Effect of Various Instruction Methods on 3D Virtual Assembly Task

Majed Dhbeel Aldhamari
Colorado State University
Fort Collins, Colorado
majeda@colostate.edu

ABSTRACT
UPDATED—May 10, 2021. This study examines virtual assembly task completion rate when using different modes of instructions. The experiment conducted will evaluate the difference in completion time of a three-dimensional (3D) puzzle when using written instructions versus using a video for instruction. Understanding the benefit of various instruction modes can provide important insight for manufacturing companies who wish to train their employees in the most effective and efficient method.

Author Keywords
Human-Centered Computing; Human-Computer Interaction; Virtual Reality (VR); Augmented Reality (AR); Assembly Task; Instruction Modes; Manufacturing

CCS Concepts
•**Human-centered computing → Human computer interaction (HCI);** Assembly Task; User studies; Instruction Mode;

INTRODUCTION
Most manufacturing companies assemble their products using a blueprint for its design. A set blueprint for an object in the manufacturing industry means that each object manufactured must have a uniform shape, structure, and functionality. The people and teams that design the object to be manufactured are not always the same people that assemble the product. Also, the people and teams manufacturing the objects need not always understand the underlying design principles of the object nor the underlying functionality. Rather, assemblers only need to understand how to build the object efficiently and according to the provided blueprint. For this reason, communication between the product designers and assemblers must be as effective as possible. In order for this to occur, those that are directly manufacturing the object must receive instructions that are not only clear, but also efficient in their delivery. If instructions are unclear or difficult to parse, this could cost the manufacturing company a significant amount of capital. Therefore, reducing the time it takes to assemble the object is of high importance to the cost of development. Keeping assembly time low while strictly abiding by the presented object blueprint could prove highly beneficial to a modern manufacturing company. Most manufacturing companies must train their employees so that they are able to assemble their products correctly. As employees start to learn how to properly assemble products according to the designer's blueprint, they are likely to perform the process incorrectly during their training. This can lead to wasted materials which is a key contributor to added cost of development [6].

A possible solution to this issue is the use of virtual assembly tasks during the training of manufacturing company employees. Virtual assembly tasks can provide the necessary training for employees to learn how to assemble products without the added waste that occurs when using real materials. Furthermore, employees can be given more attempts to assemble the products as it will not be as costly. This will result in employees having faster assembly times. This will also result in more accurate final products due to the increased practice.

It is therefore vital to make sure that employees are given the highest quality instruction. One of the most important factors in assembling a three-dimensional (3D) object accurately and within a given time constraint is the quality of the instructions given to the person or team that will be assembling that object. This applies most directly to modern manufacturing industries where the object to be assembled has a set blueprint and functionality.

The original idea for this study was going to be testing different modes of instruction on assembly task completion time in a Virtual Reality (VR) or Augmented Reality (AR) environments where the participant assembles a completely virtual object. This is because VR and AR have shown a lot of promise regarding allowing a transfer of skills from a virtual environment to real life, especially when it comes to assembly of objects. This is mainly due to the spatial awareness that is necessary in VR and AR tasks as well as the ability to replicate the real world environment [1]. This approach would have required participants to complete the assembly task in this study using a VR or AR headset to view and interact with the environment. Due to the COVID-19 pandemic, it would have been difficult to attain a group of participants large enough to see significance from the results since each participant would have had to either have their own VR headset or meet for in-person testing. For this reason it was best to instead utilize a 3D environment without VR or AR to test the effectiveness of instructions. This approach allowed me to upload the entire playable game to the Unity Play Store using WebGL. I was also able to code a camera that rotates along all axes and the ability to move the game pieces using the computer mouse.

In this study, I will examine how different modes of instruction affect the assembly rate of a 3D puzzle cube. This 3D cube is similar to a 3D version of the game Tetris, but there will be a set number of pieces. A video as well as a pictorial manual of the puzzle solution is provided to participants of the study. Time and accuracy of the participants attempt using each instruction method will be recorded and evaluated.

LITERATURE REVIEW

Studies on the different methods of instructions that are given to a user for a 3D virtual assembly task focus primarily on the effectiveness of each mode of instruction, specifically how it affects the rate of task completion, the overall user experience, and how that virtual task can translate to the assembly of a similar object in the real world. Many studies have found that the method of instruction can make a significant difference in the rate of completion of a virtual task. One study found that the least effective method of instruction for both virtual and real-world tasks is a list of typed out steps [5]. Furthermore, there are reports from multiple studies that found that the instructions for a task should resemble the task itself in order to be the most effective. These studies found that a three-dimensional (3D) task should have instructions given in 3D to be most effective [2, 3, 9]. Another similar study found that instruction for an Augmented Reality (AR) task were most effective when the instructions themselves were in AR [11, 12]. Additionally, previous research has also found that the length of instructions should be minimal as to maximize a user's cognitive efficacy. In other words, a user should spend more time completing the task at hand rather than parsing through lengthy instructions [11]. While some complex tasks require some written instruction, presenting a user with a diagram, video, or an interactive tutorial could reduce the length of instructions as well as provide the most relevant information to the task itself [8].

The overall user experience of completing a virtual task can be significantly improved by modifying the instructions given to a user. It is important to mention that a large portion of ensuring an optimal user experience during a virtual task is focused around making the user interface as realistic as possible [12, 13]. One study found that a virtual task had higher user ratings when the objects in the task were restricted by the laws of physics which increases the realism of the virtual environment which allows for a better user experience [4]. Research has shown that the most effective instruction can differ based on the user's demographic. One study found that older adults showed better outcomes and gave more positive feedback about the task when given more encompassing instructions, while these same instructions for younger adults were seen as redundant [10]. Therefore, it is important to allow users to choose the amount and medium of instruction as different demographics may have differing optimal instructions sets.

Many of the most beneficial applications of a 3D assembly task is to train people on the assembly of an object virtually prior to them assembling the same or a closely related object in the real world. A prior study has shown that in order to make a virtual task transferable to the real world, the timing of when the instructions are given can make a significant difference. The instructions given prior or during the task also make a significant difference. One study found that giving instructions throughout the task is beneficial only in a virtual environment, while this served as a cognitive interference in the real world [7]. Another study found that the use of animated instruction that allow interaction between real world and virtual components leads to a much better rate of transfer to real world assembly skills. [1]

METHODS

Environment

This project was built in a 3D environment using Unity 2020. Unity is a cross-platform game engine that was first released in 2005 and can be used to create 2D, 3D, VR, and XR games and models. I had multiple designs for the puzzle itself. One of the designs was a replica of the Roman Parthenon which is a Roman temple that is now a church. The Roman Parthenon level became the practice level in which participant would get a feel for the controls of the game. The second level I designed was assembling a B-52 Bomber airplane. This level ended up being too complicated for participants during a usability test and was therefore not included in the final edition of the game.

When the game is first started, the player is presented with a main menu that has the title of the game as well as two buttons that start either the practice Parthenon level or the main Tetris Cube level. The main Tetris Cube puzzle was the level in which assembly time and error rate were to be observed. The entire puzzle was made from the t-block tetris tetrimino from the original Tetris game released in 1984 and still popular today (see Figure 1). The game objects to create these t-blocks were two cube game object primitives that were scaled and combined together. These puzzle pieces had the RigidBody component applied to them which gave them realistic physical properties such as responding to the forces of gravity. These game pieces were also given a Collider component which gave them realistic reactions to collisions with both the ground and other game pieces.

Figure 1. Tetris t-block tetromino from the 1980's game Tetris

The controls to move the game object was clicking and dragging the puzzle pieces. To allow for such control, I wrote a script in C# which compared the object position to both the mouse position and camera view. The camera was a cinemachine that could be controlled along each axes, that is the x, y, and z planes. Two cubes were combined and manipulated

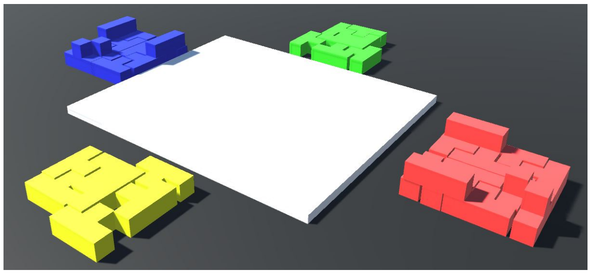

Figure 2. Tetris Cube scene upon entering the level

to create the platform on which the puzzle was to be built and serve as the focal point for the camera and the center of the scene. The camera to view the puzzle is controlled with the keyboard arrow keys. Participants are able to rotate around the center of the platform in the horizontal plane. Participants were also able to move the camera up and down on the vertical plane in order to have access to both a side view and a bird's eye view of the puzzle.

In the top right of the screen within each level was a timer that starts as soon as the player enters the level. There were also two buttons to stop and start the timer. To include this timer I had to write a script in C#. The timer consisted of multiple primitive user interface (UI) game objects.

Participants

Participants were recruited from a variety of sources. Most participants will be students at Colorado State University. The student participants were mainly juniors and seniors, however there were a few that were freshmen and sophomores. There were also some students recruited from Colorado School of Mines. Some of the participants were also non-students. Requests for participation were sent online via email. People who responded to the email were assigned a group of either "written first" or "video first." Those in the written first group would solve the puzzle using the written instructions first, then solve the puzzle the second time using the video instructions. Participants that were in the video first group solved the puzzle using the video instructions first then solve the puzzle the second time using the written instructions. The grouping of participants using a latin-square method was done in order to counter act the priming effects of solving the puzzle the second time.

Solution

Each layer of the Tetris cube was colored a different color. There was a total of four layers. The first layer was blue, the second layer was green, the third layer red, and the top layer was yellow. Puzzle pieces belonging to each layer were frozen at a certain height (y-axis). This height was only revealed after the game piece was selected. Before the puzzle piece is selected all the pieces were at the same level on the terrain (see Figure 2). Players needed only to find the appropriate two dimensional position along the x and z axes after discovering the appropriate height of the puzzle piece.

Each layer must be constructed in the correct order for the puzzle to be solved. Participants begin with the blue layer by dragging each piece to its appropriate position. Once the blue layer is completed, participants assemble the green layer on top of the blue layer, the red layer on top of the green layer, and finally the yellow layer on top of the red layer. Some pieces from the previous layers will be part of the next layer. For example, layer 2 will include both blue and green pieces. Once all layers are assembled, a cube will be the result with no pieces being left over (see Figure 3)

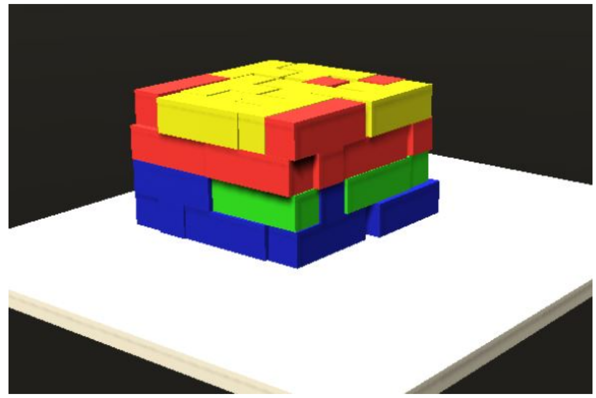

Figure 3. Tetris Cube after completing the level

Experiment Design

Once the puzzle is solved, participants send a screen shot of the game for both attempts. Time is recorded by the onscreen timer. Accuracy of the final solution can the be evaluated. The evaluation of accuracy is based on how many layers were assembled correctly. Since there are four layers, accuracy was measured from 1-4. For example if a participant assembled all four layers correctly, there accuracy score would be a 4 but if they only assembled 2 layers correctly, their score would be a 2.

This is a within subject study in which each participant will have their scores for each level of the independent variable compared. There is one independent variable that is instruction method with two levels that are written instruction and video instruction. There are two dependent variables that are completion time and accuracy as described above. The null hypothesis is that there will not be any difference in puzzle completion time when participants are using either instruction mode. The alternate hypothesis is that there will be a difference in puzzle completion time depending on whether participants are using video instructions versus written pictorial instructions. I will use a two-tailed t-test with alpha set at .05 to examine the significance of any differences I may find.

RESULTS

A total of 18 individuals participated in this study. All participants were between the ages of 18 to 28 and had relatively decent amount of exposure to video games and computers. Accuracy was also recorded for reference to outlier times. No hypothesis test was conducted on accuracy level due to the subjective means of data collection for this variable. Results for each participant can be seen in Table 1 and Figure 4. The mean puzzle completion time while using written instructions

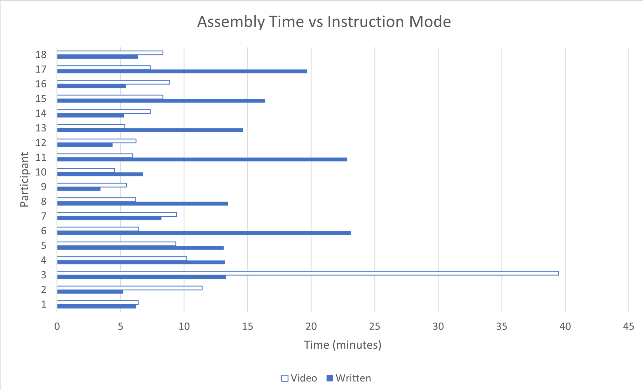

Figure 4. Video versus written instructions for each participant

was 11.16 minutes with a standard deviation of 6.39. The mean puzzle completion time while using video instructions was 9.25 minutes with a standard deviation of 7.77.

I used a two-tailed paired t-test to evaluate the significance of the difference in means of puzzle completion time in each instruction mode. I set $\alpha = .05$ and obtained a t statistic of 0.8259 (df = 17) $p = 0.4203$. As a result, I failed to reject the null hypothesis and could not conclude that the difference in puzzle completion time based on instruction mode was significant.

	Written		**Video**	
Name	*Time (minutes)*	*Accuracy (1-4)*	*Time*	*Accuracy*
Mark L.	6.22	1	6.39	1
Boston G.	5.21	4	11.41	4
Abby W.	13.29	4	39.47	2
Leah C.	13.21	4	10.22	4
Parker J.	13.11	4	9.34	4
Muath E.	8.21	4	9.41	4
Ming Z.	13.44	4	6.19	4
Kordell C.	23.12	3	6.43	4
Corey D.	3.41	1	5.45	2
Moe M.	6.76	3	4.52	3
Ali M.	22.84	4	5.95	4
Ahmed A.	4.37	2	6.22	3
Sayud G.	14.63	4	5.33	4
MEAN	**11.16**	**3.056**	**9.25**	**3.223**

Table 1. Completion time and accuracy for each mode of instruction.

DISCUSSION

There was a very large distribution of puzzle completion times among participants in both instruction modes. The lowest puzzle completion time with written instruction was 3.40 minutes and the highest was 23.12 minutes. The lowest puzzle completion time with video instruction was 4.52 minutes and the highest was 39.47 minutes. The standard deviations for each instructional methods puzzle completion time were also high, SD=6.38 for written and SD=7.76 for video. The puzzle completion times for both instruction modes had outliers that contributed to this large distribution of results.

While the mean puzzle completion time was higher in the written group than then video group, these results were not found to be significant. Some things that could have contributed to this lack of significance could be a small participant pool. Another possible contributor to this lack of significance could be the large distribution of results.

Puzzle accuracy was a score that I gave each participant based on what the puzzle looked like when they stopped the timer. For each level that was correctly assembled, the participant would receive 1 point for a maximum of 4 points. A score of 4 indicates perfect assembly while a score of 1 indicates that only one layer was assembled correctly. Most participants received a score of 3 or higher with both instruction modes. This score ended up being too subjective to analyze, however it still gives some insight on the reason for some of the abnormally small puzzle completion times.

LIMITATIONS

The wide distribution of results leads me to suspect that there were various confounding variables affecting participants at the time of data collection. Participants were able to access the experiment online at the Unity Playstore. This allowed participants to take the experiment in the comfort of their own homes, at the university library, or even a coffee shop. Although this experiment design was necessary due to the ongoing COVID-19 pandemic, it resulted in the environment of the experiment not being controlled, which means that the testing environment could have been a confounding variable.

Another possible confounding variable in this experiment was the absence of a win condition in the game. Participants had a subjective definition of completing the puzzle. Other than the assembly instructions, there was not a device in game that informed the participant that they had in fact completed the puzzle the correct way. Some of the shorter finish times could be attributed to this in both the written and video instruction groups. Furthermore, participants were instructed to stop the timer once they believed that they completed the puzzle. Self timing may have skewed the final results as well.

A third possible limitation of this study was the small amount of participants recruited. Invitations to participate in this study were sent over a two week window. A total of 18 participants sent back their results within this time frame. This is considered a small number of participants which is likely to produce insignificant results.

One last, and probably the most important limitation, was the within subjects design of the experiment. Since the same puzzle was played twice by the same participants, puzzle completion time was typically faster the second attempt. Although the instruction mode was alternated for participants using a latin square counterbalancing in an attempt to reduce priming effects, a between subjects design would have removed the priming effects of the first trial on the second trial completely.

CONCLUSIONS AND FUTURE WORK

This study, although not providing significant results, allowed for interesting insight into how people interpret and apply assembly instructions. There was a clear advantage when solving the puzzle the second time regardless of whether participants used the written instructions or video instructions first. Once participants completed the first trial, they gained some practice and had much lower completion times and higher build accuracy in the second trial as a result.

There are a few recommendations for future work. The first is to control the testing environment by having it take place in a lab under supervision. This would reduce the environmental effects on the participants while they are completing the puzzle. Another recommendation would be to set an in-game win condition so that the finish times are not skewed by each participants subjective definition of solving the puzzle. This win condition would also automatically stop the timer which would remove the confounding effects of participants self timing. Another recommendation would be to increase the recruitment time window as to recruit more participants. This leads me to the final recommendation which would be to conduct a between subjects experiment in which each participant only uses one type of instruction mode to complete the puzzle. This experiment design would require double the amount of participants in this study or more. A final recommendation would be to conduct this study in an AR or VR environment as this would be of more benefit to the manufacturing industry as AR and VR both show an increased transfer of skills when compared to 3D.

REFERENCES

[1] Andrijana Bocevska and Zoran Kotevski. 2017. Implementation of interactive Augmented Reality In 35 Assembly Design Presentation. *AIRCC's International Journal of Computer Science and Information Technology* 9 (2017). DOI: http://dx.doi.org/10.5121/ijcsit.2017.9213

[2] Jorge Camba, Manuel Contero, Michael Johnson, and Pedro Company. 2014. Extended 3D annotations as a new mechanism to explicitly communicate geometric design intent and increase CAD model reusability. *Computer-Aided Design* 57 (2014). DOI: http://dx.doi.org/10.1016/j.cad.2014.07.001

[3] J. Geng, S Zhang, and B. Yang. 2015. A Publishing Method of Lightweight Three-Dimensional Assembly Instruction for Complex Products. *ASME. J. Comput. Inf. Sci. Eng.* (2015). DOI: http://dx.doi.org/10.1115/1.4029753

[4] Nicolas Imbert, Frederic Vignat, Charlee Kaewrat, and Poonpong Boonbrahm. 2013. Adding Physical Properties to 3D Models in Augmented Reality for Realistic Interactions Experiments. *Procedia Computer Science* 25 (2013). DOI: http://dx.doi.org//10.1016/j.procs.2013.11.044

[5] S. A. Konz and G. L. Dickey. 1969. Manufacturing Assembly Instructions; A Summary. *Ergonomics* 12 (1969), 370–382. Issue 3. DOI: http://dx.doi.org/10.1080/00140136908931061

[6] Mike Oren, Patrick Carlson, Stephen Gilbert, and Judy Vance. 2012. Puzzle assembly training: Real world vs. virtual environment. *Mechanical Engineering Conference Presentations, Papers, and Proceedings.* 84 (2012). DOI: http://dx.doi.org/lib.dr.iastate.edu/me_conf/

[7] F. D. Rose, E. A. Attree amd B. M. Brooks, D. M. Parslow, and P. R. Penn. 2000. Training in virtual environments: transfer to real world tasks and equivalence to real task training. *Ergonomics* 43 (2000). DOI:http://dx.doi.org/10.1080/001401300184378

[8] T Salonen and J Saaski. 2008. Dynamic and Visual Assembly Instruction for Configurable Products Using Augmented Reality Techniques. *Advanced Design and Manufacture to Gain a Competitive Edge* (2008). DOI: http://dx.doi.org/10.1007/978-1-84800-241-8_3

[9] T Shao, D Li, and Y Rong. 2015. Dynamic Furniture Modeling Through Assembly Instructions. *ACM Transactions on Graphics* 35 (2015). DOI: http://dx.doi.org/10.1145/2980179.2982416

[10] Marion Verneau, John van der Kamp, Geert J.P. Savelsbergh, and Michiel P. de Looze. 2014. Optimising assembly learning in older adults through the manipulation of instruction. *Ergonomics* 57 (2014). DOI: http://dx.doi.org/10.1080/00140139.2014.924573

[11] Z Wange, X Bai, and S Zhang. 2020. Information-level AR instruction: a novel assembly guidance information representation assisting user cognition. *Int J Adv Manufacturing Technology 106* 106 (2020). DOI: http://dx.doi.org/10.1007/s00170-019-04538-9

[12] Z.B. Wange, S.K. Ong, and A.Y.C Nee. 2013. Augmented reality aided interactive manual assembly design. *Int J Adv Manufacturing Technology* 69 (2013). DOI:http://dx.doi.org//10.1007/s00170-013-5091-x

[13] M. L. Yuan, S. K. Ong, and A. Y. C. Nee. 2008. Augmented reality for assembly guidance using a virtual interactive tool. *International Journal of Production Research* 46 (2008). DOI: http://dx.doi.org/10.1080/00207540600972935