Are mobile Augmented Reality instructions better? Comparing the effects of AR instructions and paper instructions to guide an assembly task

Student Name: Yumeng Yang
Student Number: s2005158
First Supervisor: Joyce Karreman
Second Supervisor: Menno de Jong

UNIVERSITY OF TWENTE.

Acknowledgment

Foremost, I would like to express my sincere gratitude to my supervisor Joyce Karreman and Menno de Jong for the continuous support of my master study and research, for their patience, motivation, enthusiasm, immense knowledge, and sharp comments. Their guidance always encourages and reminds me that life is finite but knowledge is infinite.

Besides my supervisors, I would like to thank 72 participants who joined my research experiment. Without their help, I could not have collected data and finished this research. My sincere thanks also go to Jan Kolkmeier and Qian Li, for offering me extra research support.

In addition, I would like to thank my idols. They are Arashi, a Japanese boy band, Taylor Swift, an American singer, and Avril Lavigne, a Canadian singer. Every time when I feel unmotivated or discouraged, their music and movies inspire me to keep holding on.

Last but not the least, I would like to thank my parents, for encouraging me to study aboard and supporting me spiritually throughout my life. I really miss my dad because he is a "Chinese Cuisine Master" and he cooks delicious food for my mom and me all the time. Although my mom always bothers me and wants to have a video chat when I am studying, I know she also misses me too much.

Content

Abstract

Purpose: Augmented Reality (AR) has gained increasing attention as a means to provide user support and assistance in the domain of manufacturing assembly. To date, the usefulness of mobile AR instructions used in assembly contexts has not been systematically investigated and there are very few empirical studies. This research aims to bridge the gap by comparing a paper-based manual to a mobile AR manual. The purpose of this research is to evaluate the usability and users' perception of a mobile AR instruction in guiding assembly tasks.

Method: A mobile AR instruction and a paper-based instruction for a LEGO assembly task were created. 72 participants were recruited. They were divided into 2 groups randomly, namely, a paper-based group and an AR group. Participants had a training session so as to familiarize themselves with instructions firstly. They were then required to finish a LEGO™ assembly task with the assigned instruction to measure *effectiveness* and *efficiency*, fill in the questionnaires to evaluate *perceived cognitive load*, *motivation,* and *instruction experience*. All participants were observed by the researcher. Finally, a short semi-structured interview was performed.

Result: Although the mobile AR instruction did not show significant differences in *overall cognitive load*, it increased the participants' *mental demand* and satisfaction of *performance* when compared to paper instructions. In addition, the AR instruction improved task *effectiveness* significantly. Furthermore, the mobile AR instruction increased users' *positive experience* significantly such as the feelings of playful, surprised, and joyful. However, this mobile AR instruction did not show significant differences in *efficiency, and motivation* when compared with a paper instruction.

Conclusion: This research suggested that current mobile AR instructions are indeed capable of improving task performance and improve the positive experience of users. At least, this study has shown that for people who have no prior experience with an assembly task, AR instructions increase their task accuracy and positive experience. It is advised that designers of assembly instructions consider the mobile AR instruction as an alternative version.

Keywords: Instructions; Augmented Reality; mobile AR instruction; assembly task; usability

1. Introduction

Augmented reality, an emerging technological tool, has been applied in the domain of technical documentation in recent years. AR instructions become another new type of technical documentation. El Sayed, Zayed, and Sharawy (2011) explained augmented reality as a technology that adds virtual objects to real surroundings. It should have supplementary data that is overlaid on the real world context. In this research, AR means virtual images made by computers are merged with the real view. Not only virtual images but also more information such as graphics, audio, and touch are superimposed over a real environment.

With the advent of AR, it has been used in many domains such as engineering, medical education and so on. Therefore, it is not difficult to understand what AR is in daily life. A most common example is Pokémon Go (See Figure 1 left). This AR application allows users to use their mobile phones to browse and search surroundings and then catch virtual Pokémon, thereby viewing virtual objects in a real environment. Another popular application is IKEA Place (See Figure 1 right). Users use mobile devices to scan the floor and then place the virtual furniture at the right place. In this way, they can preview how their house looks like after placing the new furniture.

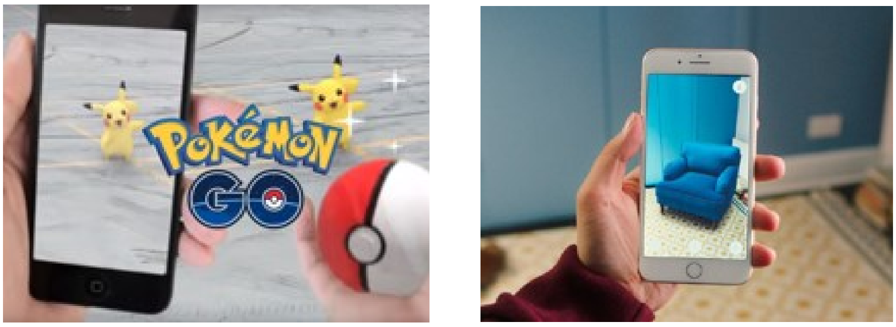


Figure 1 Examples of AR Application

Apart from those AR applications mentioned above, another promising application of AR is in the industry environment, including manufacturing assembly, equipment maintenance, and procedural instructions. Traditionally, new workers have two main ways of training before entering the workflow of maintenance and repair (Funk, Kosch, & Schmidt, 2016). One way is to learn from more experienced colleagues (McCalla et al., 1997). Another way is to refer to paper manuals or printed blueprints. However, these instruction channels have drawbacks. With the increasing number of produced variants and turnover of staff, it is unrealistic to consult experienced colleagues continuously. As for paper-based assembly instructions, searching for the correct manuals is cumbersome because of a large number of products. In order to increase the productivity of assembly, AR assistive systems for assembly have been proposed by many researchers (Funk et al., 2016; Hou, 2013; Herrema, 2013; Tang, Owen, Biocca, & Mou, 2003a).

With the development of AR technology and mobile hardware, mobile augmented reality has become a new type of AR tools. Mobile devices such as smartphones and tablets are becoming popular augmented reality tools as they fit users' needs such as portability, positional sensors, tracking and networking capability (Kim, 2013). In addition, these mobile devices are relatively cheap, flexible and accessible. Users can download AR applications on their mobile devices without any effort and cost. Therefore, mobile augmented reality has been used in cultural heritage onsite guides such as in museums and art galleries. More and more games are embedded with augmented reality to improve users' engagement and immersion. It can be assumed that mobile augmented reality can also be used in assembly instruction.

However, few studies focus on the usefulness of mobile augmented reality in assembly training. To bridge the gap, this research will investigate mobile AR instructions as an instructional medium in assembly tasks: What is the effect of AR instructions on *task performance*, *instruction experience*, *cognitive load*, and *motivation*? This study aims to provide four key contributions to our understandings of the AR instructions:

1. Do mobile AR instructions improve task efficiency and effectiveness when compared with traditional paper instructions?
2. What is the effect of mobile AR instructions on cognitive load when compared with traditional paper instructions?
3. What is the effect of mobile AR instructions on motivation when compared with traditional paper instructions?
4. How do users perceive mobile AR instructions and paper instructions?

To answer the aforementioned questions, a mobile AR instruction is developed in this research. By using mobile devices (tablets), the relevant information about assembly tasks such as textual information and 3D models can be available all the time. The purpose of this research is to compare the usability of a mobile AR instruction and a paper-based instruction in guiding basic assembly procedural tasks. The effects are measured on five metrics: *effectiveness, efficiency, cognitive load, motivation,* and *instruction experience*.

In the next section, the antecedent literature that has investigated the impact of AR instruction on task performance, cognitive load, and motivation are reviewed. Based on previous literature, hypotheses and research questions of this research are also presented in this section. In Section Three, the methodology used in the study are explained, followed by the result in Section Four. The study is discussed in Section Five. Finally, Section Six and Section Seven present the conclusion and limitations.

2. Theoretical Framework

In this part, key concepts in this thesis and their definitions are discussed, such as assembly task, cognitive load, and motivation. Besides, existing theories and literature that are relevant to the research questions of Introduction part are reviewed, which provides a theoretical basis for the hypotheses and research questions.

2.1 Activities of Assembly Tasks

An assembly task is a process of joining components or parts together to perform specific functions. In practice, assemblers refer to an assembly manual to perform assembly steps (Laperriere & ElMaraghy, 1992). In this context, the implementation of an assembly task can be divided into two type activities, namely, non-assembly-related activities and assembly-related activities (Neumann & Majoros, 1998). Assembly-related activities are directly related to operation activities that are physical, while non-assembly-related activities tend to be information-related and cognitive activities (Neumann & Majoros, 1998). For instance, an assembler not only executes physical operations such as alignment and installation but also mental activities (e.g., reading, translating and retrieving information).

The drawbacks of consulting an assembly manual have been identified by many researchers. Specifically, three defects of assembly manuals are suggested. The first drawback suggested by Zaeh and Wiesbeck (2008) is that using a manual during assembly task introduces more attention switching. Continual visual transitions could be distractors, which results in operational suspensions. Secondly, due to the limited size of papers, information context of procedures is scattered on different pages, increasing the difficulties of information orientation. As a result, apart from the necessary movements like picking up components from workpiece areas to assembling areas, assemblers have to undertake kinetic operations that are non-assembly-related actions such as paging up/down and comparing information on different pages to understand the whole process. (Hou, Wang, Bernold, & Love, 2013). The last drawback of paper-based instruction is that a planar representation has limited ability to visualize procedural information such as motion path and assembly process and etc., forcing readers to spend more time on information interpretation and analysis. The visual transitions, extra non-assembly-related actions, and limited information visualization caused by manuals reflect the time-consuming nature of paper instructions. All these drawbacks hinder people's information retrieval activity.

What is an effective retrieval activity? According to research, an effective retrieval activity is defined as a set of fast mental behaviour, including searching, analysing and interpreting information (Hou, 2013). Normally, operating information is separated from tools so that assemblers have to search a certain type of medium for information. The medium to access information is often a printed manual. However, as the drawbacks mentioned in last paragraph, information retrieval while using paper instructions seems to be unfavourable. Besides, Veinott, Kanki & Shafto (1995) also noticed that the shifts of assemblers were mostly spent on retrieval, reading procedural information when assembling components, which contributes to productivity losses. Furthermore, Watson, Curran, Butterfield & Craig (2008) pointed out that such increasing information retrieval for a complex process can trigger tiredness and the tendency to commit errors. Similarly, Veinott et al. (1995) identified that 60% of the errors are caused by misunderstanding. Such misunderstanding mostly arises from the unfavourable information retrieval.

Based on the discussion above, it is clear that a paper-based instruction has a time-consuming nature and the tendency to trigger more errors. These features potentially hinder an assembler's task effectiveness and efficiency. Therefore, the way in which assembly procedural information is presented to a worker significantly influences operational effectiveness and efficiency. How about an Augmented Reality instruction? Can we use AR technology to decrease the drawbacks of traditional instructions?

Several research groups have explored the benefits of AR in assembly industry. Tang et al. (2003) compared paper-based instructions with AR instructions. They concluded that 3D instructions overlaying virtual objects on the actual assembly area reduced error rate significantly. Likewise, it is also identified that AR instruction allowed assemblers to finish tasks quickly and resulted in less head movement. (Henderson & Feiner, 2011; Hou et al., 2013; Marner, Irlitti, & Thomas, 2013).

Compared with paper-based manuals, three benefits of an AR instruction can be summarized. Firstly, the information retrieval activity can be integrated with assembly operating via dynamic animation as guidance. Such integration can significantly smooth visual transitions and reduce the time consumed. The reason behind this is that AR visualization provides a consistent and dynamic representation of information context, which allows users to spend less time on information retrieving. This feature is particularly useful for people who have low information retrieval capacity. Secondly, AR instructions have the potential to reduce operation errors, which especially benefits novices a lot. Since AR instructions adopt stereoscopic components that are designed as real-scale objects in size and shape, it is easy for non-experienced users to distinguish the physical components. Furthermore, special hints such as motion cues and arrows enable the important dimensions of components more distinct. Specifically, virtual objects can be selectively rendered or omitted so that the superfluous parts are less distracting. For instance, more important parts can be animated while less important ones can be static. The theoretical support to this is that irrelevant stimuli and dispensable retrieval behaviours lead to poor task performance. Assemblers should be aided to focus on relevant objects and ignore irrelevant objects (Haider & Frensch, 1996). Therefore, AR instructions facilitate ongoing tasks and reduce errors. Thirdly, AR instructions with animation can contribute to the 3D representation of a procedural step concretely, which can aid users to interpret information when performing a procedural task. Users only need to mimic each assembly step. It is identified that animated digital manuals allow participants to accomplish the task faster and more accurately (Lee & Shin, 2012).

AR instructions also show advantages when compared with video instructions. Although video tutorials are effective since they harness the power of animated visualization, it is hard for users to control such animation. For instance, users have to go back and forward repeatedly to verify their understanding of each step since most animation tend to be fast. In addition, videos have limited interactions with users, since they only allow users to stop, play, speed up, and speed down. Users cannot interact with the content such as rotate the model, change steps, and etc. By contrast, AR provides more opportunities for users to control the content and even interact with the content. It can be assumed that AR instructions perform better than video tutorials.

The previous literature mentioned mainly used Head-Mounted Display equipment to implement AR instructions. The conclusions may not be generalized to a mobile AR instruction. The different experiment equipment may lead to different conclusions. When considering an animated mobile AR instruction, the findings may be different. Therefore, more research is needed before we can conclude with certainty that mobile AR instructions benefit users in terms of effectiveness as well as efficiency. In this study, two hypotheses are formulated as follows:

H1: When compared with a conventional paper-based instruction, a mobile animated AR instruction will result in an increase of *efficiency* for an assembly task.

H2: When compared with a conventional paper-based instruction, a mobile animated AR instruction will result in an increase of *effectiveness* for an assembly task.

Besides the measurements mentioned above, it is also interesting to know *how users perceive the instructions they use.* How do they think of an AR instruction? What other influence can instructions bring to users except for effectiveness and efficiency? Extant studies restrict research on traditional evaluation indicators such as efficiency and effectiveness when evaluating AR instructions. New evaluation indicators have suffered from insufficient consideration. To fill the research gap, a new exploratory indicator called *instruction experience* is proposed in this research. This indicator aims to evaluate mobile AR instructions from four aspects: *ease of use, positive experience, negative experience,* and *behaviour intention* (See Appendix VII). This exploratory approach allowed the researcher to compare the effects of mobile AR instructions and paper instructions from a new perspective.

2.2 Cognitive load

Cognitive Load Theory (CLT) explains the relationship between learning and human cognitive architecture (Sweller, 1994). CLT consists of three types of cognitive load: extraneous cognitive load, intrinsic cognitive load, and germane cognitive load (Sweller, Van Merrienboer, & Paas, 1998). Extraneous cognitive load is caused by the format and manner in which instructional materials are shown to users. Intrinsic cognitive load is determined by the complexity of the learning materials. Learning materials with high complexity require users to hold more mental resources. Germane cognitive load refers to working memory resources that are used to deal with intrinsic rather than extraneous cognitive load (Sweller et al., 1998). In this study, cognitive load is defined as the amount of mental processing required to process an assembly task.

As an integration of real and virtual environment, AR possesses three unique features such as representing information spatially, adding multiple sensory modalities and eliminating the split-attention effect. According to these properties, researchers proposed that AR has the potential to reduce the learners' working cognitive load caused by mental rotation when processing spatial information, optimizing cognitive load for users. This is because different presentations of instructions will induce various working memory load and mental processing. The goal of adopting AR is to reduce extraneous cognitive load and make instructions easier to understand, and then the germane load is optimized. Some researchers have evaluated the cognitive effect of AR. However, their conclusions about whether AR instructions can reduce cognitive load are mixed.

Haniff and Baber (2003) performed a user evaluation comparing paper-based instructions with video see-through AR instructions on a computer monitor. Results indicated there was a less cognitive load when using AR instructions. This is because participants had to translate information mentally more when consulting traditional paper manuals. This was, however, not the case with AR instruction. The AR system offered a complete and concrete representation of the task such as motion direction and spatial structure of an object, participants gained a better understanding of operations and distinguished each to-be-assembled object more easily. Therefore, such full representations with 3D information relieve the cognitive load. This result is in line with the finding of another study from Wickens & Hollands (2015). Wickens, Hollands, Banbury, & Parasuraman (2015) found that two-dimensional representation of information required more mental effort when constructing a three-dimensional world. AR provides full representation with 3D information, which can reduce extraneous cognitive load. In this way, learners have more working memory resources to deal with germane processing.

In a similar vein, Tang et al. (2003) tested the effectiveness of AR when it is used as an instructional medium. The result showed that the AR system was less mentally demanding. One reason is that AR can reduce the mental effort of object location since virtual cues such as arrows and motion path ease the ongoing tasks. Another reason could also be observed in the elimination of split-attention effect that has been discussed previously.

However, Blattgerste, Strenge, Renner, Pfeiffer, and Essig (2017) found a contradictory conclusion. They claimed that the perceived cognitive load of paper instructions was the lowest, while AR instruction resulted in relatively higher cognitive load. A presumably reasonable explanation is that the equipment they used in the experiment has a limited field of view, making AR instructions more mentally demanding. Similarly, Funk et al. (2017) suggested that the in-situ instruction slowed down workers’ task completion speed and increased the perceived cognitive load.

This negative impact can be seen especially for expert workers who have already known how to perform a procedural task. Researchers pointed out that the possible reasons could be that expert workers were used to the old assembly line and the in-situ instructions were distracting them.

In another AR system, Dunleavy, Dede, & Mitchell (2009) reported that AR increased the cognitive load of participants. This high cognitive load could be attributed to the insufficient preparation and unfamiliarity of system and task. Likewise, Huk, Steinke, & Floto (2003) found that learners with different spatial ability perceived different mode of visual representation. For instance, some learners prefer simple modes of representation, such as 2D pictures rather than 3D content or animations. This may be because a complicated presentation mode leads to information overload so that learners are unable to extract information that they need. Therefore, they chose to exclude additional animations and 3D objects.

To summarize, current research has not unequivocally shown whether AR instructions reduce cognitive load and conclusions are mixed. Besides, there is no empirical research that has investigated the effect of mobile Augmented Reality on cognitive load when performing an assembly task. *What is the effect of mobile AR instructions on cognitive load when compared with traditional paper instructions?* Will a mobile AR instruction cause higher cognitive load when compared with a paper-based instruction? Will animations and 3D virtual objects with detailed texture and shadows lead to information overload? These questions still need to be answered. To fill the research gap, more empirical research should be conducted to find out the effect of AR instructions on cognitive load.

2.3 Stimulation of Motivation

In technical communication, motivation plays an important role, which can promote an effective communication process. Goodwin(1991,100) pointed that "*technical communicators should keep a reader reading long enough and carefully enough to become competent at specific tasks*." In another word, a good user manual should possess an engaging and attractive experience which can retain users' attention. Technical communicators should strive to design instructions that motivate users to notice and put enough effort into performing their tasks. In this research, the definition of motivation is stated that "*motivation provides a source of energy that is* *responsible for why learners decide to make an effort, how long they are willing to sustain an activity, how hard they are going to pursue it, and how connected they feel to the activity*." (Rost, 2006).

Researchers have reached a consensus that if learners are motivated to learn and do tasks, they are more likely to persist and spend more effort when completing tasks (Chickzentmihalyi, 1990; Efklides, Kuhl, & Sorrentino, 2001; Keller, 1979; Schmidt, 2007). When people experience a pleasant emotion such as motivation and interest, they are more likely to view surrounding things with a positive state of mind. By contrast, if a learner is not motivated, he/she is hard to engage in learning tasks, keep persistence, and patience, which will lead to more errors and lower efficiency. When it comes to assembly tasks, assemblers always need to solve mechanical problems which are complicated and repetitive. Information retrieval and continual visual transition result in impatience and therefore suppress motivation. Assemblers are more likely to feel bored and stressful. In this situation, a manual with motivational elements tend to be necessary and useful to stimulate assemblers' interests and attention. As a result, they not only want to engage in work when everything goes well and smoothly but also want to persist and keep trying when they encounter setbacks.

In the instructional manual arena, many researchers highlighted the benefits of instructional manuals with motivational elements. For instance, Loorbach, Steehouder, & Taal (2006) found that an instructional manual with motivational elements enhanced the users' appreciation for the manual, although users' task performance was not influenced by those elements. In a later study, Loorbach, Karreman, & Steehouder (2007) pointed that a motivational instruction can increase confidence and help elderly users to persist in operating the device. They advised technical writers to add some elements to make user manuals more motivational. Therefore, when composing an instructional manual, the aspect of motivation should not be ignored.

In order to make instructional manuals more motivational, researchers proposed many strategies such as adding animated pedagogical agents, showing empathy to readers, and using multimedia. Augmented reality, an emerging technology, has been also used in technical documentation field. AR technology allows users to view a computer-generated image in a physical real-world environment, interact with 2D or 3D virtual model, and enhance our sense (vision, hearing and tactile). This new technology has introduced a new way of representing technical information, which tends to increase the attractiveness of user manuals and improve users' attention and motivation.

Many researchers in education field explored the motivation effect of Augmented Reality. According to their research, we can summarize three reasons why AR instruction can increase motivation. Firstly, AR technology can increase interaction between materials and users. Augmented reality is a technology that integrates virtual computer-generated objects and physical real-world context (Milgram & Kishino, 1994). Compared with traditional paper-based instructions that are static, users have more opportunity to interact with virtual objects in a real-world environment. For instance, when using a mobile AR application, virtual objects can be rotated, moved and scaled by clicking buttons on UI interface. Through such interaction, a "natural" experience can be generated, which results in the increase of effectiveness and attractiveness of learning. Thus, the attention and motivation are both improved (Sumadio & Rambli, 2010). Dunleavy, Dede, and Mitchell (2009) also found that such physical interaction with the AR instructional materials made learning authentic and motivating. Therefore, AR technology has the potential to provide users with more meaningful interaction in an assembly environment.

Secondly, using AR systems smooth visual transition so as to reduce tiredness and improve readers' motivation. Due to the baldness and frequent repetition of traditional reading materials, the task motivation is suppressed to some extent (Locke, 1968). Unlike paper-based interaction and computer-based interactive technology that require users to focus their attention on paper or a screen, AR instruction uses a tangible interface to view virtual objects in a real environment, which can smooth transition between reality and virtuality (Billinghurst, 2002). Thus, AR is a promising technology to improve the motivation and interest of learners (Pérez-López, Contero, & Alcaniz, 2010).

Last but not least, AR technology has the ability to facilitate immersion, which can foster learners' motivation and engagement (Barab, Thomas, Dodge, Carteaux, & Tuzun, 2005; Huang, Rauch, & Liaw, 2010; Shen & Eder, 2009). Chignell and Waterworth (1997) suggested that multimedia can convey information as well as increase motivation and interest of users or operators because of rich sensory experience and multiple modalities. Augmented reality uses digital methods to superimpose virtual and natural information on the real work. This can not only enhance our senses such as vision, hearing, and tactile but also provide immersion that might help to engage learners in learning activities (Azuma, 1997). In this way, learners are more likely to maintain high attention and interest towards learning content. For instance, lighting, object shadow, animations, UI interface, and other elements can be included so as to make the AR visualization more natural and realistic.

Based on the aforementioned studies, researchers in the educational field have developed and evaluated their AR educational applications successfully. When students were using these AR applications, they thought AR was fun, engaging and interactive. Researchers suggested that AR instructional material is a good alternative to conventional paper materials. However, these conclusions are all drawn from the education field which is a reading-to-learn setting. When it comes to instruction manuals, a typical read-to-do setting, can their conclusions still be applied? *What is the effect of mobile AR instructions on motivation when compared with traditional paper instructions?* Can AR technology produce the similar motivation effect when it is used to represent procedural information to assemblers? What values can AR instruction bring to us? These questions are still unknown.

There is no research that evaluates whether mobile AR instruction can increase users' motivation. To fill the gap, this study will investigate the motivation effect of this new type of instruction on an assembly task.

3. Methodology

3.1 Design

The objective of this experiment is to compare the differences between an AR instruction and a paper-based instruction in terms of *efficiency, effectiveness, cognitive load, motivation,* and *instruction experience*. A LEGO™ model assembly task is chosen. Studies showed that the LEGO™ model can be regarded as an abstraction for industrial assembly tasks, which has a high similarity with construction assembly (Tang et al., 2003; Sakata, Kurata, & Kuzuoka, 2006; Lei Hou, 2013; Funk, Mayer, & Schmidt, 2015; Funk et al., 2016) Besides, due to the small size of LEGO™ bricks, this assembly task can be a reasonable downscaled version, which is easy to control and duplicated. In this way, the distracting factors can be controlled.

The experimental design consists of *four* distinct phases:

1. Introduction session
2. Training session
3. Main experiment session
4. Evaluation session

This study was approved by the Ethics Committee of the Faculty of Behavioral, Management and Social Science at the University of Twente. The *independent variable* is the type of instruction: *paper-based instruction* or *mobile animated AR instruction*. The *dependent variables* are *effectiveness, efficiency, cognitive load, motivation,* and *instruction experience*. The introduction session is to introduce the experiment background. Participants need to sign a consent form (See Appendix III). The content of the consent form is to inform participants what data will be collected and how their information will be used. After signing the consent form, participants continue to the training session and perform a training task so as to get familiar with experiment equipment and materials. Later, the main experiment is executed to compare the differences between two types of instruction: a paper-based instruction and a mobile AR instruction.

3.2 Material

In this research, a car model from LEGO™ Creator 31055, online questionnaires and a tablet were used. In addition, two types of instruction were developed, namely, a paper-based instruction, and an animated mobile AR instruction. For each type of instruction, two versions were prepared, including a training version, and an experiment version. The training version is only used during the training session to introduce participants how to use materials and equipment, while the experiment version is used in the main experiment.

$$\begin{cases} \text{Paper Instruction} \begin{cases} \text{training version: 14 steps} \\ \text{experiment version: 30 steps} \end{cases} \\ \\ \text{AR Instruction} \begin{cases} \text{training version: 14 steps} \\ \text{experiment version: 30 steps} \end{cases} \end{cases}$$

LEGO™ Model

The LEGO model used in this experiment is LEGO Creator 31055 set (See Figure 2). This model was chosen for two reasons: 1) Since the model consists of diverse bricks with different shapes and color, the high complexity of this task made the user instruction highly needed for assembling. In other words, participants could not finish the task without an instruction. 2) The assembly procedure has 30 steps and the whole process takes about 10 minutes, which is a good range of time for this experiment.

Figure 2 Car model (From LEGO™ official website)

AR marker

In this experiment, an augmented reality marker was used. This marker is an official marker that is provided by Vuforia. It is rich in feature points, which makes tracking more sensitive and accurate. Appendix I shows the AR marker we used in this research. Users use camera to scan a marker first and then see augmented reality content that is overlaid on the marker.

Paper-based instruction

Since the size of the official manual is too small, the paper-based instruction was re-created by using a set of LEGO official instruction software which is free available. Firstly, LEGO™ Digital Designer[1] was used to design 3D models of bricks. LeoCAD[2] was then used to convert 3D models into 3D images. Finally, editing step timeline and delivering instruction content were achieved via LPub3D[3].

When re-creating the paper-based instruction, the same design guidelines used in LEGO™ official manuals were followed. To be more specific, the instruction shows a picture of component that needs to be picked and step number in the upper left corner. Furthermore, the instruction shows the bricks' assembly position in the middle of page. The background is also the same as in official instruction (See Figure 3). In order to avoid the influence of content size, the same tablet that is used to display AR instruction was chose to display the paper-based instruction. Participants can only navigate this instruction by touching and swiping the screen. There is no animation and interaction in this condition. Participants are trained firstly to use a training version instruction that consists of 14 steps to assemble a different model, and then an experiment version instruction when they are ready for the main experiment. In the experiment version, participants preview a complete car model that needs to be assembled so as to clarify what they need to achieve. Then, they see single steps on each page.

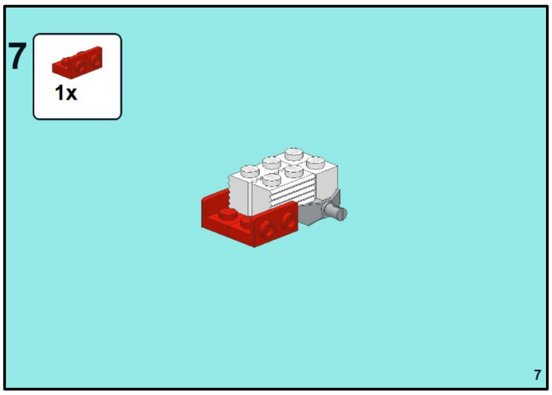


Figure 3 Step 7 of paper-based instruction (experiment version)

[1] LEGO Digital Designer: https://www.lego.com/en-us/ldd
[2] LeoCAD: https://www.leocad.org/
[3] LPub3D: https://sourceforge.net/projects/lpub3d/

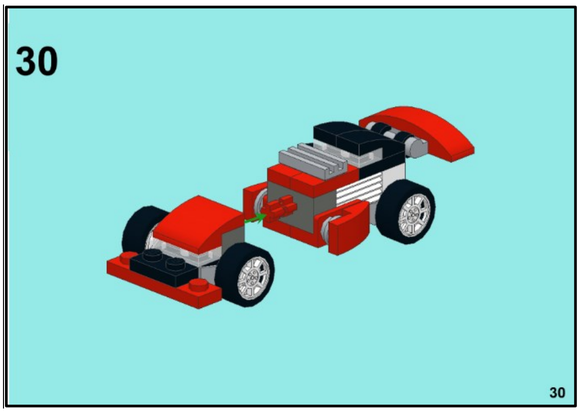


Figure 4 Step 30 of Paper-based instruction (experiment version)

Animated AR instruction

The animated AR instruction was created as an android application by using Vuforia and Unity. Vuforia is an Augmented Reality library that allows developers to make AR applications for diverse platforms. Unity is a game engine, which is typically used to develop both 2D and 3D content. 3ds Max, a 3D modelling and rendering software, was used to model 3D model of bricks. These pieces were imported to Unity as a prefab and then programmed in Unity to mimic the steps used in the paper-based instruction. Each piece was attached with animation that dynamically demonstrates the assembly process.

Considering most participants may be unfamiliar with AR technology, two AR manuals were prepared, including a training version (LegoTrial) and an experiment version (LegoAR). Both two instructions were designed in the same way and downloaded on a tablet by researchers before the experiment. The training version allows participants to scan a marker (See Appendix I) and then see a model which is different from the model used in the main experiment. In order to reduce the learning curve effect of tools and AR technology, subjects are trained to interact with AR technology as often as they want. After they get familiar with the operation of AR instruction, they begin to use the experiment version.

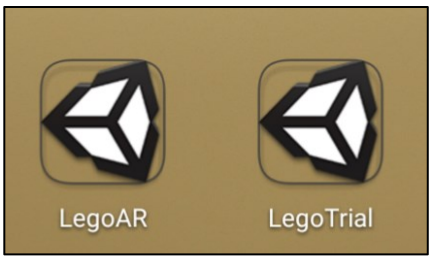


Figure 5 Two versions of AR instruction

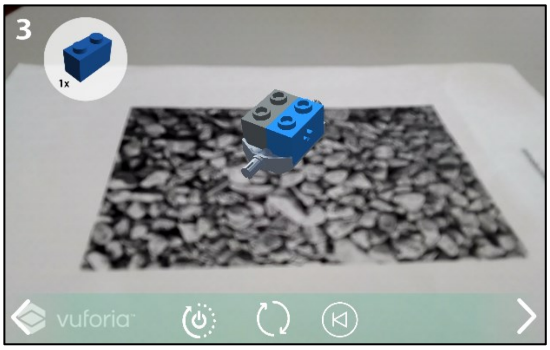


Figure 6 Step 3 of training version AR instruction

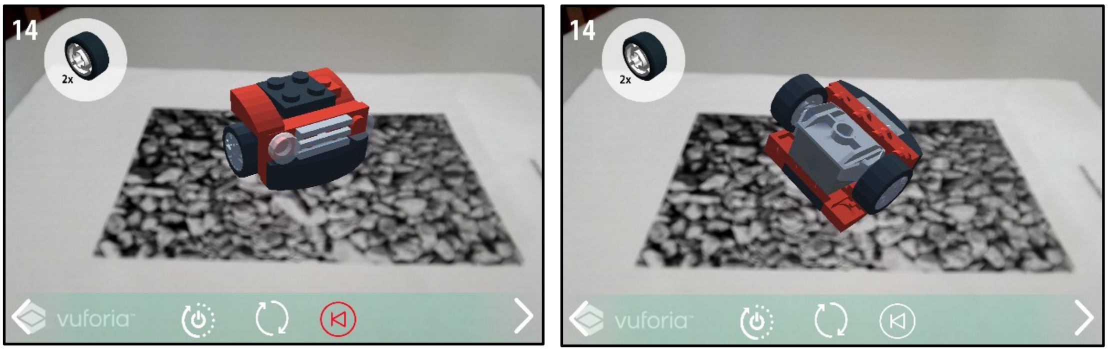


Figure 7 Step14 of training version AR instruction (left: before rotation; right: after rotation)

The experiment version AR instruction consists of two pages. In the first page (See Figure 8), participants scan a marker (See Appendix I) and see a complete car model that needs to be assembled. This car model allows participants to preview what they need to assemble during the experiment and clarify the goal of the task. Participants can click buttons to control the rotation and stop of augmented content. They can also click button "GO" to enter the second page if they are ready for starting the main experiment. After entering the second page, participants see the procedure information of assembly task. The design guideline of official instruction was used when developing AR instruction interface. On the top left corner, there are a step number and an image of the to-be-assembled component. In the middle screen, it shows animated step information. Participants can click Next, Back, Rotate, and Stop to control the 3D content. In addition, this application allows users to rotate and scale the 3D model by touching screen. If users want to replay the animation of one procedure that has been finished, they can press Reply button. Table 1 shows the function and icon of each button.

Table 1

Button icon and function

				
Back	Next	Reply	Rotate	Stop

Besides, sound effects were also included in the AR instruction. Users hear sound effects when animation shows a brick is assembled at the right position. For instance, users hear sound effects twice if they need to assemble two bricks within one step.

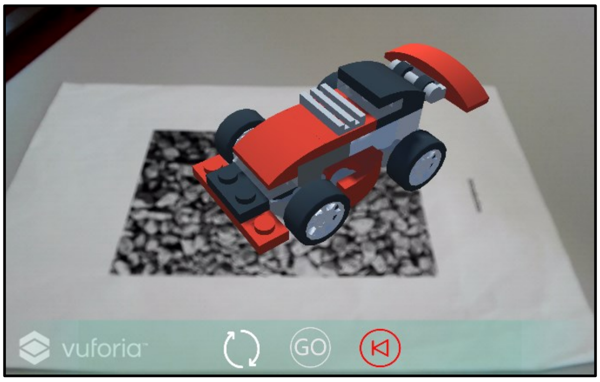


Figure 8 Preview of experiment version AR instruction

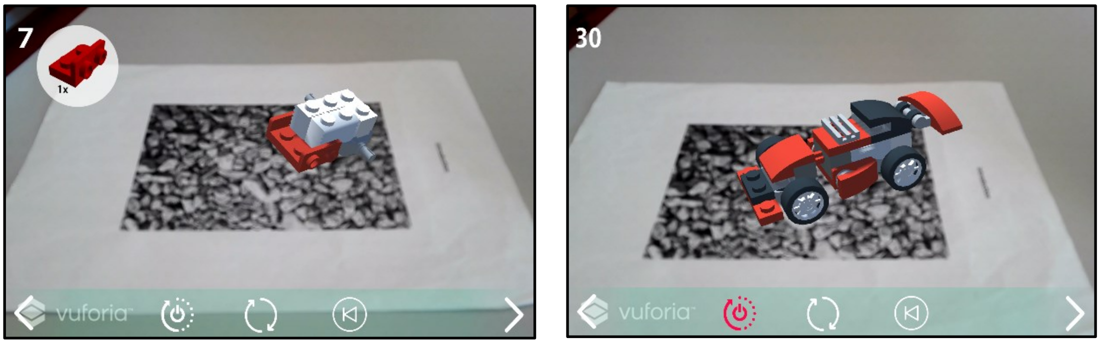


Figure 9 steps of experiment version AR instruction

3.3 Measures

Efficiency: Efficiency in this experiment refers to the total time to complete all 30 steps. Only data of participants who finished the task were recorded and analysed. A stopwatch and a camera was used to assess the total time taken to finish all 30 procedures.

Effectiveness: Effectiveness is measured by errors during the assembly process. In this experiment, we defined three types of errors according to Hou (2013) : (1) a component with wrong colour or wrong shape is selected; (2) a component is installed at the wrong location or with wrong orientation; (3) a step is skipped. If participants recognize the mistake during assembly and fix the error, the original errors are still regarded as errors instead of success. In order to record errors accurately, we composed an Error Recording Table (See Appendix X) for this experiment which contains procedure images and three types of errors. During the observation, researchers can tick (√) the box next to each type of error. Finally, by counting the number of ticks, the total errors were calculated. We hope this table is also helpful for following researchers.

Cognitive load: Considering the complexity of measuring equipment and technical limitation, we did not use psycho-physiological measures to measure cognitive load. Instead, the NASA Task Load Index (Hart, 1986) was used in this research. This questionnaire is not only inexpensive but also provides decent and reliable measurement, which has been cited in over 4400 studies. In NASA Task Load Index, the cognitive load is divided into six categories, including *mental demand, physical demand, temporal demand, effort, performance,* and *frustration* level. The definition of each category and description are listed in Appendix V. Subjects need to read a short instruction (Appendix V) before rating, which enables subjects to answer accurately. Each category is rated within a 100-points range with 5-points steps. Firstly, subjects rate each category according to their experience. Then, they need to perform a pair-wise comparison, pointing out which category contributes more to the workload of that task. After getting the rate and weight of each category, the sum of rating is calculated by multiplying each rate by its weight. The sum of the weighted ratings is divided by 15 to get the final cognitive load value. Taking into account the complexity of data calculation, an online version of NASA Task Load Index (See Appendix V) was adopted, which facilitates automatic computation.

Motivation*:* A Reduced Instructional Materials Motivation Survey (Loorbach, Peters, Karreman, & Steehouder, 2015) is used to measure motivation (See Appendix VI). This survey consists of 12 items scaled from not true (1) to very true (5), measuring motivation from four aspects (*Attention, Relevance, Confidence,* and *Satisfaction*). Each aspects consists of three items. Here is an example of questions that aim to measure Attention: *The quality of the instruction helped to hold my attention.*

Instruction experience: instruction experience refers to the how people think of the instruction they use. This survey is self-designed by the researcher and it consists of 11 items four aspects: *ease of use, positive experience, negative experience,* and *behaviour intention* (See Appendix VII). Here is an example of questions that aim to measure *Ease of use*: *This instruction is easy to use*.

3.4 Participants

72 participants were recruited in this study. All of them had never played with the car model used in this experiment. All participants come from over 27 countries. Most of them are students from University of Twente and some come from Saxion University of Applied Science.

All participants were divided into 2 groups, that is, a paper-based group and an AR group. Considering the gender and education background may influence the result, these two variables were controlled when assigning participants. As Table 2 shown, each group has 36 participants respectively, including 18 females and 18 males, 12 participants with non-technical study background and 24 participants with technical study background. The average age of the paper group is 24.44, while the average age of the AR group is 25.28 (See Table 3).

An independent-samples t-test was conducted to compare the age in the paper group and the AR group. There was no significant difference in age between two groups, t (70) = -.84; p=.40. This result suggested that two groups are comparable.

Table 2
Participants Distribution

	Paper Group	AR Group	Total
Female	18	18	36
Male	18	18	36
Non-technical	12	12	24
Technical	24	24	48

Table 3
Means (with standard deviations) of Age

Group	Paper Group	AR Group	Sig.(2-tailed)
Mean Age (Std.Deviation)	24.44(3.31)	25.28(4.92)	.40

3.5 Procedure

The total time of experiment is about 30 minutes. The whole experiment is divided into ***four*** stages. The flowchart (Figure 10) shows the workflow of this experiment.

Introduction session: At the beginning of experiment, a researcher gave a short introduction of the experiment (See Appendix II). Then participants were required to sign a consent form (See Appendix III) and fill in a demographic questionnaire (See Appendix IV) to collect their background information.

Training session: After the introduction session, the training session started. The researcher gave a short explanations about instructions assigned. Participants conducted a simple task that consists of 14 steps to familiarize themselves with the equipment. They started the experiment when they were acquainted with necessary operations. This session was particularly important for subjects from the AR group, since most participants have never used AR instructions before. The training task could be repeated until participants were ready to proceed.

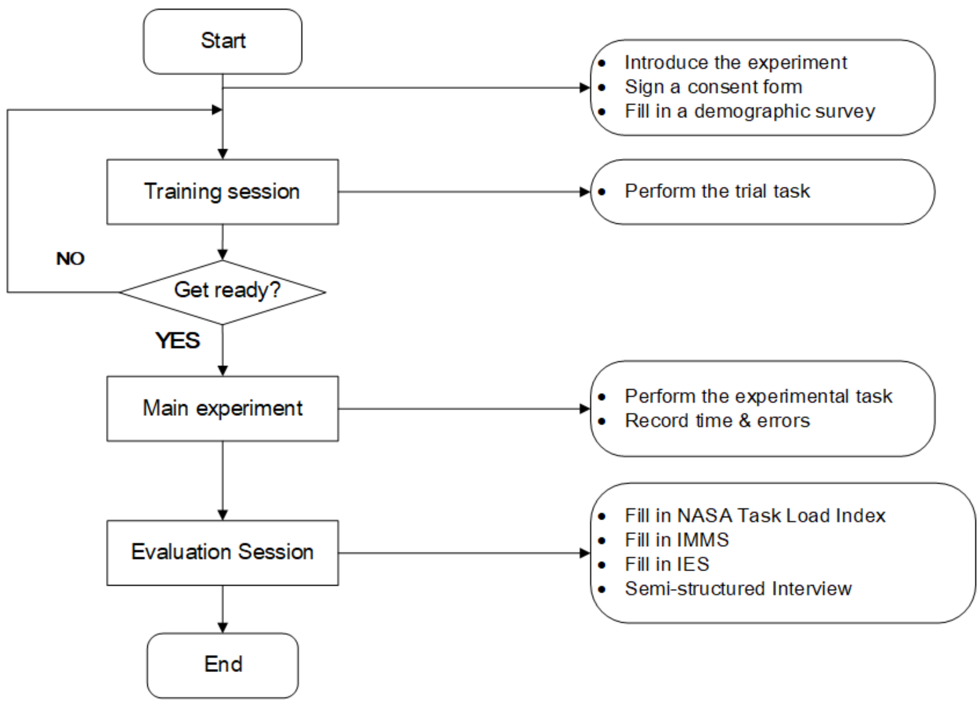


Figure 10 Experiment procedure

Main experiment session: The third stage is the main experiment. Various assembly components were randomly placed on the surface of the workplace. Participants from two groups needed to assemble all parts to form a LEGO^{TM} car model by following assigned instructions. When the subject was ready for the experiment, she/he would say "begin" to tell the researcher to start recording. A stopwatch recorded the time taken to assemble the LEGO^{TM} car model. Errors were counted by using Error Recording Table (See Appendix X). In order to guarantee the accuracy of time and error, a camera also worked together to record the whole assembly process of each participant.

Evaluation session: In the last stage, participants were provided with three questionnaires, including NASA Task Load Index, Reduced Instructional Materials Motivation Survey (RIMMS), and Instruction Experience Survey (IES) to measure their *perceived cognitive workload*, *motivation,* and *instruction experience*. Finally, a short semi-structured interview that consists of 3 questions was conducted to get additional qualitative data. Here is an example of interview questions: *How do you feel after using this instruction? Please describe your experience.* The whole interview process was voice recorded. The interview time was less than 3 minutes (See Appendix VIII).

During the experiment, participants were allowed to ask questions that were irrelevant to assembly task. After finishing data collection, the video tape of each participant was reviewed to verify assembly time and errors. All quantitative data were analysed by using SPSS. Qualitative data from the interview were summarized together with researcher's observation.

3.6 Pre-test

In order to make sure the procedure and instructions of the experiment worked smooth, a pre-test was conducted before the formal experiment. Six participants were invited to join this pre-test, including three females and three males. They were divided into two groups randomly. One group was paper group, another group was AR group. The pre-test was identical to the formal experiment that consists of four stages. During the pre-test, all participants finished the task and evaluation successfully. All surveys and equipment worked smooth and properly, which means the formal experiment could follow the procedure.

3.7 Data Analysis

After collecting all the data, validity and reliability test were conducted to make sure the quality of the data. When doing validity and reliability test, all data were pre-processed in Excel and then imported to SPSS. Three kinds of surveys used in this research were tested, namely, Reduced Instructional Material Motivation Survey (RIMMS), NASA Task Load Index (NASA-TLX), and Instruction Experience Survey (IES).

In terms of Reduced Instructional Material Motivation Survey (RIMMS), it consists of four subscales Attention, Relevance, Confidence, and Satisfaction. The reliability test shows the Cronbach's alpha scores for each subscale: *Attention* α=0.76, *Relevance* α=0.66, *Confidence* α=0.82, *Satisfaction* α=0.77. Since all scores are over 0.6, the data can be further processed. As for the NASA Task Load Index (NASA-TLX), the reliability test shows that Cronbach's alpha score is 0.70, which is acceptable for research.

As for the Instruction Experience Survey (IES), factors analysis and reliability test were both conducted to make sure the validity and reliability of such a self-designed survey. The score of KMO and Bartlette's test is 0.75, which means the data is suitable for factor analysis. After conducting the factor analysis, it is found that this survey contained four constructs, namely, *ease of use, positive experience, negative experience,* and *behavior intention*. The result is showed in Appendix IX. Question 1 was deleted since it belonged to two constructs and both correlation values were less than 0.5, which is trivially small. The final version of each construct are listed in Table 4. Finally, a reliability test of each subscale was conducted. The scores for each subscale are: *Ease of use* α=0.73, *Positive Experience* α=0.81, *Negative Experience* α=0.72, *Behavior Intention* α=0.75. Since all scores are over 0.6, which means this survey is acceptable and can be further processed.

Table 4 Instruction Experience Survey

This instruction…….
Ease of Use
Is easy to use.
Operation is clear and understandable
Positive Experience
Made me feel playful
Made me feel positively surprised
Made me feel joyful
Influenced my interest in the brand positively
Negative Experience
Made me feel dull
Made me feel unpleasant
Made me feel boring
Behavior Intention
Made me want to try more products of LEGO than I usually consider
Influenced my intention positively to purchase a LEGO product.

4. Results

4.1 Efficiency

Efficiency in this research refers to the *total time* to complete the LEGO^{TM} assembly task that consists of 30 steps. During the experiment, all participants finished the task. To test the hypothesis that a mobile animated AR instruction is able to reduce the completion time for an assembly task, an independent samples *t*-test was performed. As can be seen in Table 5, the *t*-test demonstrated no significant difference ($t\ (70)\ = -.368, p = .357$) in task efficiency. The mobile animated AR instruction did not reduce the completion time when compared with a paper instruction. Therefore, the first hypothesis is not confirmed.

Table 5
Results of Efficiency

	Mean (SD)			
	Paper	**AR**	**t**	**Sig.**
Efficiency	5'13"(2'22")	5'24"(1'36")	-.368	.357

Note: Efficiency refers to the total time to complete the task. The format of time in this table means Minutes'Second''(MM:SS).

4.2 Effectiveness

It was expected that participants from the AR group make fewer errors than participants of the paper group. Overall, the Independent Sample *t*-test indicated a significant difference between the instructions regarding the *Total errors*, $t\ (70) = 2.256, p = .015, d = 0.532$. Likewise, significant differences did exist for the *Selection Error*, $t\ (70)\ = 2.228, p = .015, d = 0.525$ and *Skip Error*, $t\ (70)\ = 1.781, p = .042, d = 0.420$. However, the *t*-test demonstrated no significant difference in the number of *Installation errors,* $t\ (70)\ = 1.482, p = .072$.

The results confirmed the second hypothesis. As expected, participants using the AR instruction made fewer *Total errors*, *Selection errors*, and *Skip errors* compared with participants using the paper instruction. A tendency towards a comparable effect was revealed for the *Installation error* (*p* <.10).

Table 6

Results of Effectiveness

	Mean (SD)				
	Paper	AR	t	Sig.	Cohen's d
Installation Error	0.72(1.19)	0.39(0.64)	1.482	.072	
Selection Error	0.36(0.59)	0.11(0.32)	2.228	**.015**	0.525
Skip Error	0.08(0.28)	0.00(0.00)	1.784	**.042**	0.420
Total error	1.17(1.63)	0.5(0.70)	2.256	**.015**	0.532

Note: Installation error is counted when a component is installed at a wrong location or a wrong orientation. *Selection Error* is counted when a component with wrong color or wrong shape is selected. *Skip Error* is counted when participants skipped a step. *Total Error* refers to the sum of three types of errors.

4.3 Motivation

Motivation was evaluated by RIMMS that consists of four aspects: *Attention, Relevance, Confidence,* and *Satisfaction*. In order to find out how different types of instruction influence motivation, the total means of motivation as well as scores on its subscales were shown in Table 7.

Table 7

Results of Motivation

	Mean (SD)			
	Paper	AR	t	Sig.(2-tailed)
Attention	3.98(0.89)	4.30(0.63)	-1.726	.089
Relevance	4.11(0.82)	4.10(0.82)	.048	.962
Confidence	4.68(0.62)	4.75(0.41)	-.598	.551
Satisfaction	4.13(0.86)	4.36(0.61)	-1.319	.191
Motivation	4.23(0.67)	4.38(0.43)	-1.153	.253

Note: Scores were measured by a five-point scale (1=not true, 2=slightly true, 3=moderately true, 4=mostly true, 5= very true)

Overall, the results showed that all participants experienced high motivation, since the means of *motivation* for the paper group and the AR group are both higher than 4. Independent sample *t*-test showed that the total motivation score did not differ between two groups ($t\ (70)\ =-1.153, p=.253$), nor did scores on its subscales *Relevance* ($t\ (70)\ =.048, p=.962$), *Confidence* ($t\ (70)\ =-.598, p=.551$), and

Satisfaction ($t\ (70)\ =\ -1.319, p = .191$). The scores on the subscale *Attention* tended to differ ($t\ (70)\ =\ -1.726, p < .10$).

According to the result, all participants showed high motivation when using either paper instructions or AR instructions. By comparison, AR instructions tended to increase attention, although such a difference was not significant.

4.4 Cognitive load

Cognitive load was measured by the NASA Task Load Index. In the NASA Task Load Index, the cognitive load is divided into six categories, including *mental demand, physical demand, temporal demand, effort, performance,* and *frustration* level. According to the usage instruction of NASA Task Load Index, a high rating indicates a negative trend.

Table 8

Results of Cognitive Load

	Mean (SD)				
	Paper	**AR**	**t**	**Sig. (2-tailed)**	**Cohen's d**
Mental Demand	4.35(3.58)	7.33(5.86)	-2.602	**.012**	-0.613
Physical Demand	1.58(2.48)	3.08(4.90)	-1.637	.108	
Temporal Demand	4.03(4.79)	4.70(5.40)	-.556	.580	
Performance	4.05(4.53)	2.19(2.42)	2.182	**.034**	0.514
Effort	4.17(4.09)	4.98(3.93)	-.854	.396	
Frustration	1.65(3.91)	0.98(3.08)	.815	.418	
Total Cognitive Load	19.84(12.40)	23.25(12.26)	-1.173	.245	

Note: Total cognitive load is the sum of six categories. Appendix V shows the definitions of each categories and calculation method.

Table 8 elaborated the means of total cognitive load and its six subscales. Overall, the total cognitive load did not show a significant difference, $t(70) = -1.173, p = .245$. Likewise, the scores of four subscales: *physical demand*($t(70) = -1.637, p = .108$), *temporal demand* ($t\ (70)\ = -1.173, p = .245$), *effort* ($t\ (70)\ = -1.173, p = .245$) and *frustration* ($t\ (70)\ = -1.173, p = .245$) did not show any differences between two groups. However, statistical differences did exist for the other two subscales:

mental demand ($t\ (70) = \ -2.602, p = .012, d = -0.613$) and *performance* ($t(70) = 2.184, p = .034, d = 0.514$).

As results shown, participants using paper instructions experienced less *mental demand* compared with participants using AR instructions. However, paper instructions caused a higher score on *performance*. The higher the score, the more poorly the subjects thought they had performed. This also means that the AR instruction had an advantage over paper instruction in making participants feel more successful. When it comes to the overall cognitive load and the other four subscales, two types of instruction showed no differences.

4.5 Instruction experience

For the evaluation of instruction experience, a self-designed questionnaire was used to collect the perception of participants towards two types of instruction. Based on reliability analysis and factor analysis, there are four constructs in this survey. They are *ease of use, positive experience, negative experience,* and *behaviour intention*. The results are shown in Table 9.

Table 9
Results of Instruction Experience

	Mean (SD)		t	Sig. (2-tailed)	Cohen's d
	Paper	**AR**			
Ease of Use	4.60(0.56)	4.58(0.53)	.108	.914	
Positive Experience	3.85(0.72)	4.15(0.44)	-2.111	**.039**	-0.498
Negative Experience	1.69(0.75)	1.44(0.42)	-1.742	.087	
Behavior Intention	3.67(0.73)	3.47(1.00)	.090	.348	

Note: Scores were measured by a five-point scale (1=strongly disagree, 2=disagree, 3=neutral, 4=agree, 5= strongly agree). *Ease of use* measured whether instruction is easy to use. *Positive experience* measured positive emotions, such as playful and joyful. Negative experience measured negative emotions, such as boring and unpleasant. Behavior intention measured users' product intention and purchase intention.

When it comes to the *ease of use*, no significant differences were found ($t(70) = .108, p = .914$). Considering the *positive experience*, a significant difference did exist in *positive experience* ($t(70) = -2.111, p = .039, d = -0.498$). The *negative experience* tended to differ ($t(70) = -1.742, p =< .10$). In terms of *behaviour* *intention*, there was no statistical difference between the paper condition and the AR condition ($t(70) = .090, p = .348$).

The results demonstrated that the AR instruction and paper instruction were both of good quality. However, AR instructions evoked more *positive experience* such as playful, surprised, and joyful emotions when compared with paper instruction.

4.6 Interview and Observation

Besides the quantitative results, qualitative results in form of interview and observation were also collected.

Paper group
Considering the paper group, almost all participants gave positive feedback after using the paper instruction. They pointed out that the paper instruction is very clear, easy to understand and straightforward. For instance, P8 said, "*it shows all the information, you don't have to analyse and think a lot*." Another participant responded as "*Even though they are just pictures, they are enough to help me. They tell you what you need to do and what pieces need to go where. I don't think I would be able to assemble the car without this instruction*. Most participants also mentioned that the images of the paper instruction are consistent and standard, which looks really professional. It helped them to finish the task in a logical and organized way.

However, some participants still gave suggestions and hoped to improve this paper instruction. For instance, 8 out of 36 participants mentioned that it was not easy to locate the position of a new brick when switching to a new page sometimes. They want animations to show where it comes from and where to install. One of the instances said: "*if there are animations for some tricky steps, it would be great. But we don't want animations for all steps*." 24 out of 36 participants indicated that when the model became more complex, they hoped to see the model from different angles so that they could get extra information. One participant even said "*with more interaction such as rotation and animation, I will feel more motivated. It gives you a lot of power to the person who uses the instruction. The instruction I use now is quite boring and plain to me*."

In a word, nearly all participants thought that the paper instruction in this research was of high quality. It demonstrated assembly information in a concise, clear and accurate way, making participants feel mentally effortless and relaxed. Some participants, on the other hand, gave their expectations for the future improvement of paper instruction. When the model became more complicated, they hoped that there is an option to rotate and animate the model.

AR group

Similarly, participants also gave positive responses to the AR instruction. In these responses, all participants thought the AR instruction used in this research was clear and intuitive. In addition, they also identified that such AR instruction was more attractive and interesting than paper instructions. 17 out of 36 participants pointed that 3D visualization of this AR instruction was more attractive than traditionally 2D visualization. It made them feel motivated and stick to the assembly task. For example, one of the participants mentioned "*I don't feel boring. It's quite different from the traditional instructions I used before. This one is just like playing a game, which is new and joyful. It makes me feel refreshed and motivated*". In addition, participants noticed that 3D visualization was vivid and realistic, which helped them understand the model clearly. For instance, three participants said:

- *It's really close to what you really have in your hand. The 3D interactive model makes my experience more realistic and engaging.*

- *I think the 3D model is very clear. It's also really realistic. I like the 3D model since I can make more confirmation.*

- *I think the AR instruction is really cool and visual. You actually see what you are going to do. The 3D model is identical to the real model.*

Considering the attitudes towards interaction controls such as animation, rotation and sound effects of this AR instruction, 28 out of 36 participants pointed that they liked the rotation function. For instance, one participant said: "*I like the rotation. You can choose the perspective that you want to see. For a video, you cannot adjust it according to your preference.*" Likewise, another participant commented: "*The rotation is useful since I feel really frustrated when I cannot rotate pieces when using a paper instruction. For this AR manual, you can see the model from every angle. You don't have to mentally rotate the model*". In addition, 16 out of 36 participants identified that animation function was helpful for them. Most of them said when the model became more complex, the animation showed the assembly path so that they could understand where bricks came from and how two bricks were connected with each other. Only 2 participants noticed the sound effect used in this AR instruction: "*I* *can know how many pieces I need to assemble in this step when I am listening to the background sound*". However, most participants said sound effect did not help them a lot. Some participants even criticized that the sound effect was noisy and desynchronized with the movement of bricks.

Apart from those positive comments mentioned before, some participant also gave useful suggestions to improve the AR instruction. 8 participants suggested that the rotation function should be more intelligent and less sensitive. For instance, one participant said: "*When showing a new step for the first time, I hope the model can reset to the best angle automatically so that I don't need to adjust it by myself*". Another participant also pointed out that "*Currently, the screen touch is too sensitive, which makes it hard for me to control the model when I want to rotate it*". When it comes to the animation function, some participants thought the speed of the animation is not proper for them. Sometimes animations were a little bit fast, they had to replay animation once again. Sometimes animations were slow, they had to wait for the animation to be done before they went to the next step. This tended to increase the assembly time and hinder the task efficiency. Participants suggested that we should add an option that allows them to adjust animation speed according to users' preference.

Overall, the perceptions of participants regarding the AR instruction were encouraging and positive. They showed an interest in using this new type of instruction. They were curious and motivated about their LEGO assembling experience using an untraditional AR instruction. However, there are a few issues that were highlighted by participants. In general, participants enjoyed such a novel instruction to guide them to perform an assembly task and most of them were positively surprised by AR technology. Some participants also suggested that such interactive instruction would be particularly useful when the assembly task became more complex so that the AR instructions would benefit them more.

5. Discussion and Implications

This study evaluated the usability of a mobile AR instruction and a paper instruction in guiding a LEGO assembly task. Two hypotheses and two questions were proposed. In this section, the main findings are discussed and explained, followed by the theoretical implication and practical implication of these findings.

5.1 Main findings

Efficiency

The first hypothesis is that a mobile animated AR instruction will result in an increase of *efficiency* when compared with a paper instruction. Contrary to the hypothesis, results suggest that there is no statistical difference between the paper instruction and the AR instruction on *efficiency*, which is inconsistent to the study by Baldassi et al. (2016).

A potential explanation for the lack of differences in efficiency is the speed issue of animations in the AR instruction. When using the AR instruction, participants cannot adjust the animation speed according to their own preference. Some participants who assembled quickly complained that they could finish the task faster without animations since sometimes animations seemed to be slower than their expectation, decreasing their task performance. Other participants, however, said the speed of animations was a little bit fast. They had to replay animations to verify them once again.

This finding can be explained by previous studies of Tekušová & Kohlhammer (2008) and Griffin et al. (2006). Their studies showed that it is critical to take speed into account so that participants can identify changes well in the animation. An improper animation speed could hinder the task efficiency and make participants feel impatient. However, as Meyer, Rasch, & Schnotz (2010) and Kriglstein, Pohl, & Stachl (2012) pointed out, finding out an optimal presentation for dynamic visualizations is a difficult issue and a huge challenge. There is no experiment that evaluates the best speed of animation in AR yet. It tends to be beneficial that when participants can adjust the speed of animation, which is in line with the expectation of some participants who use AR instruction in this research. In addition, designing the animation with customized speed seems to be another solution. For example, the animations can diagnose the complexity of tasks and adjust the speed automatically.

Effectiveness

The result shows that the AR instruction does affect people's task *effectiveness*. This result confirmed the second hypothesis. Participants using AR instructions performed more task correctly than participants using paper instructions. In other words, the AR instruction has a positive influence on task *effectiveness*.

It is not surprising that AR instructions reduce errors. There are three reasons why this result existed. Firstly, the 3D visualization of AR instruction is more realistic than a paper instruction, making it easier to distinguish the components. Although components in 2D drawings can be expressed in images or lines, it is still difficult to represent more details (Perdomo, Shiratuddin, Thabet, & Ananth, 2005). Compared with 2D drawings, 3D visualization not only shows colours and shapes but also the rendering that represents the various materials of components, which gives a "true feel" and a detailed representation of the model. Such virtual model is really close to what participants have in their hands, makings their experience more realistic. Therefore, it is easier for users to fully understand and distinguish what kind of brick they need and where it should be installed. With such a realistic representation, it is hard to make errors. It can also be assumed that the significant differences in selection error are caused by this realistic representation.

Secondly, the interactive function of AR instructions allows users to adjust the model according to their needs. Participants said the *rotation* function was very useful especially when the model becomes more complex. For instance, participants can use the rotation function to manipulate the model to a unique position in space, which adds flexibility of viewing. They can adjust the 3D model around any axis or even zoom in/out in any direction. Through such interaction, more details information can be clarified and enhanced. They can also use rotation to figure out how bricks are interconnected with each other. This interaction is similar to what we normally do when we observe an object. We pick the object up, sometimes we tile the head to see another side, or hold it close to us to make it bigger and see details (Santos et al., 2014). Besides, the *animation* function shows the motion path and the whole assembly process, which enables participants to mimic each step and reduce the difficulty of operations. By contrast, the paper instruction that shows static and oversimplified images to present 3D objects is sometimes abstract and perceptually inefficient (Baldassi et al., 2016). Paper instructions only show the beginning and the final result, which is difficult for participants to know what happened in the middle process. This finding is in accordance with the previous literature. As Salomon(1979) indicated, animations facilitate information processing since they help learners mentally visualize a process, compensating for a learner's inability to imagine motions of a mental representation. Through an animated 3D model, the AR instruction can eliminate the abstraction of 2D images when representing 3D objects and actions.

Last but not least, the AR instruction tends to engage participants more when compared with a paper instruction. It was observed that during the experiment some participants interacted with the model very frequently, even though they had already selected correct bricks and installed them correctly. When asking the reason, participants said that they want to make more confirmation and make sure they do it in a right way. By contrast, participants who use the paper instruction tended to only glance at the instruction very quickly and performed the task. Most of them neglected details and made errors. Apart from more installation errors and selection errors in the paper group, 3 participants even overlooked some steps, leading to *Skip Errors*. In a word, it is assumed that more engagement of the instruction brings more correctness even though the number of skip errors in this research was not very high.

Cognitive load

When it comes to the effect of two instructions on cognitive load, no statistically significant differences were found in the AR version and paper version. Significant differences in subscale *mental demand* and *performance* were found. For the *mental demand*, the AR instruction was rated higher than the paper instruction. However, participants using paper instructions felt less successful than participants using AR instructions. In other words, although the AR instruction causes more mental demand, it gives participants more sense of accomplishment and satisfaction with their performance in achieving goals.

There are two potential reasons for the *higher mental demand* of AR instructions. The first reason could be attributed to the participants' being unfamiliar with AR instructions. It is found that 70 participants (totally 72 participants) have never used such an AR instruction before. During the experiment, many participants invested much effort to learn the operation of AR instructions at the beginning. Although they knew how to click buttons and where to click, most of them could not operate the instruction very smoothly. By contrast, all participants were quite familiar with the paper instruction, since they often use this kind of instruction in daily life. Therefore, the learning curve of AR instructions may contribute to the higher mental demand.

The second reason could be the improper speed of animation and too sensitive screen-touch function. This is the same reason as *efficiency*. Although most participants said animation helped them a lot, there is still a risk of such dynamic information presentation. Hegarty (2004) pointed out that providing non-stable and transient information may impose a higher mental load on learners. The information of changes is demonstrated only for a short time and has to be kept in working memory to integrate with a mental model, which leads to a higher mental load. Like some participants complained, the speed of animation and automatic rotation was too fast for them, making them feel confused or even a little bit dizzy. In addition, the screen-touch rotation function tends to be too sensitive. It was observed that when participants touched the screen slightly, the virtual model changed the viewing a lot. As a result, participants tried hard to rotate the model several times so as to adjust it to the best angle. Participants expected that the speed of animation can be adjusted according to their preference and the screen-touch function can be less sensitive.

However, participants using AR instructions experienced more success and achievement about their performance. According to the results of effectiveness and efficiency, participants using AR instructions finished the task with fewer errors and similar time when compared with participants using paper instructions. As a whole, participants from the AR group performed better than participants from the paper group. Therefore, it is not surprising that they felt more successful since they actually did a better job. This result also verified the findings of efficiency and effectiveness.

It is important to mention that some participants felt that the description of NASA-TLX was ambiguous to them even though statistic result did not show such confusion. During the evaluation session, participants pointed out that the definition of subscale *effort* and the definition of subscale *mental demand* tended to be overlapped with each other, which made them feel confused and vague when evaluating these subscales and giving weight points. Although the researcher gave explanations when participants felt confused, the definition of NASA-TLX that tends to be ambiguous may still impact the results.

Motivation

Considering the effect of the two instructions on *motivation*, there is no significant difference in motivation even though the scores of subscale *Attention* tended to differ. There are three reasons why participants experienced high motivation and no significant differences existed when using two types of instruction.

Firstly, the LEGO assembly task used in this research is full of fun and motivation, which arouses the intrinsic motivation of participants. According to qualitative results, most participants said they enjoyed playing with LEGO since it was a really interesting and creative toy. They had large interests in LEGO so they joined the experiment. Previous literature shows that intrinsic motivation is closely related to intrinsic value, which refers to subjective interest or enjoyment of performing a particular task (Lai, 2011). Thus, it is assumed that such intrinsic motivation and attraction aroused by LEGO made participants in both two groups feel self-motivated. As a result, the motivation evoked by AR technology may not add too much value. However, assembly tasks from other fields may be different, which tends to be more boring and less attractive when compared with LEGO tasks. Therefore, more types of tasks from other assembling context should be evaluated in the future study to find out the effect of an AR instruction and see whether it can increase users' motivation.

Another reason for high motivation may be the good quality of two instructions. According to participants' feedback, two instructions were evaluated as clear and easy to understand, which made the assembly task simple and understandable. Due to good-quality instructions, participants did not encounter lots of confusion and failures which decrease their motivation.

The last possible reason is that the complexity of the task is not high enough. During the experiment, it was observed that all participants finished the task even though some of them met difficulties during assembly. Based on the result of interview, many participants mentioned that the assembly task was less difficult than they expected. Therefore, it is expected that tasks with different complexities will be investigated in the future study.

Instruction experience

Results indicated that AR instructions aroused more positive experience when compared with paper instructions. Specifically, participants using AR instructions felt more playful, joyful, surprised, and positively interested in the LEGO brand. However, in terms of other aspects such as *ease of use, negative experience,* and *behaviour intention*, two types of instructions showed no statistical differences.

Although there are no significant differences in *negative experience*, a tendency is found in this measurement. Participants using paper instructions tend to experience more negative emotions such as boring, dull and unpleasant when compared with participants using AR instructions. It could be assumed that traditional paper instructions tend to be boring and less attractive.

Considering the *ease of use*, it is observed that both two types of instruction are evaluated with high quality. Participants from the paper group pointed out that the paper instruction was very clear, and straightforward, while participants from the AR group said the AR instruction was intuitive and easy to understand. According to participants' positive responses, it is reasonable to conclude that both instructions are easy to use.

In terms of *behaviour intention*, no significant differences were found. It is really hard to say a good instruction can influence users' behaviour intention such as purchase intention and instruction intention because that depends on other factors. More influential factors should be taken into consideration when measuring the influence of instructions on behaviour intention.

It is not surprising that AR instructions bring more *positive experience.* Throughout the observation and the following interview, participants expressed a wide range of positive emotions regarding their experience with AR instructions. Nearly all participants have never used AR instructions before. It is no wonder that participants felt surprised with this fancy and novel instruction. In addition, the interactive features of AR instruction bring more engagement and entertainment to participants. In a word, the experience of joy, playfulness, and surprise dominated participants' AR instruction experience. These findings are in line with the results of other researchers. For example, Cehovin & Ruban (2017) found that the interaction and augmentation of AR were full of playfulness and joy. In a similar vein, Huang & Tseng (2015) suggested that AR with interactive experiences caused users to show a high degree of playfulness and positive experience.

It is necessary to assume that the feeling of surprise may be evoked by the novelty of AR. This means such emotion may appear strongly when participants encounter AR for the first time but fade in strength with longer the participants use AR technology. As Berlyne(1970) pointed out, once users get more used to AR, they may no longer have notable excited when interacting with AR applications, even though they exactly do so in the beginning. Users tend to regard such interactive technology as the most common one, exhibiting less strong surprise experience to AR instructions. Therefore, it is interesting to see how users perceive AR instruction in a long-term period.

Nonetheless, it is believed that the *higher positive experience* of AR instructions cannot be solely attributed to the novelty of AR. At least parts of the positive experience are accredited to the fact that AR has changed users' interaction with instructions. This new type of instruction allows users to view 3D models, see dynamic animations, and interact with content instead of only reading static 2D image and texts. It can be said that such new, novel, and visualized instruction boosts users' curiosity and interests, making reading instructions less boring and unattractive.

5.2 Theoretical implications

Altogether, the theoretical implications of this thesis are summarized as three aspects. This research contributes (1) to the understanding of the mobile AR instructions to guide an assembly task, (2) to scientific knowledge from new research dimensions with regard to the effect of AR instructions (3) to supplement the empirical evidence with regard to the effect of AR instructions on task performance and cognitive load.

First, this thesis provides empirical research about the effects of mobile AR instructions to guide an assembly task. To date, the literature lacks the studies that investigate mobile AR instructions. Since extant studies restrict their research to traditional AR instructions that use Head-Mounted Display equipment, this new type of mobile AR instructions has suffered from insufficient consideration. The research presented in this thesis contributes to the closure of this research gap. In more detail, a systematic research with five dependent variables was conducted to compare the effects of mobile AR instructions and paper instructions.

Second, previous studies have often only focused on traditional dimensions such as efficiency, effectiveness and etc. when evaluating AR instructions. Apart from including traditional evaluation dimensions, this research also considered other new dimensions to investigate the effect of mobile AR instructions and paper instructions. These new dimensions include *motivation* and *instruction experience*, which contributes to the closure of the research gap. This exploratory approach allowed the researcher to take a more systematic picture of comparing the effects of mobile AR instructions and paper instructions. In more detail, empirical evidence suggested that mobile AR instructions aroused more positive experience when compared with paper instructions. Two types of instructions showed no differences in motivation. Yet, to the best of my knowledge, the study presented here is the first to show how mobile AR instructions influence users' *motivation* and *instruction experience.*

Thirdly, this research offers empirical evidence for the effect of mobile AR instructions on *task performance* (efficiency and effectiveness), and *cognitive load*. As already mentioned, there is empirical evidence for the effect of traditional AR instructions that use Head-Mounted Display equipment. A number of scholars have already shown that AR instructions significantly improve task efficiency and effectiveness. This research considered a new type of mobile AR instruction rather than traditional AR instructions. It also provides empirical evidence to explain how mobile AR instructions influence task efficiency and effectiveness. This study found mobile AR instructions did not increase task efficiency, which rejects the findings from Baldassi et al. (2016) about efficiency. By contrast, it validated and underpinned the studies from Tang et al. (2003), since mobile AR instructions did increase task effectiveness. In addition to task performance, this study added to the argument regarding the influence of mobile AR instruction on *cognitive load*. The finding in this thesis provided empirical evidence that mobile AR instructions neither increase nor decrease *cognitive load* when compared with paper instructions.

5.3 Practical implications

The main findings of this research have clear practical implications for instruction designers. First of all, instruction designers can consider this new type of instructions as an alternative version or a complementary version for paper instructions, displaying user instructions in a dynamic way. As results shown, mobile AR instructions are indeed capable to improve task effectiveness and positive experience of users, and they do not hinder task efficiency, motivation, and cognitive load. Compared with traditional paper instructions, mobile AR instructions provide more vivid 3D visualization and more interaction controls. This type of instructions could be particularly useful in some tasks that emphasize more on accuracy instead of time. It is also advisable that AR instructions can be an addition for some tricky tasks instead of replacing paper instructions completely.

In addition, instruction designers should consider the design of controls of AR instructions, optimizing the interaction of instructions and users. As this study shown, rotation function and animation function are crucial to users especially when models become more and more complex. An improper animation speed or unintelligent rotation control could cause high mental demand. Therefore, a more intelligent and customized interaction should be considered when designing mobile AR instructions. For example, the animation can change its speed according to the complexity of the task. Also, designers can give more freedom to users so that users can set the interaction way by themselves, such as adjusting speed, stop/activate animations and etc.

Last but not least, participants in this research showed large interests and curiosity towards this new type of AR instructions. The managers of organizations and companies can try to adopt AR technology to update product instructions since this new digital instruction bring more positive experience and customers’ interests in brands and companies.

6. Limitations and Future Research

There are four limitations of this study.

First of all, although all participants have never played the LEGO model used in this experiment, three quarters of participants have played other LEGO models before. The experience with other LEGO models may influence their task performance. Due to the limitation of research time, it is hard to find more participants. Therefore, it would be useful to conduct the study with more participants who completely have no experience with LEGO model assembly. It is assumed that participants with no experience will get more benefit from AR instruction. When they get more experience, the benefit of AR instructions may decrease (Funk et al., 2017). In addition, participants in this study are all students. More participants with various background should be included in later research.

Secondly, although AR instruction is evaluated positively by all participants, it still has some technical issues due to the limited development time schedule. For instance, the screen-touch function is too sensitive, which tends to influence users' operation. Therefore, this function should be improved in the updated version.

Third, due to the laboratory limitation, the assembly task used in this study is a LEGO model. More studies need to be conducted in other assembly tasks that are from real construction context. It is hoped that this mobile AR instructions can be implemented into real construction projects to investigate the benefits of AR instructions.

Last but not least, this experiment only investigates the impact of AR instructions in a short-term period. In future research, it is expected that the effects of a mobile AR instruction in a long-term study with a runtime of several months should be explored. It is interesting to see whether the positive effects of such AR instructions can retain or decrease over a longer period of time.

There are other interesting research directions that are not addressed in this thesis. Firstly, it could be interesting to conduct an experiment to investigate static/dynamic mobile AR instructions and how animation speed affects users' cognitive load. Another research direction can be that whether gender and spatial ability influence users' cognitive load regarding AR instructions. Furthermore, the complexity of tasks may also influence the experience of AR instructions. Thus, further studies could investigate the effect of AR instructions on assembly tasks with different complexity levels. It is assumed that AR instructions may benefit more in a more complex task than a simple task.

7. Conclusion

Although the benefits of mobile Augmented Reality have been discussed by researchers from other fields such as education, tourism, and gaming, its usage in assembly instruction has not been investigated so far. The aim of this thesis is to compare the differences between a mobile AR instruction and a paper-based instruction when guiding an assembly task. Compared with paper instructions, AR instructions increase task *effectiveness* to complete an assembly task. However, AR instructions do not show an advantage in improving task *efficiency*. Considering *cognitive load* and *motivation*, AR instructions demonstrate similar influences as paper instructions. Considering *instruction experience*, AR instructions bring more *positive experience* such as playful, joyful, and surprised feelings, which facilitates participants to show interests in the brand.

Overall, the results of this research suggest that current mobile AR instructions are indeed capable of improving task performance and positive experience of users. It is advised that designers of assembly instructions can consider mobile AR instructions as an alternative version. At least, this study has shown that for people who have no prior experience with an assembly task, an AR instruction increases their task *effectiveness* and *positive experience* and it does not hinder task *efficiency*, *motivation*, and *cognitive load*. Therefore, it is meaningful to try such a new type of instruction.

8. Literature

Azuma, R. T. (1997). A survey of augmented reality. *Presence: Teleoperators and Virtual Environments*, *6*(4), 355–385.

Baldassi, S., Cheng, G. T., Chan, J., Tian, M., Christie, T., & Short, M. T. (2016). Exploring immersive AR instructions for procedural tasks: The role of depth, motion, and volumetric representations. *Adjunct Proceedings of the 2016 IEEE International Symposium on Mixed and Augmented Reality, ISMAR-Adjunct 2016*, 300–305. https://doi.org/10.1109/ISMAR-Adjunct.2016.0101

Barab, S., Thomas, M., Dodge, T., Carteaux, R., & Tuzun, H. (2005). Making learning fun: Quest Atlantis, a game without guns. *Educational Technology Research and Development*, *53*(1), 86–107.

Berlyne, D. E. (1970). Novelty, complexity, and hedonic value. *Perception & Psychophysics*, *8*(5), 279–286.

Berthomieu, B., & Menasche, M. (1983). An enumerative approach for analyzing time Petri nets. *Proceedings IFIP*. https://doi.org/10.1007/3-540-45014-9

Blattgerste, J., Strenge, B., Renner, P., Pfeiffer, T., & Essig, K. (2017). Comparing conventional and augmented reality instructions for manual assembly tasks. *Proceedings of the 10th International Conference on Pervasive Technologies Related to Assistive Environments - PETRA '17*, 75–82. https://doi.org/10.1145/3056540.3056547

Cehovin, F., & Ruban, B. (2017). *The impact of augmented reality applications on consumer search and evaluation behavior.* (Master Thesis, Copenhagen Business School, Copenhagen, Denmark). Retrieved from http://studenttheses.cbs.dk/handle/10417/6270

Chickzentmihalyi, M. (1990). *The psychology of optimal experience*. New York, Harper Collins.

Dunleavy, M., Dede, C., & Mitchell, R. (2009). Affordances and limitations of immersive participatory augmented reality simulations for teaching and learning. *Journal of Science Education and Technology*, *18*(1), 7–22.

Efklides, A., Kuhl, J., & Sorrentino, R. M. (2001). *Trends and prospects in motivation research*. Springer.

El Sayed, N. A. M., Zayed, H. H., & Sharawy, M. I. (2011). ARSC: Augmented reality student card. *Computers & Education*, *56*(4), 1045–1061.

Funk, M., Bachler, A., Bachler, L., Kosch, T., Heidenreich, T., & Schmidt, A. (2017). Working with augmented reality? A long-term analysis of in-situ instructions at the assembly workplace. *ACM International Conference Proceeding Series*, *Part F1285*, 222–229. https://doi.org/10.1145/3056540.3056548

Funk, M., Kosch, T., & Schmidt, A. (2016). Interactive worker assistance: Comparing the effects of in-situ projection, head-mounted displays, tablet, and paper instructions. *Proceedings of the ACM International Joint Conference on*

Pervasive and Ubiquitous Computing, 0–5. https://doi.org/10.1145/2971648.2971706

Funk, M., Mayer, S., & Schmidt, A. (2015). Using in-situ projection to support cognitively impaired workers at the workplace. In *Proceedings of the 17th international ACM SIGACCESS conference on Computers & accessibility* (pp. 185–192). ACM.

Goodwin, D. (1991). Emplotting the reader: Motivation and technical documentation. *Journal of Technical Writing and Communication*, *21*(2), 99–115.

Griffin, A. L., MacEachren, A. M., Hardisty, F., Steiner, E., & Li, B. (2006). A comparison of animated maps with static small-multiple maps for visually identifying space-time clusters. *Annals of the Association of American Geographers*, *96*(4), 740–753.

Haider, H., & Frensch, P. A. (1996). The role of information reduction in skill acquisition. *Cognitive Psychology*, *30*(3), 304–337.

Haniff, D. J., & Baber, C. (2003). User evaluation of augmented reality systems. In *Information Visualization, 2003. IV 2003. Proceedings. Seventh International Conference on* (pp. 505–511). IEEE.

Hart, S. G. (1986). *Task Load Index*. Retrieved from https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20000021488.pdf

Herrema, J. S. (2013). *Evaluation of graphical user interfaces for augmented reality based manual assembly support*. (Doctoral Dissertation, Iowa State Univerisity, Iowa, United States). Retrieved from https://lib.dr.iastate.edu/etd/13269/

Hou, L. (2013). *Evaluating the use of augmented reality to facilitate assembly,*. (Doctoral Dissertation, Curtin Univeristy, Perth, Australia). Retrieved from https://espace.curtin.edu.au/handle/20.500.11937/2125

Hou, L., Wang, X., Bernold, L., & Love, P. E. D. (2013). Using animated augmented reality to cognitively guide assembly. *Journal of Computing in Civil Engineering*, *27*(5), 439–451. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000184

Huang, H.-M., Rauch, U., & Liaw, S.-S. (2010). Investigating learners' attitudes toward virtual reality learning environments: Based on a constructivist approach. *Computers & Education*, *55*(3), 1171–1182.

Huang, T.-L., & Tseng, C.-H. (2015). Using augmented reality to reinforce vivid memories and produce a digital interactive experience. *Journal of Electronic Commerce Research*, *16*(4), 307.

Huk, T., Steinke, M., & Floto, C. (2003). The influence of visual spatial ability on the attitude of users towards high-quality 3D-animations in hypermedia learning environments. In *E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education* (pp. 1038–1041). Association for the Advancement of Computing in Education (AACE).

Keller, J. M. (1979). Motivation and instructional design: A theoretical perspective. *Journal of Instructional Development*, *2*(4), 26–34.

Kim, M. J. (2013). A framework for context immersion in mobile augmented reality. *Automation in Construction*, *33*, 79–85.

Kriglstein, S., Pohl, M., & Stachl, C. (2012). Animation for time-oriented data: An overview of empirical research. *Proceedings of the International Conference on Information Visualisation*, 30–35. https://doi.org/10.1109/IV.2012.16

Lai, E. R. (2011). Motivation : A literature review research. *Research Reports*, (April), 43. https://doi.org/10.2307/3069464

Laperriere, L., & ElMaraghy, H. A. (1992). Planning of products assembly and disassembly. *CIRP Annals-Manufacturing Technology*, *41*(1), 5–9.

Lee, D. Y., & Shin, D.-H. (2012). An empirical evaluation of multi-media based learning of a procedural task. *Computers in Human Behavior*, *28*(3), 1072–1081.

Locke, E. A. (1968). Toward a theory of task motivation and incentives. *Organizational Behavior and Human Performance*, *3*(2), 157–189.

Loorbach, N., Karreman, J., & Steehouder, M. (2007). Adding motivational elements to an instruction manual for seniors : Effects on Usability and Motivation. *Technical Communication*, *54*(3), 343–358.

Loorbach, N., Peters, O., Karreman, J., & Steehouder, M. (2015). Validation of the instructional materials motivation survey (IMMS) in a self-directed instructional setting aimed at working with technology. *British Journal of Educational Technology*, *46*(1), 204–218. https://doi.org/10.1111/bjet.12138

Loorbach, N., & Taal, E. (2006). The effects of motivational elements in user instructions. *Journal of Business and Technical Communication*, *20*(2), 177–199. https://doi.org/10.1177/1050651905284404

Marner, M. R., Irlitti, A., & Thomas, B. H. (2013). Improving procedural task performance with augmented reality annotations. *2013 IEEE International Symposium on Mixed and Augmented Reality, ISMAR 2013*, 39–48. https://doi.org/10.1109/ISMAR.2013.6671762

McCalla, G. I., Greer, J. E., Kumar, V. S., Meagher, P., Collins, J. A., Tkatch, R., & Parkinson, B. (1997). A peer help system for workplace training. *Artificial Intelligence in Education: Knowledge and Media in Learning Systems*, 183–190.

Meyer, K., Rasch, T., & Schnotz, W. (2010). Effects of animation's speed of presentation on perceptual processing and learning. *Learning and Instruction*, *20*(2), 136–145. https://doi.org/10.1016/j.learninstruc.2009.02.016

Milgram, P., & Kishino, F. (1994). A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Systems*, *77*(12), 1321–1329.

Neumann, U., & Majoros, A. (1998). Cognitive, performance, and systems issues for augmented reality applications in manufacturing and maintenance. *Proceedings. IEEE 1998 Virtual Reality Annual International Symposium (Cat. No.98CB36180)*, 4–11. https://doi.org/10.1109/VRAIS.1998.658416

Perdomo, J. L., Shiratuddin, M. F., Thabet, W., & Ananth, a. (2005). Interactive 3D visualization as a tool for construction education. *2005 6th International*

Conference on Information Technology Based Higher Education and Training, FB4-23-FB4-28. https://doi.org/10.1109/ITHET.2005.1560307

Pérez-López, D., Contero, M., & Alcaniz, M. (2010). Collaborative development of an augmented reality application for digestive and circulatory systems teaching. In *Advanced Learning Technologies (ICALT), 2010 IEEE 10th International Conference on* (pp. 173–175). IEEE.

Rost, M. (2006). Generating student motivation. *WorldView*.

Sakata12, N., Kurata12, T., & Kuzuoka, H. (2006). Visual assist with a laser pointer and wearable display for remote collaboration.

Salomon, G. (1979). *Interaction of media, cognition, and learning: An exploration of how symbolic forms cultivate mental skills and affect knowledge acquisition.* San Francisco: Jossey-Bass Publishes.

Santos, M. E. C., Chen, A., Taketomi, T., Yamamoto, G., Miyazaki, J., & Kato, H. (2014). Augmented reality learning experiences: Survey of prototype design and evaluation. *IEEE Transactions on Learning Technologies*, *7*(1), 38–56. https://doi.org/10.1109/TLT.2013.37

Schmidt, J. T. (2007). *Preparing students for success in blended learning environments: Future oriented motivation and self-regulatio*n (Doctoral Dissertation, lmu, München, Germany). Retrieved from https://edoc.ub.uni-muenchen.de/6561/

Shen, J., & Eder, L. B. (2009). Intentions to use virtual worlds for education. *Journal of Information Systems Education*, *20*(2), 225.

Sumadio, D. D., & Rambli, D. R. A. (2010). Preliminary evaluation on user acceptance of the augmented reality use for education. In *Computer Engineering and Applications (ICCEA), 2010 Second International Conference on* (Vol. 2, pp. 461–465). IEEE.

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, *4*(4), 295–312.

Sweller, J., Van Merrienboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, *10*(3), 251–296.

Tang, A., Owen, C., Biocca, F., & Mou, W. (2003a). Comparative effectiveness of augmented reality in object assembly. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 73–80). ACM.

Tekušová, T., & Kohlhammer, J. (2008). Perception of direction changes in animated data visualization. In *Proceedings of the 5th Symposium on Applied Perception in Graphics and Visualization* (pp. 205-205). ACM.

Veinott, E. S.,& Kanki, B. G. (1995). Identifying human factors issues in aircraft maintenance operations. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 39*(14), 950-950. doi:10.1177/154193129503901440.

Watson, G., Curran, R., Butterfield, J., & Craig, C. (2008). The effect of using animated work instructions over text and static graphics when performing a small scale engineering assembly. *Collaborative Product and Service Life Cycle Management for a Sustainable World*, 541–550.

Wickens, C. D., Hollands, J. G., Banbury, S., & Parasuraman, R. (2015). *Engineering psychology & human performance*. Psychology Press.

Zaeh, M. F., & Wiesbeck, M. (2008). A model for adaptively generating assembly instructions using state-based graphs. *Manufacturing Systems and Technologies for the New Frontier*, 195–198.

Appendix

Appendix I: Augmented Reality Marker

Appendix II: Introduction Script of Experiment

1. Purpose of the research

Thank you for join this experiment. Today's experiment is a part of my master thesis project. The topic is about augmented reality instruction.

2. Benefits and risks of participating

This research project has been reviewed and approved by the BMS Ethics Committee. This experiment is very relaxing and fun. You will play with LEGO by the following instruction.

3. Procedure

Today's experiment consists of three parts, lasting about 30mintues.
Firstly, I will guide you to be familiar with the equipment and experiment material. If you have any questions, just let me know.
Then, we will do the main experiment. In this session, you will be assigned an instruction and guided to assemble a Lego car model. A camera will record the process to help the researcher collect data.
Finally, you need to fill in three surveys and we will conduct a short interview together. The interview will be voice-recorded.
Before we start, could you fill in this **consent form** and a survey firstly?

4. Participant information collected

Your personal information, including gender, age, and major, will be collected by a survey. All data will be anonymous and be only identified by using a participant ID. You have the right of the participant to request access to and rectification or erasure of personal data. The retention period of research data is from 30 May to 30 September.

5. Contact details of the researcher

Name: Yumeng Yang   Email: y.yang-2@student.utwente.nl

Appendix III: Consent Form

Please tick the appropriate boxes	Yes	No
Taking part in the study		
I have read and understood the study information dated 30/5/2018, or it has been read to me. I have been able to ask questions about the study and my questions have been answered to my satisfaction.	○	○
I consent voluntarily to be a participant in this study and understand that I can refuse to answer questions and I can withdraw from the study at any time, without having to give a reason.	○	○
I understand that taking part in the study involves an audio-recorded interview, a survey questionnaire, and a video-recording. The audio recordings will be transcribed as text. The video-recordings will be used to the task performance of participants (time& error).	○	○
Use of the information in the study		
I understand that information I provide will be used for master thesis.	○	○
I understand that personal information collected about me that can identify me, such as [e.g. my name or where I live], will not be shared beyond the study team.	○	○
I understand that personal information collected about me that can identify me, such as age, gender, major, related experience will not be shared beyond the study team.	○	○
I agree to joint copyright of the Demographic Information to [Are Augmented Reality instruction better?]	○	○
I agree to be audio recorded.	○	○
I agree to be video recorded.	○	○
Future use and reuse of the information by others		
I give permission for the Demographic Information that I provide to be archived in UT Qualtrics so it can be used for future research and learning. The data will be deposited in the form of survey database. All deposited data will be anonymised and will be only identified with participant number. All data will be used only in research, excluding commercial use.	○	○
Signatures		
Name of participant Signature Date		
I have accurately read out the information sheet to the potential participant and, to the best of my ability, ensured that the participant understands to what they are freely consenting. Researcher name Signature Date		
Study contact details for further information: [*Name, email address*] If you have questions about your rights as a research participant, or wish to obtain information, ask questions, or discuss any concerns about this study with someone other than the researcher(s), please contact the Secretary of the Ethics Committee of the Faculty of Behavioural, Management and Social Sciences at the University of Twente by ethicscommittee-bms@utwente.nl		

Appendix IV: Demographic Information

Link URL: https://utwentebs.eu.qualtrics.com/jfe/form/SV_6gPA8VVEK9UkeWN

- What is your age?
- What is your major?
- What is your gender?

  ☐Female ☐Male
- Have you ever played Lego?

  ☐Yes ☐No
- Have you had experience with AR before?

  ☐Yes ☐No
- Have you had experience with mobile AR manual before?

  ☐Yes ☐No

Appendix V: NASA Task Load Index

Rating scale definition		
Title	**Endpoints**	**Description**
Mental demand	Low/high	How much mental and perceptual activity was required (e.g., thinking, deciding, calculating, remembering, looking and searching.)? Was the task easy or demanding, simple or complex, exacting or forgiving?
Physical demand	Low/high	How much physical activity was required (e.g., pushing, pulling, turning, controlling and activating)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?
Temporal demand	Low/high	How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?
Performance	Good/poor	How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?
Effort	Low/high	How hard did you have to work (mentally and physically) to accomplish your level of performance?
Frustration level	Low/high	How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?

NASA-TLX Instruction

First Part: Rating scale

We are not only interested in assessing your performance but also the experiences you had during the different task conditions. Right now we are going to describe the technique that will be used to examine your experiences, **in the most general sense we are examining the “workload" you experienced.** Workload is a difficult concept to define precisely, but a simple one to understand generally. The factors that influence your experience of workload may come from the **task itself**, your feelings about your **own performance**, **how much effort you put in**. or the **stress** and **frustration** you felt.

Since workload is something that is experienced individually by each person, there are no effective "rulers" that can be used to estimate the workload of different activities. Because workload may be caused by many different factors, we would like you to **evaluate several of them individually** rather than lumping them into a single global evaluation of overall workload. This set of **six rating scales** was developed for you to use in evaluating your experiences when using the instruction. Please read the **descriptions** of the scales carefully.

If you have a question about any of the scales in the table, please ask me about it, It is extremely important that they be clear to you. You may keep the descriptions with you for reference during the experiment.

Second Part: Sources of workload evaluation

Throughout this experiment the rating scales are used to assess your experiences in the different task conditions. Scales of this sort are extremely useful, but their utility suffers from the tendency people have to interpret them in individual ways. For example, **some people feel that mental or temporal demands are the essential aspects of workload regardless of the effort they expended on a given task or the level of performance they achieved. Others feel that if they performed well the workload must have been low and if they performed badly it must have been high. Yet others feel that effort or feelings of frustration are the most important factors in workload; and so on.** The factors that create levels of workload differ depending on the task. For example, some tasks might be difficult because they must be completed very quickly. Others may seem easy or hard because of the intensity of mental or physical effort required. Yet others feel difficult because they cannot be performed well, no matter how much effort is expended.

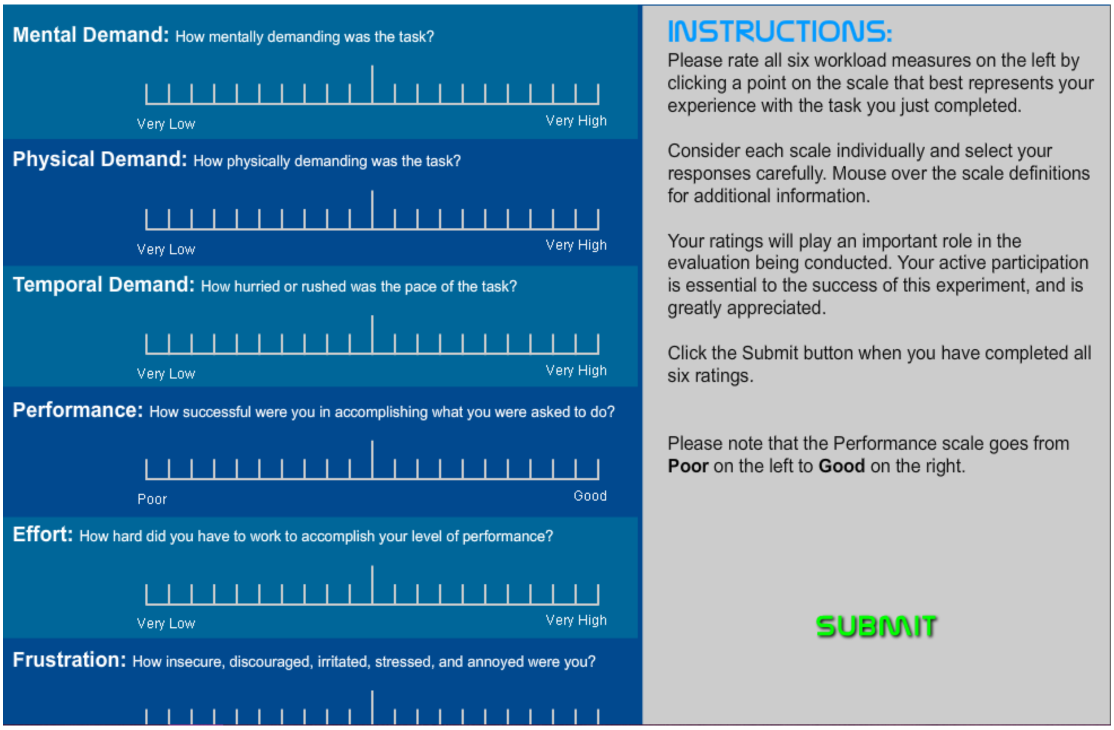
The evaluation you are about to perform is a technique that has been developed by NASA to assess the relative importance of six factors in determining how much workload you experienced. The procedure is simple: You will be presented with a series of pairs of rating scale titles (for example. Effort vs. Mental Demands) and asked to **choose which of the items was more important to your experience of workload in the tasks that you just performed**. Each pair of scale titles will appear on a separate card:

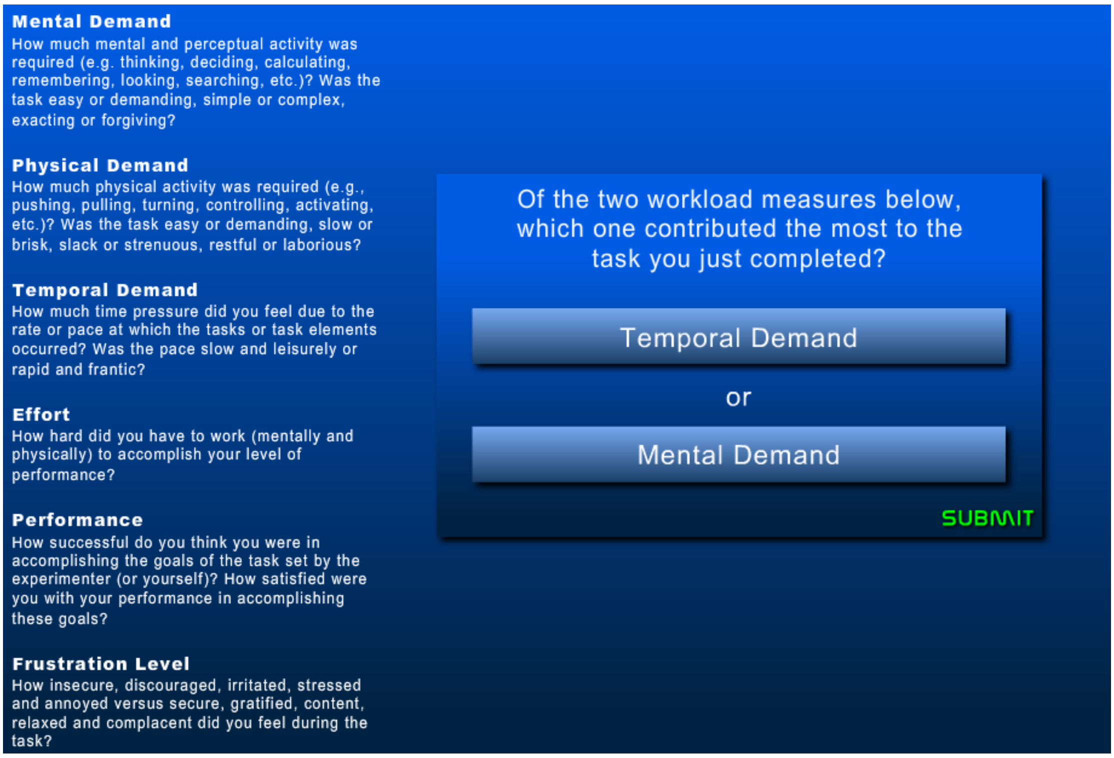
<u>Click the Scale Title that represents the more important contributor to workload for the specific task you performed in this experiment.</u>

Please consider your choices carefully and make them consistent with how you used the rating scales during the particular task you were asked to evaluate. Don't think that there is any correct pattern; we are only interested in your opinions.

If you have any questions, please ask them now. Otherwise, start whenever you are ready. Thank you for your participation.

Link URL： *http://www.nasatlx.com/*





Appendix VI: Reduced Instructional Material Motivation Survey

Link URL: https://utwentebs.eu.qualtrics.com/jfe/form/SV_eE5srcZF310LUK9

Dear participants,

Thanks for your patience.

There are 12 statements in this questionnaire. Please think about each statement in relation to the instructional materials you have just studied and indicate how true it is. Give the answer that truly applies to you, and not what you would like to be true, or what you think others want to hear. Think about each statement by itself and indicate how true it is. Do not be influenced by your answers to other statements.

Questions are answered on the following scale: 1) Not true, 2) Slightly true, 3) Moderately true, 4) Mostly true, 5) Very true

Please consider your choice carefully. Your rating will play an important role in the evaluation being conducted. Thanks for your cooperation.

11A03 The quality of the instruction helped to hold my attention.

17A06 The way the information is arranged using augmented reality technology/on the pages helped keep my attention.

28A10 The variety of images, texts, audio, and animations (if there are) etc. helped keep my attention on the instruction.

06R01 It is clear to me how the content of this instruction is related to things I already know.

23R06 The content and style of this instruction convey the impression that its content is worth knowing.

33R09 The content of this instruction will be useful to me.

13C05 As I worked with this instruction, I was confident that I could understand the content.

25C07 After working with this instruction manual for a while, I was confident that I would be able to complete the task.

35C09 The good organization of the content helped me be confident that I would learn this instruction.

14S02 I enjoyed using this instruction so much that I would like to know about this topic.

21S03 I really enjoyed using this manual.

36S06 It was a pleasure to use such a well-designed instruction.

Appendix VII: Instruction Experience Survey

Link URL: *https://utwentebs.eu.qualtrics.com/jfe/form/SV_1TjleZNQigJCTMp*

Background: You want to buy a new LEGO^{TM} car model. When entering a LEGO shop or browsing online information of LEGO^{TM} products, you see lots of car model boxes or images. You can also use the instruction (just like the instruction you used before). After using with this instruction, please answer these questions below and tell us your experience.

Questions are answered on the following scale: 1) strongly disagree, 2) disagree, 3) neutral, 4) agree, 5) strongly agree

This instruction…….

1. Impacted my overall experience positively
2. Is easy to use.
3. Operation is clear and understandable
4. Made me feel the below mentioned feelings
   - Playful
   - Positively surprised
   - Joyful
   - Dull
   - Unpleasant
   - Boring
5. Influenced my interest in the brand positively.
6. Made me want to try more products of LEGO than I usually consider.
7. Influenced my intention positively to purchase a LEGO product.

Appendix VIII: Semi-structured Interview

Introduce the purpose of the interview:

Hi, thank you for joining the experiment today. Now, we will finish the last part of our journey. I would like to ask a few questions to you. You are free to answer anything that you want and don’t be scared. I will take 3-5 minutes. Do you have any questions before we start?

Questions

1. How do you feel after using this instruction? Please describe your experience.
2. If you have to use one word to describe your experience? What would that be? Why?
3. What do you like/dislike about the application?

Appendix IX: Factor Analysis

KMO and Bartlett's Test

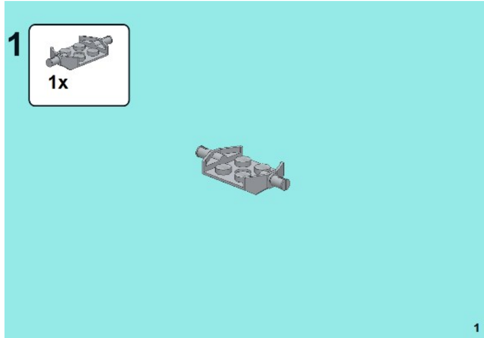
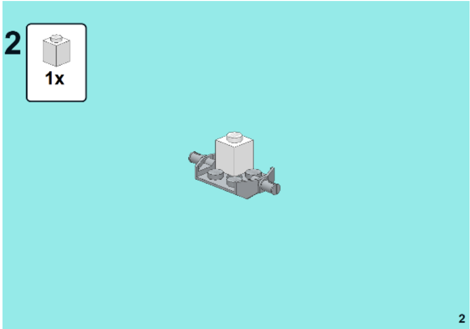
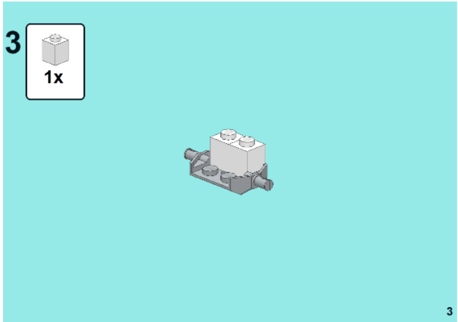
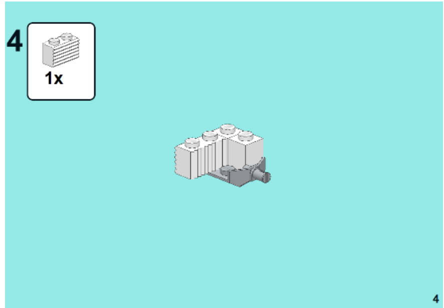
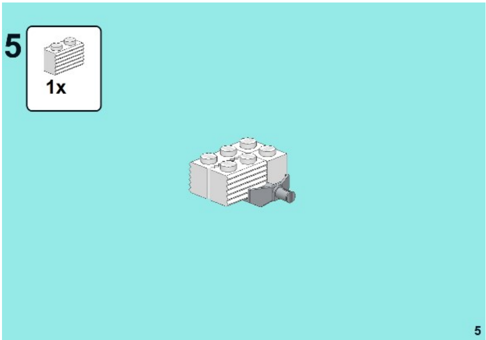
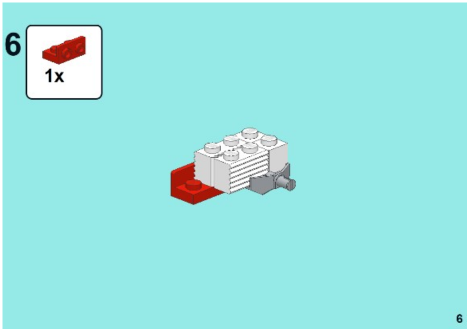
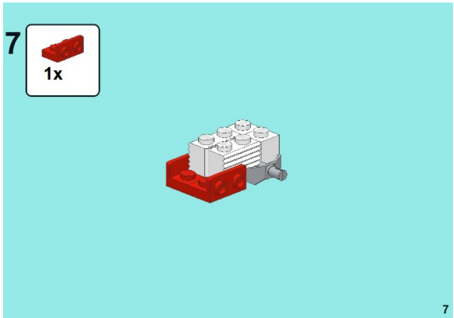
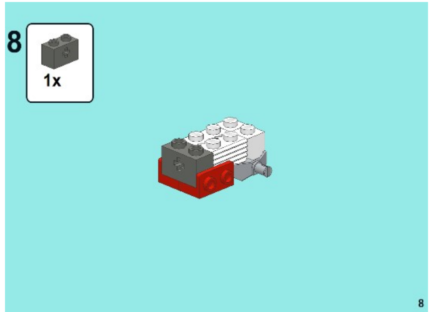
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.746
Bartlett's Test of Sphericity	Approx. Chi-Square	311.838
	df	66
	Sig.	.000

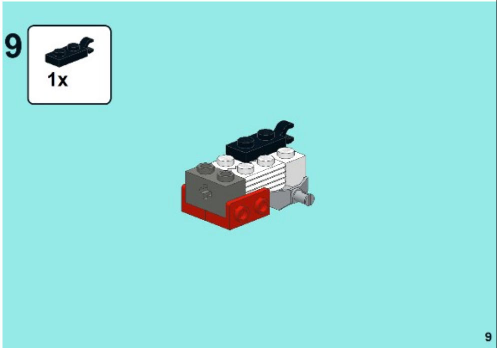
Rotated Component Matrix[a]

	Component			
	1	2	3	4
Q4_2: Made me feel the below mentioned feelings - positively surprised	.867			
Q4_3: joyful	.765			
Q4_1: playful	.719			
Q5: Influenced my interest in the brand positively	.684			
Q1: Impacted my overall experience positively	.460	-.423		
Q4_5: unpleasant		.826		
Q4_6: boring		.797		
Q4_4: dull		.685		
Q2: Is easy to use.			.884	
Q3:Operation is clear and understandable			.872	
Q6: Made me want to try more products of LEGO than I usually consider				.849
Q7: Influenced my intention positively to purchase a LEGO product.				.849

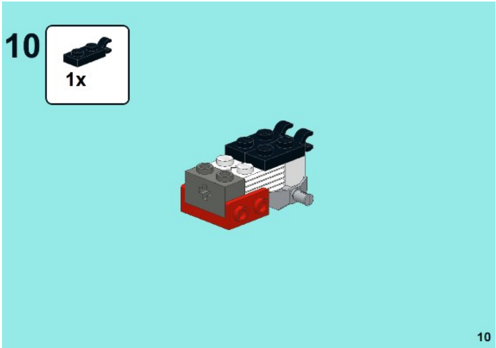
Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 5 iterations.

Appendix X: Error Recording Table

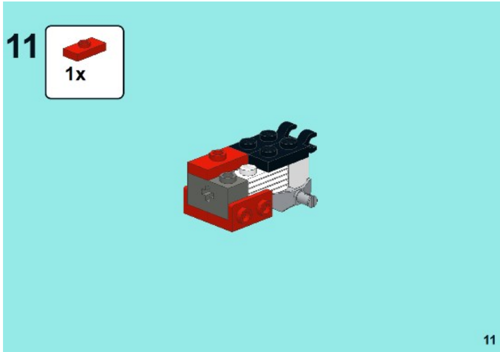
<table>
<tr><td colspan="4">Participant ID:</td></tr>
<tr>
<td>
■ Wrong selection □□□□□□□
■ Wrong installation □□□□□□□
■ Skip □□□□□□□</td>
<td>
■ Wrong selection □□□□□□□
■ Wrong installation □□□□□□□
■ Skip □□□□□□□</td>
<td>
■ Wrong selection □□□□□□□
■ Wrong installation □□□□□□□
■ Skip □□□□□□□</td>
<td>
■ Wrong selection □□□□□□□
■ Wrong installation □□□□□□□
■ Skip □□□□□□□</td>
</tr>
<tr>
<td>
■ Wrong selection □□□□□□□
■ Wrong installation □□□□□□□
■ Skip □□□□□□□</td>
<td>
■ Wrong selection □□□□□□□
■ Wrong installation □□□□□□□
■ Skip □□□□□□□</td>
<td>
■ Wrong selection □□□□□□□
■ Wrong installation □□□□□□□
■ Skip □□□□□□□</td>
<td>
■ Wrong selection □□□□□□□
■ Wrong installation □□□□□□□
■ Skip □□□□□□□</td>
</tr>
</table>



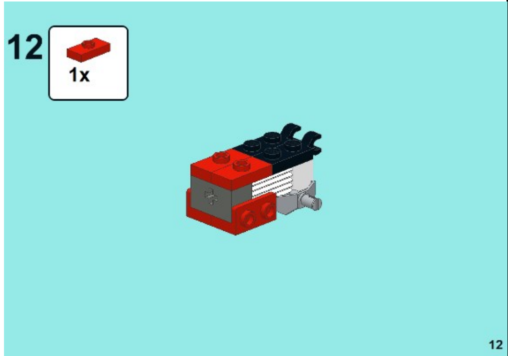
- Wrong selection □□□□□□□
- Wrong installation□□□□□□□
- Skip □□□□□□□



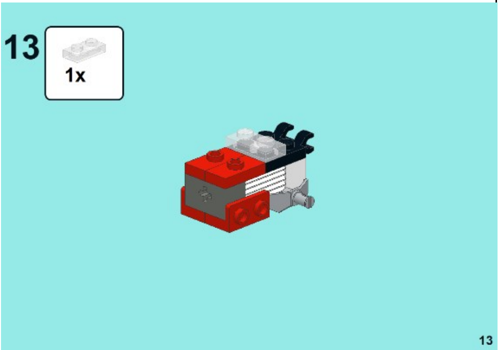
- Wrong selection □□□□□□□
- Wrong installation□□□□□□□
- Skip □□□□□□□



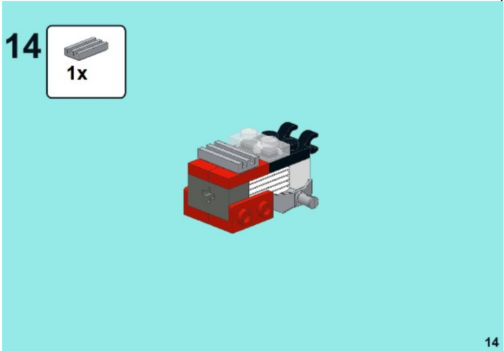
- Wrong selection □□□□□□□
- Wrong installation □□□□□□□
- Skip □□□□□□□



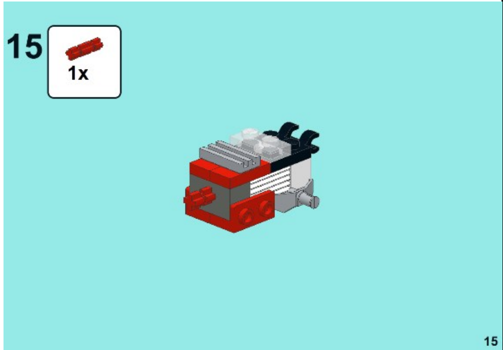
- Wrong selection □□□□□□□
- Wrong installation □□□□□□□
- Skip □□□□□□□



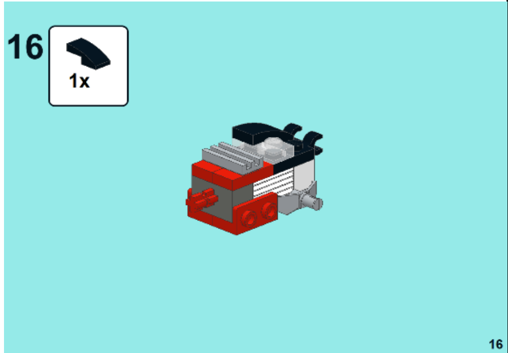
- Wrong selection □□□□□□□
- Wrong installation□□□□□□□
- Skip □□□□□□□



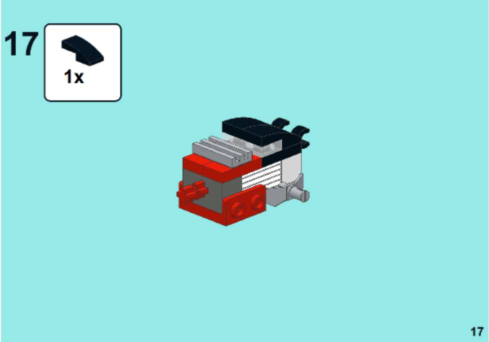
- Wrong selection □□□□□□□
- Wrong installation□□□□□□□
- Skip □□□□□□□



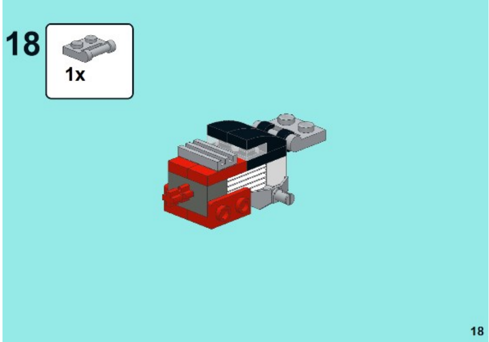
- Wrong selection □□□□□□□
- Wrong installation□□□□□□□
- Skip □□□□□□□



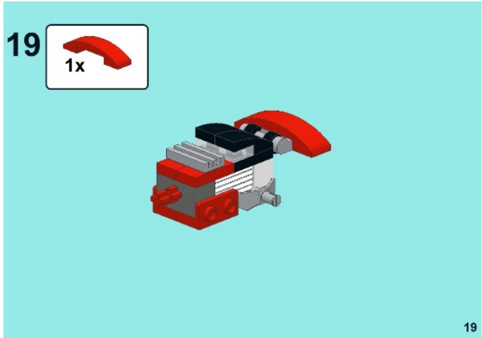
- Wrong selection □□□□□□□
- Wrong installation□□□□□□□
- Skip □□□□□□□



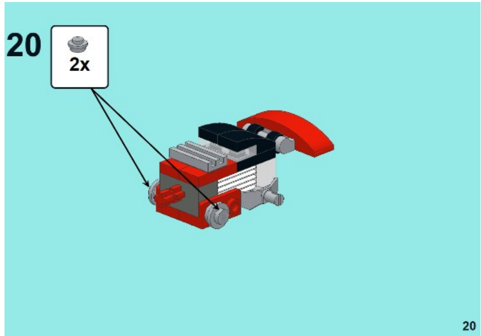
- Wrong selection □□□□□□□□
- Wrong installation□□□□□□□□
- Skip □□□□□□□□



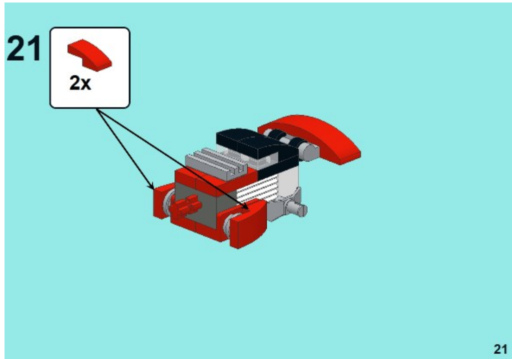
- Wrong selection □□□□□□□□
- Wrong installation□□□□□□□□
- Skip □□□□□□□□



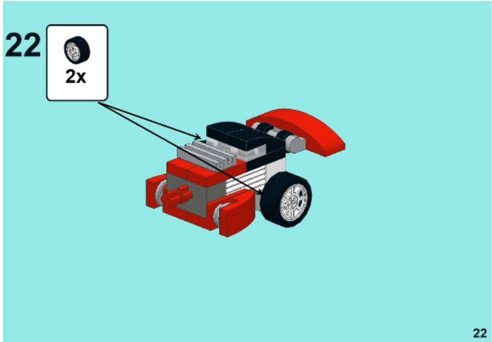
- Wrong selection □□□□□□□□
- Wrong installation □□□□□□□□
- Skip □□□□□□□□



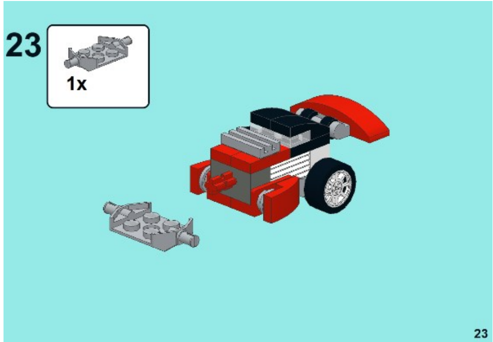
- Wrong selection □□□□□□□□
- Wrong installation □□□□□□□□
- Skip □□□□□□□□



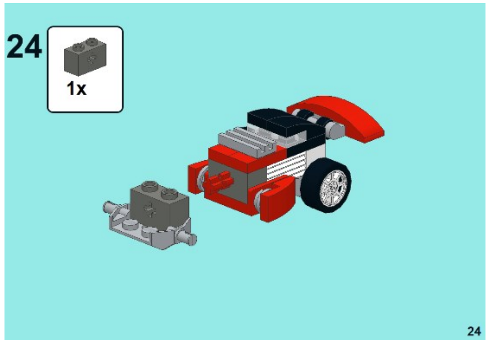
- Wrong selection □□□□□□□□
- Wrong installation□□□□□□□□
- Skip □□□□□□□□



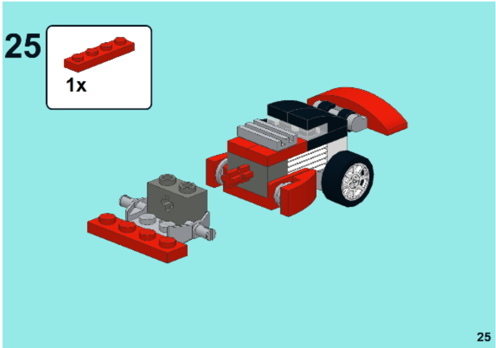
- Wrong selection □□□□□□□□
- Wrong installation□□□□□□□□
- Skip □□□□□□□□

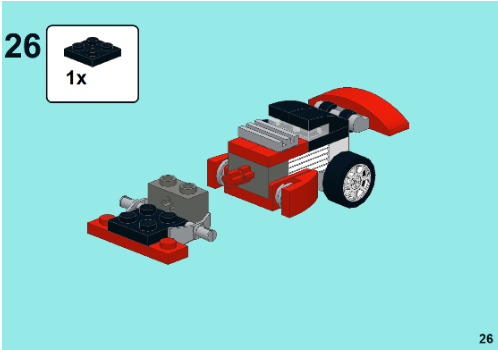


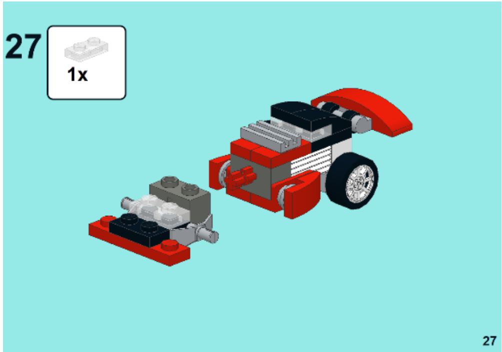
- Wrong selection □□□□□□□□
- Wrong installation□□□□□□□□
- Skip □□□□□□□□

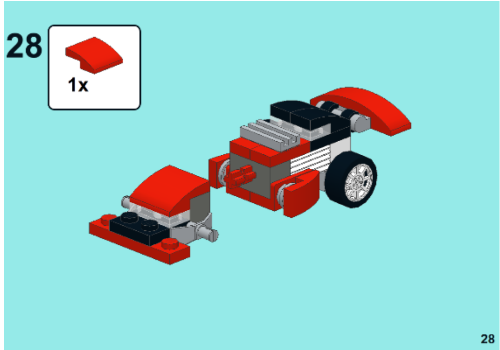


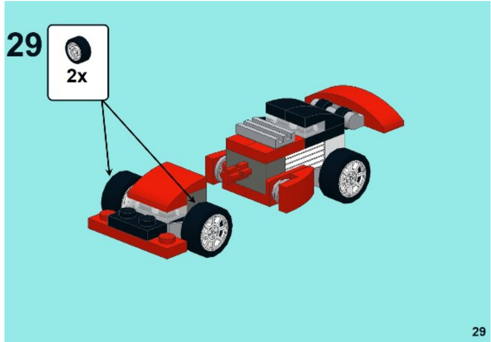
- Wrong selection □□□□□□□□
- Wrong installation□□□□□□□□
- Skip □□□□□□□□

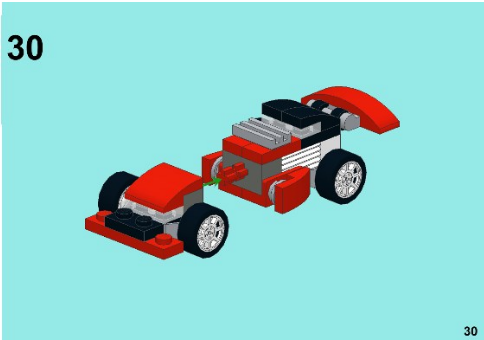
<table>
<tr>
<td>

■ Wrong selection □□□□□□□
■ Wrong installation□□□□□□□
■ Skip □□□□□□□</td>
<td>

■ Wrong selection □□□□□□□
■ Wrong installation□□□□□□□
■ Skip □□□□□□□</td>
<td>

■ Wrong selection □□□□□□□
■ Wrong installation □□□□□□□
■ Skip □□□□□□□</td>
<td>

■ Wrong selection □□□□□□□
■ Wrong installation □□□□□□□
■ Skip □□□□□□□</td>
</tr>
<tr>
<td>

■ Wrong selection □□□□□□□
■ Wrong installation□□□□□□□
■ Skip □□□□□□□</td>
<td>

■ Wrong selection □□□□□□□
■ Wrong installation□□□□□□□
■ Skip □□□□□□□</td>
<td>■ Total wrong selection
____
■ Total wrong installation
____
■ Total Skip
______________</td>
<td>Total error:
______________</td>
</tr>
</table>