International Journal of Computer Information Systems and Industrial Management Applications.
ISSN 2150-7988 Volume 3 (2011) pp. 578-585
© MIR Labs, www.mirlabs.net/ijcisim/index.html

Effectiveness of Paper, VR and Stereo-VR in the Delivery of Instructions for Assembly Tasks

John Strobel[1], **Guy W. Zimmerman**[2]

Department of Computer Science, Bowling Green State University,
235 Hayes Hall, Bowling Green, Ohio, 43402 USA
[1]*john.strobel@ymail.com*, [2]*gzimmer@bgsu.edu*

***Abstract*: A study was conducted to compare the relative performance, in terms of completion time and accuracy, of desktop virtual reality, desktop stereo virtual reality, and a paper-based approach in presenting assembly instructions for 3D construction tasks. All presentations depicted a single step of the assembly process at a time. The results showed no significant differences in accuracy between any of the three approaches. In terms of interaction, a comparison of the two computer based presentations revealed that users rotated the model less frequently in the 3D stereo graphics than in the standard graphics, suggesting that more information was communicated by the stereo approach. The results also reveal that object complexity significantly impacts user performance with regards to time and that presentations of real-world assembly tasks will benefit from enhanced attention to inter-spatial relationships.**

Keywords: Stereographic, Assembly Instructions, Usability, Spatial Ability.

I. Introduction

Technologies for displaying 3D graphics/virtual worlds stereoscopically have been available for many years; they run the gamut from very expensive, high-end systems such as the CAVE to immersive head-mounted displays to low-cost glasses such as CrystalEyes shutter glasses. The applications of these technologies have largely been the province of experts in specialized niche areas (e.g. medical imaging); with non-experts encountering them only infrequently in the guise of 3D movies as they come in and go out of fashion. It is only relatively recently that the necessary elements have come together for stereoscopic display technology to potentially become commonplace (e.g. in the workplace, in the home). While it seems clear that the entertainment industry (film, television and gaming) will continue to be responsible for any widespread use, it is reasonable to consider practical applications of this technology for non-experts, especially given its decreasing cost and concomitant increasing ubiquity. In this paper, we consider the application of this technology in aiding the understanding and execution of instructions for assembly tasks.

Assembly tasks are commonplace; examples include assembling a child's bicycle or a piece of furniture. Such tasks can consist of anywhere from 10s to 1000s of steps, with the instructions typically presented as a sequence of illustrations (on paper; see Figure 1) optionally accompanied by text. Such paper presentations are notorious for being difficult to use, for a variety of reasons. However, the availability of 3D-graphics capable display technology holds out the promise of potentially mitigating many of these difficulties. In this paper we consider the utility of interactive 3D graphics as a component in a system to deliver instructions for an assembly task. We present results from a study that compared the relative effectiveness of three presentation formats: paper, 3D and 3D stereo in completing assembly tasks. Subjects' spatial ability was measured and included in the analysis.

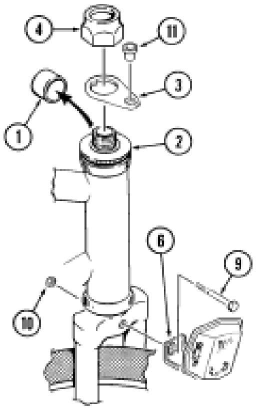

Figure 1. A diagram from a set of bicycle assembly instructions

The remainder of this paper is organized as follows: Section II will present background material and relevant results from the literature. Section III describes our study setup and procedure. Section IV reports our results, discussed in section V and section VI our conclusions.

II. Background and Related Work

As mentioned, the use of diagrams to assist people in the execution of assembly tasks is commonplace, and almost synonymous with assembly tasks themselves. Examples include assembling prefabricated furniture, children's toys, and origami figures. Despite the pervasive use of diagrams in the presentation of such instructions for assembly tasks, only recently has any research been done on the role such diagrams play in supporting the assembly task [18]. Since most such presentations are provided in printed form, much of the existing work has focused on two-dimensional diagrams. These are typically perspective drawings of the object to be assembled using a predetermined viewpoint – presumably selected by the instruction designer as being optimal for the step(s) of the assembly being illustrated. Novick [17] examined diagrams that accompany instructions for folding origami figures. Agrawala [1] presented a suite of design principles to create two-dimensional diagrams as well as a system for the automated production of those diagrams.

Heiser et al [9] empirically studied the qualities of successful assembly instructions through a set of five experiments, one of which was a continuation of earlier work done by Agrawala, et al [1]. The design principles that their subjects preferred, and which helped reduce total assembly time and perception of task difficulty, were programmed into an automated assembly instruction design system. They showed that the resulting diagrams were significantly better in terms of assembly times and error rates than the highest-rated hand-drawn assembly instructions as well as the factory-provided assembly instructions. One of the design principles they recommend is the use of action diagrams (actively showing the user what to do; see Figure 2) vs. structural diagrams (which depict a finished step, and generally requires a visual comparison to a previous step to determine the correct action; see Figure 3). The authors note that people generally prefer that instructions partition the steps of an assembly over multiple diagrams. However, it is most common for a single diagram to illustrate some minimal number of steps (>1), partly to reduce the total number of diagrams – reducing the production and printing cost.

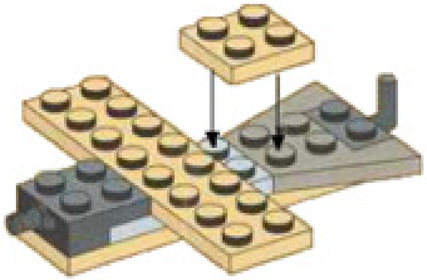

Figure 2. An action diagram with arrows showing an assembly step.

A. Spatial Ability

Using (2D) paper representations of 3D objects requires users to perform a mental translation from two dimensions into three; a person's aptitude at performing this and other mental manipulations is referred to as their spatial ability (SA). The definition of spatial ability also extends to extracting spatial information and SA is commonly measured by some combination of standard psychometric tests [15]. It has been shown [8] that those with high spatial ability are more successful in mathematics and in construction tasks such as those performed in the Heiser, et. al. [9] study. As mentioned above, paper instructions require an inherent mental translation from a 2D diagram to the 3D object being assembled. A person with high spatial ability might only be mildly impacted (or not impacted at all) by this translation, whereas a person with low spatial ability might find this translation very difficult or even impossible without extra information.

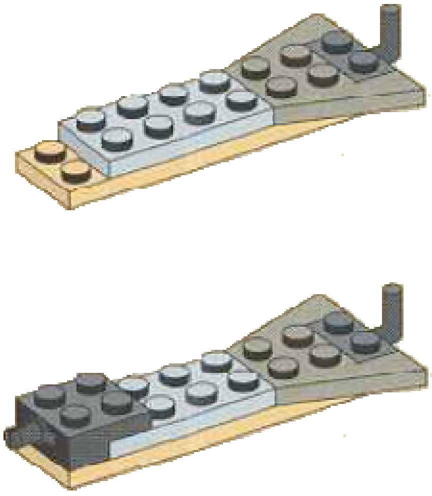

Figure 3. A structural diagram showing before state and after state.

A few studies of virtual reality hint that some individuals may be "better" at some aspects of using a presentation than others [3], [5], [12], [19], [21]. Velez et al. [21] found a positive correlation between spatial ability and performance on their computerized visualization test. Czerwinski et al [5] and Tan et al [19] did consider spatial ability but found no effect. Ritter et al. [17] observed that subjects who scored well on a standardized figurative classification had higher learning performance than the subjects as a whole.

B. Depth Cues and Stereo Graphics

Part of the problem with 2D diagrams is in the limited amount of depth-cue information that they can convey; a number of researchers have examined the extent to which various technologies can assist in providing depth-cue information. Here we focus on stereographic display technology specifically.

Lo and Chalmers [16] conducted research into what effect viewing a scene in stereo had on how real the scene seemed. In their study, subjects viewed several scenes with disparate depth cues presented. Subjects were divided into two groups randomly, one of which was presented with a stereographic presentation, the other with a standard 2D presentation (both groups wore the shutter glasses to control against those effects). They were asked to report if a scene was "real" or "not real", and the time it took to respond was measured. Through the three experiments conducted, each involving different monocular depth cues such as illumination direction, light source numbers, and viewpoint variances, those presented with the stereographic presentation took longer to respond. Lo and Chalmers claim that this is an indication that stereographics are interpreted by the user as being more realistic than standard 3D computer graphics. Stereographics were used in a study by Grossman and Balakrishnan [7] to compare the ability of subjects to perceive depth information. Their approach pitted a perspective projection onto a 2D display and a stereoscopic display against a volumetric display system. The study consisted of three tasks which

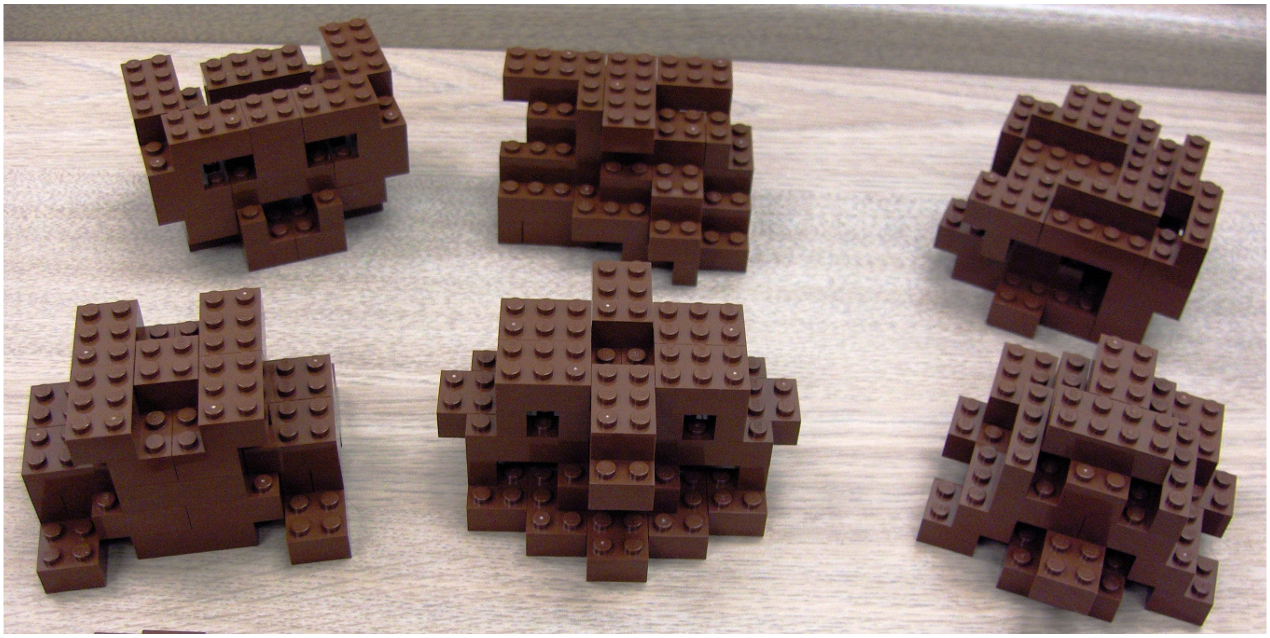

Figure 4. The six Lego models used in the study.

included depth ranking of a wire-frame sphere, path tracing, and potential collision judgment. In all three tasks, the users of the stereographic presentations were significantly more accurate than the users of the perspective projection treatment.

Additionally, Hu et al. [10] performed two studies comparing the effects of three depth cues on subjects' ability to maneuver a block to where it was as close as possible to, but not touching, a virtual tabletop. Their results indicated a clear advantage for stereographics in accuracy of depth perception. Similarly, Hubona et al. [11] compared various depth cues while having users perform two tasks: positioning and resizing of objects. The independent variables included three depth cues (shadows on/off, number of light sources, stereo/mono viewing) and three different background settings (plane, stairs, room). Of particular concern to the current study was that their results indicated that in the positioning task, stereographics resulted in faster completion and lower errors on the part of the subject.

C. Virtual Reality and Assembly Tasks

The limitations of paper diagrams outlined above suggest several approaches to potentially improve assembly instruction effectiveness, including: 1) make diagrams interactive, and 2) provide additional 3D depth cue information. Both of these can be accomplished through the use of virtual reality technologies. A number of researchers have examined the application of virtual reality systems to the problem of object assembly.

Li W. et al [14] described a system for the creation and viewing of exploded views of complex 3D models. The system allowed both direct and indirect interaction modes. Kashiwazaki [13] discusses potential advantages of 3D 'contents' as compared to 2D 'contents' in the teaching of assembly/disassembly procedures. Likewise, the RapidManuals™ application produced by Cortona3D [6]. However, in these cases only anecdotal evidence is provided as to their effectiveness. Boud [4] reports on a study where subjects were trained to assemble a pump consisting of 8 parts. The study compared several virtual reality technologies (desktop VR, desktop Stereo-VR, augmented reality) with a standard engineering drawing. Subject assembly times for all VR presentations were faster than with the standard drawing. A study by Antifakos, Michaheles, and Schiele [2] presented a proof-of-concept, using sensors to detect assembly stages and guide the user accordingly. Yuan, Ong, and Nee [23] compared screen vs. head-mounted displays for an augmented reality approach. This study focused on helping the designers of assembly instructions as opposed to determining design principles or the effect of augmented reality on instruction comprehension or performance. Zauner et al [24] describe a tool for guiding users in assembly instructions, focusing on presenting the tool rather than studying the effects of it. A study by Zimmerman, Barnes and Leventhal [26] compared desktop vs. handheld computers in the use of web-delivered VRML instructions. The task was to create an origami figure based on the interactive 3D instructions. It was determined that the handheld computer presentation was just as effective as the desktop presentation. A subsequent study [25] examined instructions for assembling Lego models on mobile devices. This study also considered subjects' spatial ability, finding differences in the ways users with different levels of spatial ability performed and used the interactive features of the presentation. Further, Zimmerman considered the impact of the objects' inherent complexity, including spacing between substructures and their general symmetry. The study found a significant negative impact on performance for asymmetric models with increased space between substructures. Finally, Tang et al [20] studied the effect of augmented reality systems by comparing computer screen-based instructions to head-mounted-display-based (HMD-based) instructions. Their results show that the HMD-based instructions drastically decreased the error rate in a Duplo block (similar to Lego block) construction task.

III. Research Study

Our study was motivated by the following question: how effective is the use of stereographic displays in the

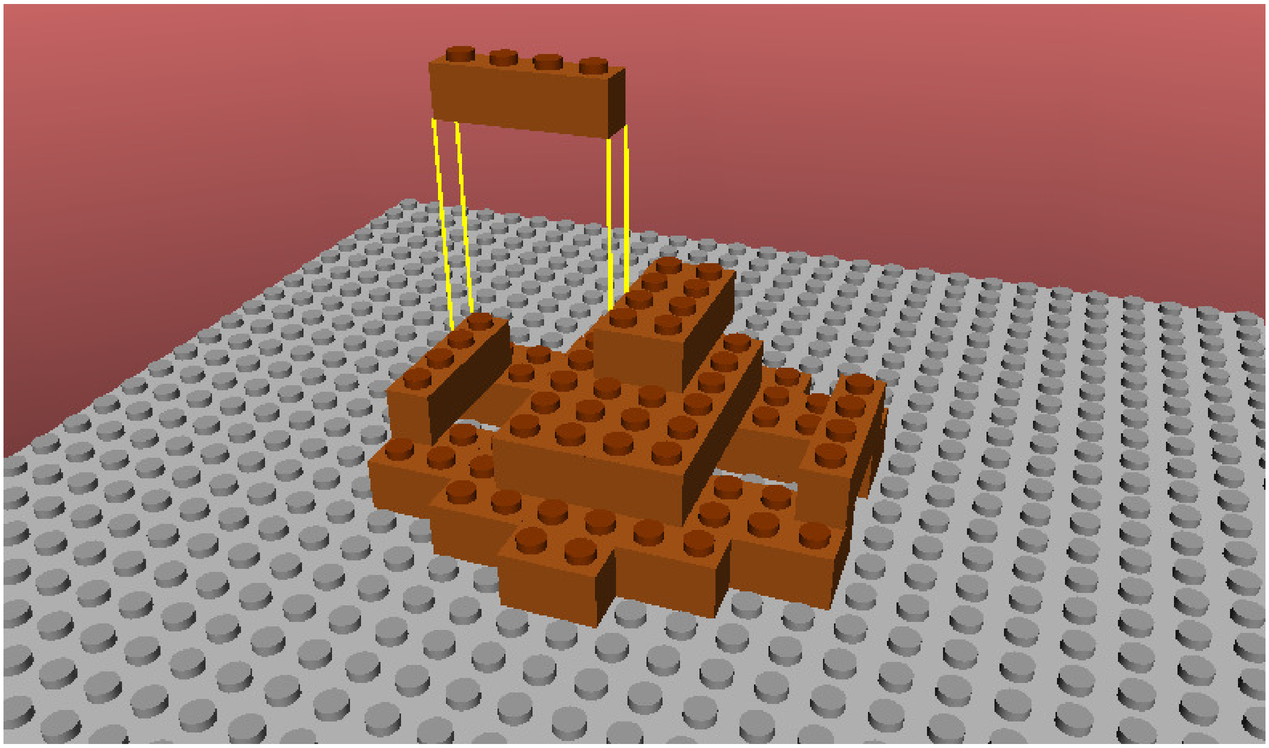

Figure 5. A step from the printed instruction presentation.

presentation of instructions for an assembly task, compared to non-stereo and paper formats? Completion time and error rates were used as measures of effectiveness; a number of secondary independent variables were considered in addition to presentation type.

We note that this study was part of a Master's thesis at Bowling Green State University which had a broader scope than the focus of this paper.

A. Materials and Task

To address the above question we selected the construction of six abstract Lego models for our assembly task. Abstract designs were used so that subjects would not be affected by preconceived notions as to what they were to build. All models consisted of the same 40 pieces in differing arrangements, each comprising 5 layers. Six models were employed in order to allow for varying degrees of model complexity: three types of symmetry (0, 1, and 2 lines of rectilinear symmetry) and two spacing conditions (packed and not-packed) were considered. See Figure 4.

Lego blocks were chosen as the building material for several reasons: they are unambiguously 3D (as compared to tasks such as origami), they are inexpensive, they are familiar (requires minimal training) and they allow for models of incremental increasing complexity to be generated. Additionally, the software for creating Lego-based model presentations was readily available, having been developed for use in prior studies.

B. Presentation Types

There were three different presentation types: paper, desktop VR, and desktop stereo-VR. All presentations were created using the Virtual Reality Modeling Language (VRML). The instructions depicted the construction process by layer, from bottom to top, with all pieces within a layer positioned before those in higher layers. Pieces were placed from the back to the front and from left to right based on the original viewpoint. All pieces were the same color: medium brown.

The paper presentation was comprised of screenshots of the VRML models printed on standard letter-sized paper. Each model required 41 images, one per step for all 40 steps and one to show the final product.

Every diagram except the final image (completed product) showed which piece was to be added as well as lines at the four corners indicating the correct placement of the piece (see Figure 5). The added lines effectively made them action diagrams and this was intended to be analogous to the information conveyed by the animations in the other two presentations. The paper presentation was secured to the top of a monitor so that the viewing angle for each presentation would be the same for each subject relative to the computer screen. The subject was asked to flip a page up over the top of the screen to move on to the next diagram. They were not restricted in their interaction with the instructions, and could flip back and forth through the diagrams at will.

Both desktop presentations (non-stereo, stereo) were rendered using the Cortona VRML browser version 6.0 within Internet Explorer 7. The Cortona browser was set to full-window 800x600, 120 Hz. Keyboard input was chosen so that interaction data could be collected reliably. Subjects were able to rotate the model on its vertical axis, zoom in and out on the center of the object, play an animation of the current piece falling into its correct position, and load the previous or next instruction step. This ability to reorient the viewpoint is illustrated in Figure 6. The figure shows the same model/step as that shown in Figure 5; however the viewpoint has been changed so that the piece to be added is now seen edge-on and is the front-most piece. As mentioned in the introduction, a subject’s ability to see the figure from different viewpoints is one of the (potential) advantages of using a full 3-D representation.

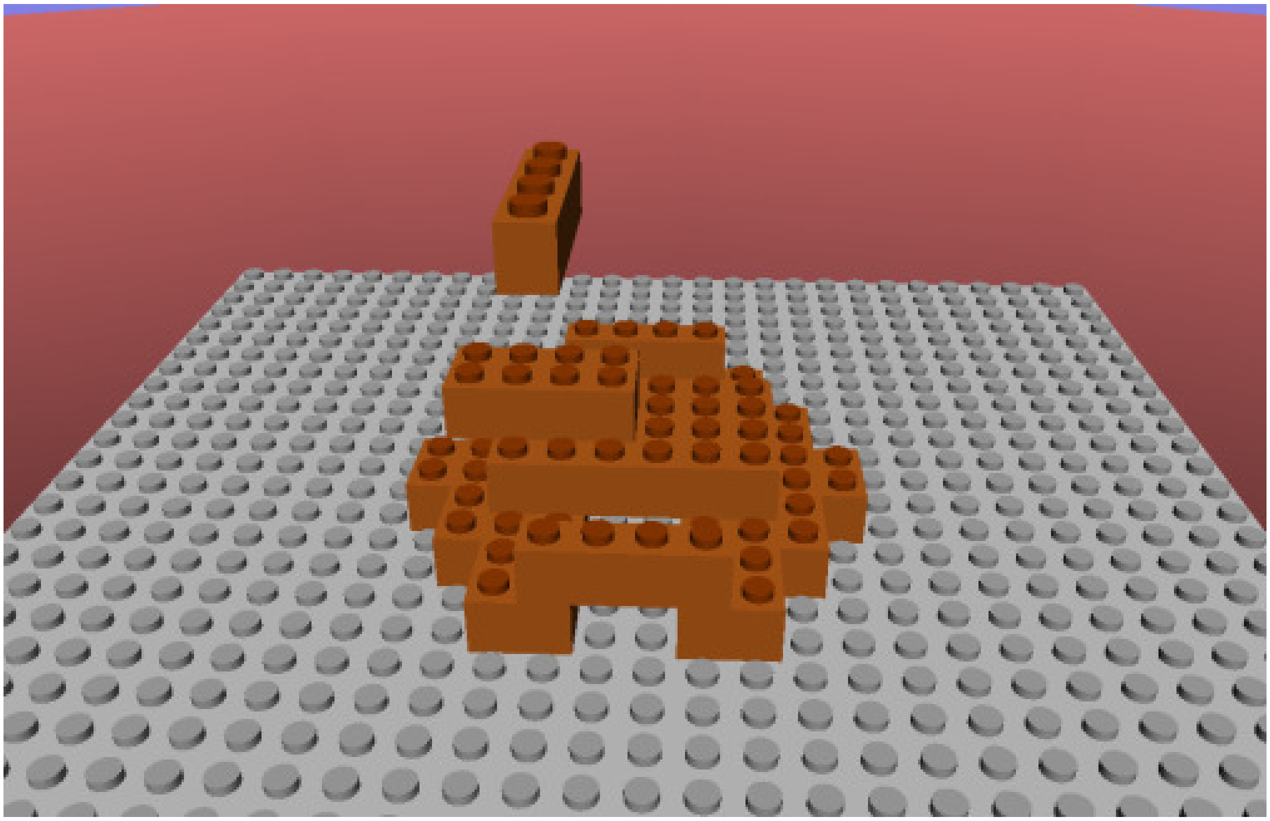

Figure 6. The computer-based presentation showing a user selected viewpoint.

The stereo presentation only differed from the non-stereo presentation in that it made use of stereographic technology. Shutter glasses (to enable page-flipped stereo images) were worn during the experiment for these subjects only.

C. Data Collection

For all presentations, subjects' hands were videotaped as they were constructing each model to provide timing data and to facilitate grading (described later). For the paper presentation only, a second camera was used to record interaction data. Interaction data for the two computer-based approaches was gathered through the VRML renderer's console (Cortona 6.0). Both the action taken and a time stamp were recorded in this way. All subjects completed a spatial ability test suite (card rotation and paper folding) before beginning the experiment. The spatial ability tests were web-based to allow immediate scoring.

D. Subjects

This study included 58 subjects, of whom 54 produced viable data for analysis. Of the four sets of data removed, three were due to mechanical failures of the videotape and one for producing models so incorrect as to make accurate grading impossible. Subjects were volunteers from two mid- to upper-level Computer Science courses offered at Bowling Green State University. They were compensated in the form of extra credit from their professors.

E. Procedure

Upon arrival at the lab, subjects were first asked to sign a consent-to-participate form, and were then asked to complete the online SA tests previously described. Based on their scores, the experimenter assigned them to one of the three presentations, according to the (then) current number of high and low SA subjects in each presentation. This was to attempt to balance each presentation in terms of high and low SA participants.

The subjects completed a demographic survey and then performed a simple training exercise, assembling a four-piece model to familiarize themselves with their respective presentation types. All subjects constructed the same training model in their presentation type. They were allowed to take as long as they needed for this task, and only continued to the actual experiment after they indicated that they were ready. The order of the six models was randomized to control for order bias and each subject completed all models. Subjects were asked to indicate that they had finished with each model before they were given the subsequent model. Subjects were provided an empty building board and a bin of pieces for each model. Subjects were allowed to leave at any time, however none chose to leave before completing all six models. After they left, the models were disassembled layer-by-layer; with photographs being taken of each layer for use in subsequent grading.

F. Grading Procedures

Subjects' SA test scores were computed as the difference between the number of correct and incorrect responses, with no penalty for not answering. The sum of the two test scores was used as the subject's spatial ability score. Timing data was collected from the video of the subjects' hands as they worked. Time was measured from the placement of the first piece in the model to the placement of the last piece in the model. Grading of model accuracy was completed from pictures of the final constructs after the subjects had left. The pictures were of each layer, and layers were graded from bottom to top. Grading was per-piece and consisted of counting the number of errors made. Two types of errors were possible: absolute and relative. Absolute errors were determined as incorrect placement relative to the first piece in each model. Relative errors were determined as incorrect placement relative to the first piece in each layer. It is possible for a single piece to be both relatively and absolutely incorrect.

IV. Results

We considered the impact of the independent variables: presentation type (paper, desktop VR and desktop stereo-VR) and spatial ability on two dependent variables, time (in seconds) and Accuracy (measured in number of errors), for all 6 models. The analyses for time and accuracy were performed with separate repeated measures analysis of variance (ANOVA) tests. For the dependent variable Accuracy, the mean number of absolute errors made per subject over all 6 models was 0.256, with a standard deviation of 1.341. The mean number of relative errors was 0.229, with a standard deviation of 1.140. These numbers were not large enough to produce any significant effect, and further analysis was deemed unwarranted.

The means and standard deviations for time were: paper (mean=2032.65, sd=378.392), desktop VR (mean=2141.33, sd=739.979), desktop Stereo-VR (mean=2201.61, sd=690.976). The difference of the means was not significant: F(2,50)=0.324, $p=0.725$.

Because the distribution of spatial ability scores of the population did not exhibit a clear bimodal distribution, the middle section (12 subjects) was removed. The remaining subjects were then characterized as having low/high spatial ability, respectively.

In terms of time, the high spatial subjects were significantly faster than those with low spatial ability. F(1.000, 39.000)= 7.858 $p<.008$ (High SA: mean= 317.579 sd= 22.107; Low SA: mean= 406.308, sd= 22.653). Also in terms of time, subjects were significantly faster on the packed condition than on the not-packed condition. F(1.0, 24.0) = 17.08, $p<0.001$ (Packed: mean = 312.96, sd = 90.87; Not-Packed: mean = 358.94, sd = 141.55). There was a significant three-way interaction between spacing, condition, and subjects' spatial ability (F(2.0, 24.0) = 5.38, $p<0.05$), summarized in Table 1.

Restricting our attention further to the two desktop presentations, 15 subjects performed at least one rotation of each model; eight were in the stereo condition, 7 in the non-stereo condition. We found a significant effect (F(1,11)=4.882, $p< 0.005$) of the independent variable spatial ability (SA) on the dependent variable time per Rotation (measured in seconds per rotation). Subjects in the stereo condition (mean 21.718, sd=15.290) rotated less often than the non-stereo condition (mean=13.023, sd-8.423). There was also a significant effect of spacing on time per Rotation (F(1.0, 11.0) = 6.408, $p<0.05$)). Subjects rotated less often for the packed condition (mean = 19.684, sd = 10.661) than the not-packed condition (mean = 15.057, sd = 7.465).

Condition * Spatial Ability (SA) * Spacing				
Condition	SA	Spacing	Mean	Std. Dev
Desktop VR	Low	Packed	300.00	44.533
		Spaced	324.00	36.320
	High	Packed	262.8	90.241
		Spaced	278.00	44.233
Paper	Low	Packed	355.4	120.728
		Spaced	378.933	76.667
	High	Packed	295.333	55.950
		Spaced	353.333	54.810
Desktop Stero-VR	Low	Packed	384.476	79.923
		Spaced	524.905	47.021
	High	Packed	279.733	202.370
		Spaced	293.867	79.375

Table 1. Means and standard deviations of all combinations of the condition, spatial ability, and spacing variables.

V. Discussion

The results reported in section IV reveal that no significant difference was observed in terms of accuracy. Likewise no significant difference was observed in terms of time. On the surface, this would seem contrary to the results in Boud [4] where completion times were faster (significantly) for all VR technologies compared to a standard engineering drawing. However, in the Boud study, the drawing illustrated all 8 steps of the assembly within the single drawing, while the VR presentations depicted the assembly steps sequentially. Also, in the Boud study, the presentations were designed for training subjects on the task – to be performed subsequently, rather than as an active aid in performing the task. The significant disparity of completion times between low and high spatial ability subjects is not in retrospect surprising; this simply supports the theory that people with high spatial ability tend to perform 3D construction tasks more easily than people with low spatial ability.

The impact observed in this study of an object's internal spatial relationships is consistent with the findings of the Zimmerman [25] study. In the packed condition, subjects were consistently able to complete construction in less time than in the non-packed condition. One explanation of this observation is that the pieces themselves provided strong cues to the location of future pieces. This is corroborated by the fact that Lego blocks are topped by pegs which are spaced regularly and can thus be counted to give an exact spatial relationship. An increase in distance also increases the possibility of a counting error, explaining the increased time for non-packed configurations. Augmenting presentations to include additional spacing-cues for such non-packed substructures could serve to ameliorate the increased times needed for subjects to complete the correct placement.

With respect to construction time, an interesting result was discovered in the interaction between spacing, presentation condition, and spatial ability. For both high and low spatial ability subjects, the non-stereo presentation was fastest (packed, then not-packed). However, Table 1 shows that the second-fastest presentation for the high spatial subjects was the stereo condition and for the lows spatial subjects, the paper condition (both packed, then not-packed, as before). This could indicate that the stereo condition gives more information than the paper presentation, which the subjects with high spatial ability are more able to tolerate. The extra information provided by the stereoscopic vision could potentially be a source of cognitive overload for subjects with low spatial ability, the effect of which can be seen in their completion times.

In terms of time per rotation, the model was rotated more frequently in the non-stereo condition than in the stereo condition. These findings are consistent with those of Lo and Chalmers [16] in that the stereo presentation seems more realistic and could, in fact, be delivering more information to the user. Again, the user might become overloaded by the amount of incoming information, explaining the longer completion times and the slower rotation rate; to get the same amount of information requires less frequent rotations. This is supported further by observing the increased time per rotation on models in a packed configuration versus those in a not-packed condition.

The low number of errors made rendered meaningful error analysis futile. One possible explanation for the low number of errors is that the models were not complex enough. This could be tested by some combination of a larger number of pieces, a larger variety of pieces (building blocks) and more complex spatial relationships among the pieces. Another explanation is that by limiting the presentation to one piece per step, the problem becomes simply too easy – independent of any delivery mode. If so, this would be significant: once one decides to use electronic delivery, there no overriding reason to overload the user with more than piece of information at a time and thus the added expense of 'exotic' technologies would not be warranted.

VI. CONCLUSION

The study described evaluated three methods for delivering assembly instructions (paper, non-stereo, and stereo). There were no significant differences in terms of time or accuracy. From a cost/benefit perspective, this suggests that the lowest cost solution should be preferred. In terms of the effects of object complexity, it is clear that the spatial relationships of the components of the object is an important factor in user performance and designers can improve the effectiveness of instruction presentations by making these spatial relationships as clear as possible; especially where components are not proximal. The inherent symmetry of an object being constructed was also shown to play a role in object complexity, though the effects were less pronounced.

There are several directions for future work suggested by this study. One would be to increase object complexity: total number of components, variety of components, as well as the spatial relationships of those components. In particular, it would useful to include non-rectilinear relationships among the component pieces, as these are common in actual applications. A second direction would be to vary the delivery device to include mobile devices (small screen) and 3D televisions. Finally, further research needs to be done on the effect of stereo and similar technology on 3D object construction.

References

[1] M. Agrawala, D. Phan, J. Heiser, J. Haymaker, J. Klingner, P. Hanrahan, and B. Tversky, "Designing effective step-by-step assembly instructions", *ACM Transactions on Graphics (TOG)*, 2003, 22(3):828-837.

[2] S. Antifakos, F. Michahelles, and B. Schiele, "Proactive instructions for furniture assembly", *Lecture notes in computer science*, 2002, pp. 351-360.

[3] R.F. Astur, M.I. Ortiz, and R.J. Sutherland, "A characterization of performance by men and women in a virtual Morris water task: a large and reliable sex difference", *Behavioral Brain Research* 93, 1998, pp. 185-190.

[4] A.C. Boud, "Virtual Reality and Augmented Reality as a Training Tool for Assembly Tasks", *Proceedings of the 1999 international Conference on information Visualization*, Washington, DC, 1998, 32.

[5] M. Czerwinski, D.S. Tan, and G.G. Robertson, "Women take a wider view". *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Changing Our World, Changing Ourselves*, 2002, pp. 195-202.

[6] Cortona3D, available at http://www.cortona3d.com/interactive_electronic_technical_manuals.aspx [accessed 20 May 2010]

[7] T. Grossman, and R. Balakrishnan, "An evaluation of depth perception on volumetric displays". *Proceedings of the Working Conference on Advanced Visual Interfaces*. Venezia, Italy, 2006, pp.193-200.

[8] M. Hegarty, and M. Kozhevnikov, "Types of visual-spatial representations and mathematical problem solving". *Journal of Educational Psychology*, Vol. 91, No. 4, 1999, pp.684–689.

[9] J. Heiser, et al, "Identification and validation of cognitive design principles for automated generation of assembly instructions". *Proceedings of the Working Conference on Advanced Visual interfaces*. Gallipoli, Italy, 2004, pp. 311-319.

[10] H.H. Hu, et al, "Visual cues for perceiving distances from objects to surfaces". *Teleoper Virtual Environ*. Vol. 11, No. 6, 2002, pp. 652-664.

[11] G.S. Hubona, et al, "The relative contributions of stereo, lighting, and background scenes in promoting 3D depth visualization". *ACM Trans. Comput.-Hum. Interact*. Vol. 6, No. 3, 1999, pp. 214-242.

[12] G.S. Hubona, G.W. Shirah, D.G. Fout, "3D object recognition with motion". *CHI '97 Extended Abstracts on Human Factors in Computing Systems: Looking To the Future*, 1997, pp. 345-346.

[13] K. Kashiwazaki, et al, "Interactive 3D learning contents based on 3D meme media architecture". *Proceedings of the 4th international Symposium on information and Communication Technologies*, Cape Town, South Africa, 2005, pp. 14-19.

[14] W. Li, et al, "Automated generation of interactive 3D exploded view diagrams". *ACM SIGGRAPH 2008 Papers*, Los Angeles, California, 2008, pp. 1-7.

[15] D.F. Lohman, "Spatial abilities as traits, processes, and knowledge". *Advances in the psychology of human intelligence*, Vol. 4. Erlbaum, Hillsdale, NJ, 1988, pp.181-248.

[16] C.H. Lo, and A. Chalmers, "Stereo vision for computer graphics: the effect that stereo vision has on human judgments of visual realism". *Proceedings of the 19th Spring Conference on Computer Graphics*. Budmerice, Slovakia, 2003, pp. 109-117.

[17] F. Ritter, et al, "Virtual 3d jigsaw puzzles: studying the effect of exploring spatial relations with implicit guidance", *Proceedings of Mensch & Computer*, 2002, 363-372.

[18] L.R. Novick, and D.L. Morse, "Folding a fish, making a mushroom: The role of diagrams in executing assembly procedures". *Memory and Cognition*, Vol. 28, No. 7, 2000, pp. 1242-56.

[19] D.S. Tan, G.G. Robertson, M. Czerwinski, "Exploring 3D navigation: combining speed-coupled flying with orbiting", *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems CHI '01*, 2001, pp. 418-425.

[20] A. Tang, et al, "Comparative effectiveness of augmented reality in object assembly". *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Ft. Lauderdale, Florida, USA, 2003, pp. 73-80.

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

[22] C. Ware, "Dynamic stereo displays". *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Denver, Colorado, USA, 1995, pp. 310-316.

[23] M. Yuan, S. Ong, and A. Nee, :Assembly Guidance in Augmented Reality Environments Using a Virtual Interactive Tool", *Innovation in Manufacturing Systems and Technology*, 2005.

[24] J. Zauner, et al, "Authoring of a Mixed Reality Assembly Instructor for Hierarchical Structures", *Proceedings of the 2nd IEEE/ACM international Symposium on Mixed and Augmented Reality*, 2003, pp. 237.

[25] G. Zimmerman, "Some Assembly Required: Effectiveness of Interactive 3D Graphics on Mobile Devices for Object Assembly", *Proceedings of the Second International Conference on Computer Graphics Theory and Applications*, Barcelona, Spain, 2007, pp. 191–196.

[26] G. Zimmerman, J. Barnes, and L. Leventhal, "A comparison of the usability and effectiveness of web-based delivery of instructions for inherently-3D construction tasks on handheld and desktop computers". *Proceedings of the eighth international conference on 3D Web technology*, Saint Malo, France, 2003, pp. 49–54

Author Biographies

John Strobel was born in Stuttgart, Germany. He earned his BS in Computer Science (2009) and MS in Computer Science (2010) both from Bowling Green State University, in Bowling Green, Ohio, USA. He is currently working as an applications programmer in Cleveland, Ohio.

Guy W. Zimmerman was born in Saginaw, Michigan, USA. His got his BSE in Applied Mathematics from the University of Michigan, Ann Arbor. An MS in Applied Mathematics (1982), an MS in Computer Science (1986) and his Ph.D. in Computer Science (1990) all from Michigan State University, East Lansing. He is currently an associate professor of computer science at Bowling Green State University, in Bowling Green, Ohio, USA. His research interests include virtual reality and graphics with applications to HCI.