring 1.8m in width, 1m in height, and 1.3m in length using a projector suspended from the ceiling directly above the table, as shown in Figure 1. The surface did not support touch interaction; however, user interactions were limited to a simple response so this was unnecessary. All scenes used a perspective camera with the center of projection directly above the center of the table. Participants stood around the table and were allowed to move freely.

Different participants took part in each experiment, with 55-56 people in each. Since groups were assigned, group members generally did not know each other. Participants were volunteers from a university student population, mostly aged 18-24, and approximately 60% male and 40% female (exact numbers unavailable). No participants reported difficulty with the tasks due to color vision deficiencies.

Groups had 1, 2, or 4 participants. Due to difficulty in recruiting participants at the same time, we did not include group size 3, and had slightly fewer of the larger group sizes. ANOVA results were calculated using Type III Sum of Squares, which is designed to handle unequal sample sizes.

Participants were instructed to complete the tasks as quickly and accurately as possible. Most experimental sessions lasted 5-10 minutes. We did not analyze learning effects as the designs were all counter-balanced for interface order.

## 4 Experiment 1 - Relative Judgment of Distance

Our previous research [17] examined users' ability to judge relative distances between objects using two representations: an animated 3D view and ExoVis (EV) [23]. The task was generally done quicker and more accurately with animation than with EV on a desktop display. We extend these results by examining how people fared on a large horizontal surface, and how results varied when people completed the task in groups.

### 4.1 Task

As shown in Figure 2, three spherical balls of equal size were presented on the computer screen. Each ball had a center point defined in three dimensions and a diameter of approximately 1.5-2 cm in each condition. The exact size of the ball on the screen varied because the camera was set to a perspective projection. The red ball represented the target. The participants' task was 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 to the yellow and blue balls was always unequal, so that there was always exactly one correct answer. Each was closer in exactly half the trials.

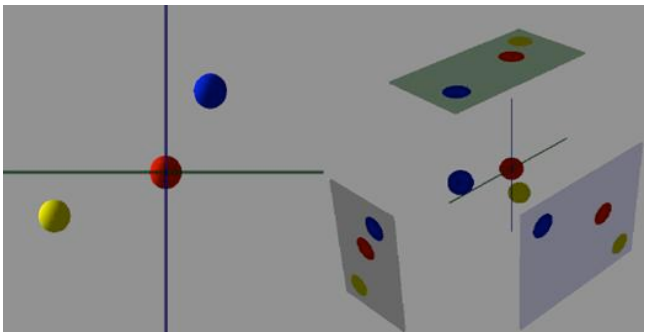

Figure 2. Spatial judgment task. Animation (left) and EV (right). Participants determined which ball was closer to the red ball.

### 4.2 Design and Procedure

We used a 2x3x3 factorial design. The first factor was presentation: ExoVis (EV) or animation. In the animation condition, a single 3D view of the scene was rendered and the camera rotated at a fixed angle around the vertical axis at a rate of 5 degrees every 55 milliseconds. In all conditions, participants attempted to select the closest ball using a keypad.

The second factor was difficulty level, defined as the distance between the closest ball and the distractor. There were three levels of difficulty: easy (1.5 cm), medium (0.75 cm), and hard (0.5cm), as in [17] . A total of 32 ball configurations were used, with 8 for

each difficulty level. Configurations were presented in random order until all 32 had been shown for each presentation condition.

The third factor was the group size. Participants completed the task either alone (N=9), in pairs (N=9), or in a group of four (N = 7). Order of presentation condition was counterbalanced.

Each participant used a keypad to indicate his or her answer. When participants were in pairs or a group of four, they were allowed to converse, and were required to agree before moving on to the next trial. Participants were informed whether they were correct or not. If they were correct, the screen flashed green. If they were incorrect, the screen flashed red.

### 4.3 Measures

We measured success rate (percent of correct responses) and time to complete the trial. Trial time started when participants were shown a scene and ended when they came to a consensus.

### 4.4 Results

For success rate (which we could not assume to follow a normal distribution), Pearson Chi-square tests were used to assess whether there were any differences between the presentation conditions and the number of people per group, and between difficulty level and the number of people per group. For elapsed time, the data was log transformed to satisfy the normality assumption. Separate two-way ANOVAs were used to assess whether there was an interaction effect between the presentation condition and the number of people per group, and between difficulty level and the number of people per group. *P* values less than 0.05 were considered significant. We did not analyze interactions between difficulty and interface since those were already assessed in our previous research [17] (hence the two-way ANOVAs instead of one 3-way). Because success rate was not normally distributed, we could not use multivariate ANOVA.

#### 4.4.1 Success Rate

There was no interaction between the presentation condition and the number of people on the success rate. Figure 3 shows that animation was significantly more accurate than EV overall, $\chi^2(2, 23) = 10.162$, $p < .01$. As shown in Figure 4, the number of people in a group significantly effected the rate of success, $\chi^2$ (2, 23) = 9.958, $p < .01$. Bonferroni-corrected post hoc tests were used to assess individual differences between group sizes. We found that two people were significantly more accurate than either one person or four people. A group of four was not significantly different than a single person. There was also an interaction between the number of people and the difficulty level, $\chi^2$ (3, 15) = 10.281, $p < 0.01$. A pair of people performed significantly better than either an individual or a group of four only at the hardest difficulty level. Other differences were not significant.

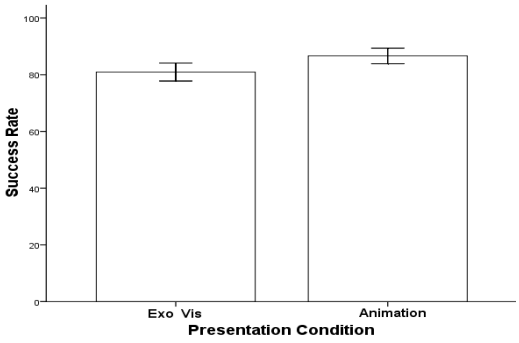


Figure 3. Success rate for each presentation condition. Error bars show +/- 2 SE.

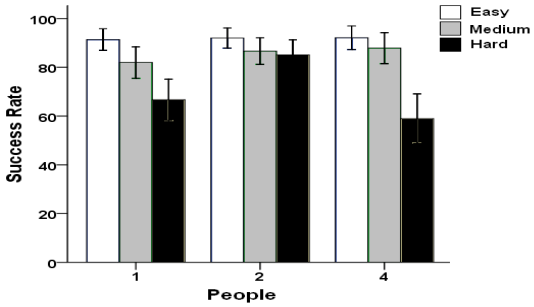


Figure 4. Success rate for each group condition and difficulty condition. Error bars show +/- 2 SE.

#### 4.4.2 Elapsed Time

As shown in Figure 5, there was a significant difference in time between groups of different sizes, F (2, 23) = 21.293, $p < 0.001$. A Bonferroni-corrected post hoc test showed that groups of two and four took significantly longer to complete the task than an individual. Animation was significantly faster than EV, F (1, 23) = 57.261, $p < 0.001$. There was also an interaction between difficulty and group size, F (2, 23) = 2.332, $p < 0.05$, as shown in Figure 6. Differences between easy, medium, and hard trials were more pronounced for groups than for an individual.

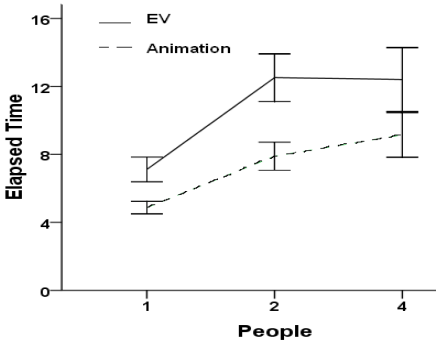


Figure 5. Elapsed time for each group and presentation condition. Error bars show +/- 2 SE.

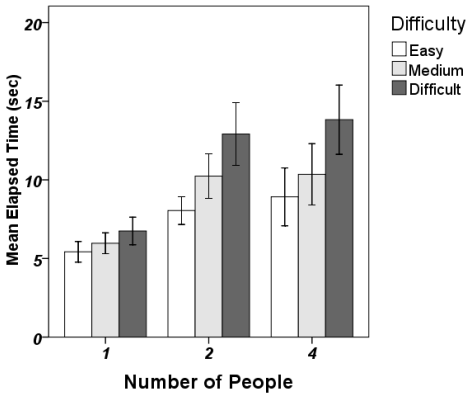


Figure 6. Elapsed time for each group and difficulty condition. Error bars show +/- 2 SE.

### 4.5 Summary and Analysis

Overall, animation had a higher success rate than EV. This contrasted with an earlier experiment that found no significant difference [17]. Possible explanations for the difference from

earlier work are the display method, i.e. larger horizontal display versus a computer monitor, group dynamics, or that the earlier study had insufficient power to show the difference. Consistent with the earlier work [17], animation was faster than EV.

Two people had a higher success rate than both an individual and a group of four. When 1 or 4 people completed the task, there was a clear drop in performance as difficulty increased that was not found in a group of two. More people also led to an increase in completion time. These results together suggest that when accuracy is important, two people may be optimal, especially for difficult spatial judgments.

## 5 Experiment 2 – Contrast Judgment

Comparing two models is a common task. For example, medical professionals compare images in cancer patients looking for differences in the size of tumors, perhaps before and after treatment. In this experiment we examine people's ability to identify differences between two similar 3D models using the two presentation conditions, animation and EV.

### 5.1 Task

Participants were presented with two different 3D block configurations, and asked to identify the number of differences between them. Each configuration was made up of 10 randomly assembled blocks. Blocks varied in both their size and their color (red, blue and yellow). Both configurations were presented simultaneously with the same display type: EV or animation. The left side acted as the source, while the right side acted as the target. There were two sources of differences between the two configurations: a target block could have a different position or a different color from the source. Figure 7 shows an example of both the animation condition (top) and the EV condition (bottom).

The target configuration either had 0, 1, 2, or 3 differences when compared to the source. Zero differences meant that the target and source configurations were actually the same. In the animation condition, both models rotated automatically at the same degree, at the same rate as in Experiment 1.

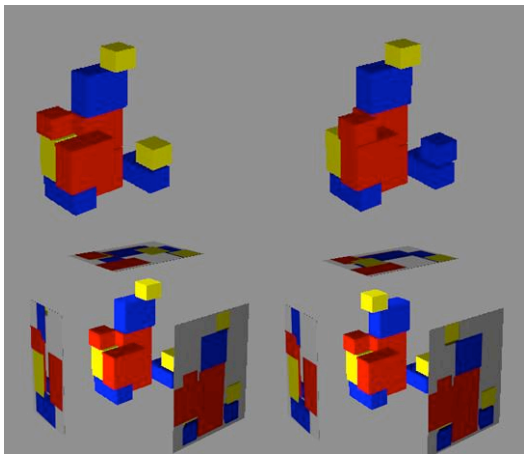

Figure 7. Contrast task. Animation is shown at the top, EV at the bottom. Top has two differences: red block at the left has moved forward and a yellow block at the right has turned blue. Bottom: the yellow block at the top has moved to the left.

### 5.2 Design and Procedure

Participants had 16 trials for each presentation condition: 4 trials where the target and the source were the same, 4 trials in which they differed in one way, 4 trials in which they differed in two ways and 4 trials in which they differed in three ways. Participants completed the task either alone (N=7), in pairs (N=9) or in a group of four (N=8). Order of presentation condition was counterbalanced; within these blocks, trial order was randomized.

Each participant used their own keypad and selected one button to indicate their answer to each trial. Participants who worked in pairs or in a group were required to come to a consensus before being allowed to continue. When participants had all selected the same answer, they were informed of the result via a green flash to indicate success, or a red flash to indicate failure.

### 5.3 Measures

We measured elapsed time and accuracy. Time was measured from the beginning of a trial to the time when all participant(s) had agreed on the same answer. Accuracy was marked on a success or failure criterion.

### 5.4 Results

For success rate, Pearson Chi-square tests were used to assess whether there were any differences between the presentation conditions and the number of people per group. For elapsed time, the data was log transformed to satisfy the normality assumption. A two-way ANOVA was used to assess whether there was an interaction effect between the presentation condition and the number of people per group. $P$ values less than 0.05 were considered significant. In preliminary analysis, the number of differences between target and source did significantly influence time or success rate; it was therefore not analyzed further.

#### 5.4.1 Success Rate

Figure 8 shows that the number of people significantly effected the success rate, $\chi^2$ (2, 23) = 17.815, $p < 0.001$. Bonferroni-corrected post hoc tests showed that a group of four people was significantly more accurate than either one or two people. An individual and a pair of people were not significantly different. Presentation condition had no observable effect: animation and EV were not significantly different. There was also no interaction between presentation condition and group size.

#### 5.4.2 Elapsed Time

As shown in Figure 9, increasing the number of people significantly increased the time to complete the task, F(2, 23) = 7.954, $p < 0.001$. Bonferroni-corrected post hoc comparisons showed that groups of four or two were significantly slower than an individual. In addition, animation was significantly faster than EV, F(1,23) = 52.21, $p < 0.001$. There was no evidence of an interaction between presentation and group size.

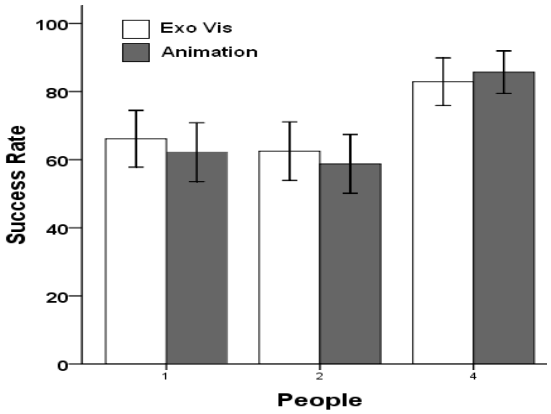


Figure 8. Success rate for the number of people and presentation condition. Error bars show +/- 2 SE.

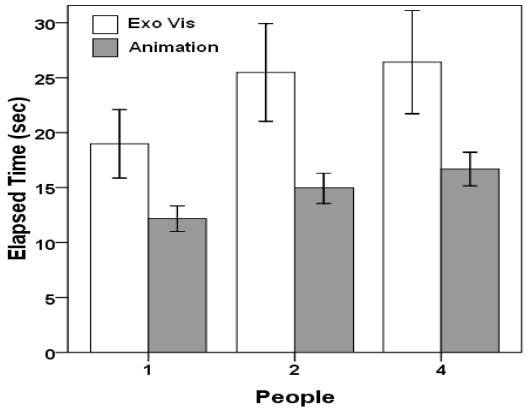


Figure 9. Elapsed time for each group and presentation condition. Error bars show +/- 2 SE.

### 5.5 Summary and Analysis

Consistent with earlier work [17], animation was faster than EV. The success rate showed no difference between EV and animation. Group size effected both accuracy and elapsed time. A group of four was found to be more accurate than both one or two people, but groups were slower than individuals. The results suggest that when accuracy is more important than completion time, a group of four performs the best at contrast judgment.

## 6 Experiment 3 - Construction

Experiments 1 and 2 evaluated the ability to understand a virtual representation. In contrast, for Experiment 3 we examined people's ability to translate the understanding of a 3D model to the real world. Figure 10 shows an example trial. Participants built a physical model that matched an on-screen visual model (i.e. the visual model served as the instructions). Unless otherwise specified, the experimental setup was the same as for the other two experiments.

### 6.1 Task

Participants were presented with a configuration of blocks on the display. The configuration acted as the instructions for assembling blocks in the real world. Participants were provided with exactly the ten necessary MegaBloks™ to recreate the visual configuration on the table.

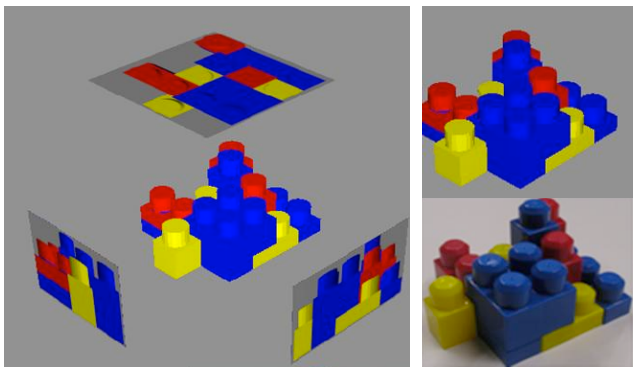

Figure 10. Example model and display for Experiment 3. Left: EV, top right: animation, bottom right: physical blocks.

### 6.2 Design and Procedures

The experiment was a 3x2 factorial design. The first factor was the number of people in the group. Participants completed the task either alone (N=8), in pairs (N=8) or in a group of four (N=8). The second factor was the presentation condition, which was either EV or animation. Each group assembled three configurations for EV and three for animation. Order of the presentation conditions was counterbalanced.

When participants were ready to begin the trial, they selected an on-screen start button. They were then presented with the block configuration on the screen. Participants then began assembling the blocks. When they had completed the assembly to their satisfaction, they signaled the experimenter and the trial was paused. If the configuration was correct, the trial was finished. If the configuration was not correct, the timer was continued and the participants were told to keep working on the configuration. When a trial was finished, the blocks were disassembled and the next trial started when signaled by the participants.

### 6.3 Measures

Participants were measured on a single variable, which was the time to complete the block assembly task.

### 6.4 Results

Elapsed time was log transformed to satisfy the normality assumption. A two-way ANOVA was used to assess whether there was an interaction effect between the presentation condition and the number of people per group. *P* values less than 0.05 were considered significant.

#### 6.4.1 Elapsed Time

Unlike the other two experiments, there was a significant interaction between presentation condition and group size, $F(5,21) = 6.137$, $p < 0.01$, as shown in Figure 11. EV showed a downward trend in elapsed time as the number of people increased. Animation, on the other hand, showed a dip in elapsed time with two people compared to one or four people. Overall, the number of people in a group significantly effected elapsed time, $F(2,21) = 4.673$, $p < 0.05$. A Bonferroni-corrected post hoc test showed that a group of two was significantly faster than one person. EV was not significantly different than animation when group size was ignored.

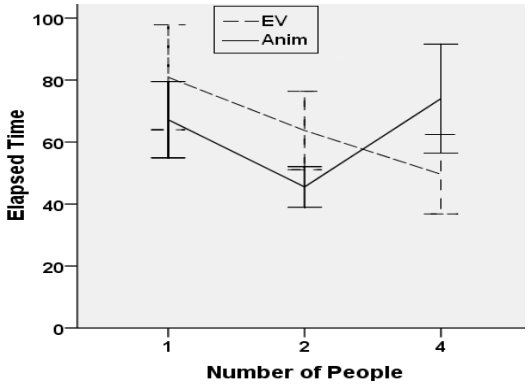


Figure 11. Elapsed Time for each group and presentation condition. Error bars show +/- 2 SE.

Details of the interaction effect were analyzed using Bonferroni-corrected post hoc tests. They showed that a group of two people using animation was significantly faster than both an individual and a group of four also using animation. For EV, a group of four was significantly faster than an individual. When

considering a group of two people, animation was significantly faster than EV consistent with the other experiments. However, when considering a group of four, the opposite was true: EV was significantly faster than animation.

## 7 Relationship to Steiner's Model of Group Productivity

Steiner's model [20] (presented in section 2.2) suggests that actual team productivity is the result of its potential productivity minus any faulty group interaction processes. Based on this definition, we can assume more participants will increase the likelihood of sub-optimal group interaction. When we analyze our results from this perspective, the distance judgment task is easily explained. This task was largely disjunctive: it required a group to choose one best answer from all the alternatives the group could generate. For this type of task, Steiner's model predicts an improvement in accuracy with increasing group size, but with diminishing returns. In cases where coordination cost is high, an inverted U shape is expected; accuracy should improve with larger groups up to some maximum, but then decrease with more people due to coordination. Note that "productivity" is typically measured simply as accuracy, as compared to accuracy per unit time.

An inverted U can be clearly seen in the success rate results for distance judgment. Note that the inverted U only occurred on difficult problems. For easy problems, participants likely all identified the correct answer so that adding more participants added little benefit, but also very little coordination was needed.

For the contrast task, groups also had an improved success rate compared to individuals, with groups of size four performing better. This task is only partially disjunctive: more eyes are likely to see the differences, which can then be easily verified by others. It remains to be seen exactly how additional group members would influence accuracy on this task; however, we might expect accuracy to quickly plateau and then fall off due to process losses.

Interpreting results of the block assembly task is more challenging. This task has the potential to be partially additive; additional hands could assemble the blocks more quickly, up to the point where there is insufficient room for more hands in the workspace. However, we observed that groups tended to complete the task in a highly coupled manner, with one person doing most of the assembly and others providing advice. Simultaneous building by more than one person was observed, but was neither common nor continuous. Thus the additive effect is likely rather small. Since participants were required to assemble each structure perfectly, there was no possible time-accuracy trade-off in this task and time is the only measure of effectiveness. Time for the animation condition followed a U shape (in this case not inverted because better performance has a lower value in comparison to accuracy where a higher value is better). This suggests an effect where there is an optimal group size of two and process losses for larger groups. Interestingly, EV did not follow this same trend, becoming in fact more effective than animation for groups of size four. A similar U shaped curve might be seen for EV but with a larger optimal group size. This of course could only be verified by repeating the experiment with larger groups.

## 8 Discussion

An interesting finding for the first two experiments is that the choice of visualization technique did not interact with the group size. Animation was clearly the better technique, being consistently faster (in both tasks) and more accurate (in the distance judgment task), consistent with earlier results [17]. This could either be because the animated view is easier to understand, or because the back of the object was not visible with EV. Regardless, group size appeared to effect these easy and hard-to-use interfaces equally, suggesting that the choice of presentation technique had minimal impact on group coordination. If true, this is helpful for design; a representation that works well for an individual also works well for a group. Thus our results generally support earlier findings that a rotatable 3D view is best for many 3D spatial tasks, even when the tasks are performed by groups.

However, the story is not quite as simple when we consider experiment 3. Here there was a clear interaction, and surprisingly, EV actually outperformed animation in some circumstances, specifically for groups of four. We pose two possible hypotheses to explain this result. The first is that having a static view made coordination easier: it could be potentially difficult to discuss a moving model. However, this would not explain why a similar effect was not seen for the other two tasks. The second hypothesis is that the multiple views provided by EV may have better mapped to individuals' different perspectives on the physical blocks. The physical blocks in this task forced individuals to take different perspective views (since they cannot all stand in the same direction relative to the structure), and there are more different viewpoints as the number of people increases. In this case, multiple static viewpoints may provide each individual with their own optimal view as well as provide an understanding of the viewpoint of others. We observed some evidence that perspectives on the physical blocks may have been important. In particular, participants tended to align the physical structure with the static 3D view in EV, and group together around the physical structure. Future experiments could further investigate these hypotheses by constraining participant movement to control or experimentally manipulate the points of view.

We also observed that groups generally achieved better results than individuals. In cases where accuracy at 3D spatial tasks is more important than time, our results imply that it is worth having two or more users work together. Spatial tasks depend on spatial ability and can be highly challenging for some individuals [10]. Presumably, working together allows users to identify and correct each other's interpretation mistakes. Including more users also increases the likelihood that one member of the group will have a high spatial ability and therefore perform better, influencing the entire group. At the same time, there is a maximum group size beyond which performance will begin to decrease due to coordination overhead. There are also obvious resource considerations when multiple users are assigned to a task that could potentially be performed by one user, although these are likely less important when decisions are expensive or safety critical (e.g., surgery planning). From a practical standpoint, a group of two may be a good choice in many cases, though we did sometimes observe higher accuracy with four users.

## 9 Limitations and Future Work

To provide sufficient control, our experiment only considered subtasks that are commonly carried out within a larger task set. However, viewers are more often concerned with a larger problem, and subtasks could be distributed among team members. Additional investigation is required to examine the effect of higher-level tasks, role divisions, and loosely coupled work styles on effectiveness of different visual representations. It would be particularly interesting to examine tasks that are very difficult to solve, to see what impact this would have on group dynamics and whether Steiner's law extends to these situations.

In considering each group as a unit whole, we did not explicitly look at the spatial ability or experience of individuals. We believe our overall results are valid because spatial ability differences should average out due to random assignment of people to conditions. However, further investigation of the relationship between individual spatial abilities and group performance is warranted and will likely lead to a richer understanding. One benefit of working in a group is that users ideally build upon each

others' skill sets. Thus, the overall accuracy of the group may reflect the performance of the member with the highest spatial ability. Simultaneously, the time to perform tasks in Experiments 1 and 2 could have been dominated by the person with the lowest spatial ability since all individuals were required to answer.

Finally, user expertise in a specific domain could impact the understandability of 3D representations. We did not control for this factor since our representations were very general. However, different presentation styles may be more or less effective depending on the experience and skills of the group members.

We plan to examine the interactive features necessary for individual and collective control of a 3D visualization. One extension could investigate private and shared views dedicated to various forms of investigations. Also, our studies suggest that if a representation works well for an individual it will probably work well for a group, but this should be further investigated with other visualizations and tasks. More generally, further investigation of collaborative visualization combined with sociological theories may be worthwhile, especially to examine factors such as group dynamics, strategy, and individual influence.

## 10 Conclusion

We explored effects of group size on comprehension of 3D spatial visualizations with two competing presentation styles: animated rotations and multiple static 3D and 2D views. Our results generally suggest that users are more effective with animated rotations than with multiple static views. They also suggest that as group size increases, so does accuracy, but with a tradeoff in task completion time. In more complex tasks, larger group sizes can facilitate more accurate judgments in 3D environments. However, this improvement in accuracy can peak with a relatively small group size; adding additional group members then negatively impacts performance, presumably due to the cost of coordinating individuals' actions and understanding. Results of our third experiment reveal that multiple perspectives may improve 3D spatial comprehension in group settings where each individual user maintains their own perspective of a physical 3D object.

## References


[1] C. Forlines, C. Shen, D. Wigdor, and R. Balakrishnan. "Exploring the effects of group size and display configuration on visual search," *Proc. ACM Conf. CSCW*, pp. 11-20, 2006.

[2] T. Grossman and D. Wigdor. "Going deeper: a taxonomy of 3D on the tabletop," *Proc. IEEE Tabletop*, pp. 137-144, 2007.

[3] M. Hancock and S. Carpendale. "Supporting multiple off-axis viewpoints at a tabletop display," *Proc. IEEE Tabletop*, pp 171–178, 2007.

[4] M. Hancock, S. Carpendale, and A. Cockburn. "Shallow-depth 3D interaction: design and evaluation of one-, two- and three-touch techniques," *Proc. ACM CHI*, pp. 1147–1156, 2007.

[5] J. Heer, F. van Ham, S. Carpendale, C. Weaver, and P. Isenberg. "Creation and collaboration: engaging new audiences for information visualization," In A. Kerren et al. (eds), *Information Visualization, LNCS 4950*, pp. 92-133, 2008.

[6] I. Heldal, M. Spante, and M. Connell. "Are two heads better than one? Object-focused work in physical and in virtual environments," *ACM Symp. VRST*, pp. 287-296, 2006.

[7] P. Isenberg and D. Fisher. "Collaborative brushing and linking for co-located visual analytics of document collections," *Computer Graphics Forum (Proc. EuroVis)*, vol. 28, no. 3, pp. 1031–1038, June 2009.

[8] P. Isenberg, A. Tang, and S. Carpendale. "An exploratory study of visual information analysis," *Proc. ACM CHI*, pp. 1217-1226 2008.

[9] M. Keehner, M. Hegarty, C. Cohen, P. Khooshabeh, and D.R. Montello. "Spatial reasoning with external visualizations: What matters is what you see, not whether you interact," *Cognitive Sciences*, vol. 32, pp. 1099–1132, 2008.

[10] M. Keehner, P. Khooshabeh, and M. Hegarty. "Individual differences among users: implications for the design of 3D medical visualizations," *User Centered Design for Medical Visualization*, F. Dong et al. (Eds), Information Science Reference, Hershey, PA, USA, 2008.

[11] G.E. Littlepage. "Effects of group size and task characteristics on group performance: a test of Steiner's model," *Personality and Social Psychology Bulletin*, vol. 17, no. 4, pp. 449-456, 1991.

[12] G. Mark and A. Kobsa. "The effects of collaboration and system transparency on CIVE usage: an empirical study and model," *Presence*, vol. 14, no. 1, pp. 60–80, 2005.

[13] G. Mark, A. Kobsa, and V. Gonzalez. "Do four eyes see better than two? Collaborative versus individual discovery in data visualization systems," *Proc. Intl Conf. InfoVis*, pp. 249-255, 2002.

[14] M.A Nacenta, S. Sakurai, T. Yamaguchi, Y. Miki, Y. Itoh, Y. Kitamura, S. Subramanian, and C. Gutwin. "E-conic: a perspective-aware interface for multi-display environments," *Proc. ACM UIST*, pp. 279–288, 2007.

[15] G.G. Robertson, S.K. Card, and J.D. Mackinlay. "Information visualization using 3D interactive animation," *Communications of the ACM*, vol. 36, no. 4, pp. 56–71, 1993.

[16] K. Ryall, C. Forlines, C. Shen, and M. Ringel Morris. "Exploring the effects of group size and table size on interactions with tabletop shared-display groupware," *Proc. ACM CSCW*, pp. 284-293, 2004.

[17] T. Sando, M. Tory, and P. Irani. "Effects of animation, user-controlled interactions, and multiple static views in understanding 3D structures," *Proc. Applied Perception in Graphics and Visualization*, 2009.

[18] G. Simmel. "The number of members as determining the sociological form of the group," *Am. J. Sociology*, vol. 8, pp. 1-46 and 158-196, 1902.

[19] R.L. Sollenberger and P. Milgram. "The effects of stereoscopic and rotational displays in a three-dimensional path-tracing task," *Human Factors*, vol. 35, no. 3, pp. 483–500, 1993.

[20] I.D. Steiner. *Group Process and Productivity*. Academic Press, 1972.

[21] M. Tobiasz, P. Isenberg, and S. Carpendale. "Lark: coordinating co-located collaboration with information visualization," *IEEE Trans. Visualization and Computer Graphics (Proc. InfoVis)*, vol. 15, no. 6, pp. 1065–1072, 2009.

[22] M. Tory, "Mental Registration of 2D and 3D Visualizations (An Empirical Study)", *Proc. IEEE Visualization*, pp. 371-378, 2003.

[23] M. Tory and C. Swindells. "Comparing exovis, orientation icon, and in-place 3D visualization techniques," *Proc. Graphics Interface*, pp. 57–64, 2003.

[24] M. Tory, A.E. Kirkpatrick, M.S. Atkins, and T. Möller. "Visualization task performance with 2D, 3D, and combination displays," *IEEE Trans. Visualization and Computer Graphics*, vol. 12, no. 1, pp. 2–13, 2006.

[25] M. Velez, D. Silver, and M. Tremaine. "Understanding visualization through spatial ability differences," *Proc. IEEE Visualization*, pp. 511-518, 2005.

[26] C. Ware and G. Franck. "Evaluating stereo and motion cues for visualizing information nets in three dimensions," *ACM Trans. Graphics*, vol. 15, no. 2, pp. 121–140, 1996.

[27] C. Ware, K. Arthur, and K.S. Booth. "Fish tank virtual reality," *Proc. ACM INTERACT'93 and CHI'93*, pp. 37–42, 1993.

[28] C. Ware. "3D contour perception for flow visualization," *Proc. Applied Perception in Graphics and Visualization*, pp. 101–106, 2006.

[29] D. Wigdor., C. Shen, C. Forlines, and R. Balakrishnan. "Perception of elementary graphical elements in tabletop and multi-surface environments," *Proc. ACM CHI,* 473–482, 2007.