ITU A|Z • Vol 19 No 1 • March 2022 • 225-242

Exploring usability tests to evaluate designers' interaction with mobile augmented reality application for conceptual architectural design

Mahmut Çağdaş DURMAZOĞLU[1], Leman Figen GÜL[2]
[1] cagdas.durmazoglu@gmail.com • Department of Informatics, Faculty of Architecture, Istanbul Technical University, Istanbul, Turkey
[2] fgul@itu.edu.tr • Department of Architecture, Faculty of Architecture, Istanbul Technical University, Istanbul, Turkey

Received: February 2021 • Final Acceptance: July 2021

Abstract

Architectural design could be defined as a process in which the information about various aspects of a design object is produced at different stages, then shared between relevant stakeholders. For the early stages of the architectural design, in order to carry out an efficient and productive design process, and create comprehensive solutions to design problems, novel digital tools have to be developed. Digital become omnipresent in the contemporary architectural practice. Even though very complex design ideas could not be realized without the help of the Computer-aided Design (CAD) tools, some may argue that the idea generation is hindered by the employment of them during the early phases. With the idea of containing the essence of analogue tools, we propose an Augmented Reality (AR) application for architects to explore 3D mass geometries in a similar immediacy and ease of designing with a pen-paper. In this paper, we present and discus the validity of two chosen usability scales System Usability Scale (SUS) and Handheld Augmented Reality Usability Scale (HARUS) measuring comprehensibility of the developed AR application (MimAR). The results of the preliminary study shows that the chosen measurement methods provide a similar tendency of scores. The findings of this study suggests that the comprehensibility of the MimAR application still needs improvement, and yet the usability of the application could be considered above the acceptable threshold.

Keywords
Architectural design, Mobile augmented reality, Usability tests.

doi: 10.5505/itujfa.2022.90688

1. Introduction

Early phases of the architectural design process, also called the conceptual design phase, is the foundation of any building design. Even though, data generated during this phase is not more precise and comprehensive than the data required to manufacture / build the design artefact at the end, most of the major decisions are made in the early phase. Nowadays architectural design projects that are complicated and big in scale could be completed in a short time by several design teams collaboratively working together without being limited to space or time. This new type of workflow has become essential to be able to work in the architectural practice in the digital era. Apparently, using analogue architectural tools and methods have already been outweighed by the benefits of using Computer Aided Design (CAD) tools and methods in terms of data storage, simulation, precision, time, and effort. Thus, digital become omnipresent in the contemporary architectural practice giving designers new ways to think, collaborate, design, build and fabricate. Evidently, digital tools have continued to be an essential part of the design practice over the years, and their influence on both theory and practice is going to increase in the coming years. Naturally, being proficient in using at least one CAD tool as well as being familiar with additional rendering and image-video editing programs became a requirement for a new graduate.

As a continuing process, to harness the potential and fully benefit from the recent advancements of digital era, and enhance the architects' ability to develop artefacts, new digital modes of design need to be defined. Based on lessons learned from previous studies, there is a consensus on the inherent characteristics of the analogue design tools such as speed of articulation, ambiguity and familiarity that make them preferable in the conceptual design phase. Any proposed new tool for the conceptual phase should at least have a combination of those characteristics. With this motivation, the very promising and rapidly growing field of advanced digital technologies is investigated to propose a new digital design mode for enhancing, in a way "augmenting", common working environment of the conceptual designer, we named the proposed system as MimAR. MimAR is an Augmented Reality (AR) application where designers overlay the virtual design artefact on the site plan to study 3D mass geometry of a building in easy and fast manner. The development of the MimAR and the results of the pilot study with the participation of expert and novice designers are presented in this paper. In order to find a suitable method for evaluating the usability of the MimAR and gain the necessary insight into the various methods for assessing the computer/digital applications, the field of Human Computer Interaction (HCI) is reviewed.

The main contribution of the paper is to define and adapt a reliable evaluation method to be used in the development process of AR applications. The findings of this research can be used to measure the usability of a system and make necessary revisions to AR applications in order to increase their comprehensibility and efficiency.

2. Views on analogue vs digital tools and methods in design

Defining design activity as a problem-solving strategy and understanding the employment of different design environments and their impact both on designer's cognitive processes and on design artefact are well studied (Cross, 2011; Kalay, 2006; Simon, 1995; Simon, 1996). Design process described as the evolution of design ideas (Goel, 1995) and during the course of design process, several representation methods expressing the design idea as well as containing the information for building and manufacturing of the design artefacts are used (Akın & Weinel, 1982; Goldschmidt, 2004; Gül 2018).

The focused body of the work could be examined on the comparison between digital and analogue design media and tools. In terms of the types of design representation, almost every other design tool has something different to offer to its users, summarized as follows:

First, the advantages of using analogue design over digital design methods and tools in the conceptual design phase are stated by Gross & Do (1996); Aliakseyeu, Martens and Reuterberg

(2006). In the early phases of the architectural design process, expressing design ideas by using freehand sketches is the frequently preferred method for designers. Because using freehand sketches provide speed, flexibility, ambiguity, and instinctive interaction with them. Additionally, the use of those traditional design methods and tools allows designers to reflect in action (Schon, 2008) to make new connections, and re-interpret their design ideas, and finally formulate new solutions by fostering creative thinking with ambiguity (Goel, 1995; Goldschmidt, 1991; Suwa, Purcell, & Gero, 1998). Sachse, Leinert and Hacker's (2001) study on the effects of sketching on design problem solving shows that sketching improves the quality of design solutions, and reduces the overall time spent on design. Won (2001) studied cognitive visual thinking by comparing designers' behavior in conventional and digital media and reported that even though designers' cognitive behavior was much more complex while using computer media to generate concepts or ideas, they could generate more concepts while using conventional media than using computers. Although it was done in the context of graphical design, Stones and Cassidy's study (2007) also showed that the paper-based sketching was proved to be a more productive environment for the exploration of design solutions than the digital environment. Here, the immediacy, speed and ambiguity of the used tools become particularly vital to facilitate design thinking. Thus, we consider that the proposed systems for design activity should provide such speed, flexibility, and ease.

Second, several studies stated the advantages of simultaneously or sequentially using both design environments (Chen, 2007; Ibrahim & Rahimian, 2010; Shih, Sher, & Taylor, 2017). Even though, it was reported (Tang, Lee, & Gero, 2011) that using either environment would not make a difference in terms of cognitive activities of designers, Chen (2007) suggested that the employment of analogue and digital media and tools simultaneously could foster a more productive design environment in terms of creative thinking and cognitive activities, and improve the quality of the design in the conceptual design process. Ibrahim and Rahimian (2010) stated that the designers using both traditional sketching and CAD modelling tools produced significantly higher quality design solutions, and spent less time on formulating design solutions compared to the subjects using either the analogue or CAD modelling tools. Shih et al. (2017) also reported that both analogue and digital design media and tools play very similar roles in design process.

Lastly, several studies addressed and challenged the established notion of necessity, and the superiority of using analogue design media and tools in the early stages of the design process over the use of digital design media and tools (Boeykens, Santana Quintero, & Neuckermans, 2008; Reffat, 2007). During the design process, design ideas are represented in various scales. Using digital design media and tools in the design process make it possible for architects to work on a drawing in various scales without the need to reproduce the original drawing from the scratch. In addition, digital design tools make it possible to create clear and understandable design documentation by eliminating imperfections resulted by working with analogue tools such as inconsistent drawing quality, imprecise registration and low graphic quality (Mitchell & McCullough, 1995). Apart from the flexibility, detail and precision in the drafting process, many current digital tools such as Rhino and AutoCAD also provide spatial comprehensibility by fostering a suitable environment to evaluate design ideas from different viewpoints on the fly. Traditional architectural design methods and tools, albeit proven to be very effective in other facets of the design process throughout the time, are not considered as a suitable method to be used in the formation, evaluation and representation of complex forms (Lin, 2001). Studies show that students who were working with the Computer Aided Architectural Design (CAAD) tools had a better grasp of the quality of the internal spaces, on the contrary, students who were working with the analogue tools rarely explored their designs in 3D (Knight et al., 2005). Thus, thinking about CAD not only as a tool to produce technical drawings, but also as the conceptual tool to develop new design ideas become very common (Salman, Laing, & Conniff, 2008). In the conceptual design phase, employing digital media and tools also improves designers' spatial cognition compared to using analogue media and tools (Rahimian & Ibrahim, 2011). In addition, some studies also shows that different digital design environments afford different types of ideation and design actions in relation to the tools in the environment (Gül, 2008, 2018).

Based on above review and considering continuing advancements on technology, clearly the debate on the use of tools and technology on architecture would exist in future. There would be a constant need for research projects, which confront the established tools and technology with developing new systems that incorporate the strength of analogue design tools. With this idea in mind, an AR application that can augment its users' ability to be able to create meaningful spatial layouts of 3D geometries as fast, easy, and intuitive as manipulating forms with the analogue tools is developed.

3. MimAR for conceptual architectural design

Augmented reality is an advanced information technology that first developed by Sutherland (1968) and coined as "Augmented Reality" for the first time by Caudell and Mizell (1992). Even though AR technologies had been studied since the early 70's, due to their reliance on technological advances in the hardware and software, the progress on the AR research has been slow and limited. The recent advances in the computer technologies have made it possible for the end user to afford very powerful mobile devices with internal sensors. Azuma (1997; 2001) explained AR as a technology that makes it possible for an artificial image to be generated in the real environment while allowing both embedding the artificial image into real environment and interacting with it at the same time. The interaction with the system can be through intangible gesture based (Ens et al., 2017; Funk, Kritzler, & Michahelles, 2017; Hürst & van Wezel, 2011) or with wearable display devices (Gruenefeld et al., 2017; Zimmer et al., 2017).

We argue that AR is beneficial to use in developing design support tools that could incorporate useful features of analogue design media and tools such as flexibility, speed, intuition, and ambiguity in design. In addition, in the early stages of the architectural design process, a fruitful design process in which designers define massing proposals and evaluate their spatial relationships could be performed by using the AR application (Gül, 2018). MimAR allows users to work on and evaluate virtual 3D geometries that could be superimposed in a real-world environment. The inherent ability of AR allows its users to be able to work with virtual objects within a physical environment providing a conceptual design environment where mostly 'massing studies took place' (Gül, 2017).

The main aim is to develop a user-friendly AR application for architects so that they could quickly learn how to operate the system and able to create mass design and evaluate it in the conceptual phase of architectural design through a mobile device. However, developing the MAR application with just architecture graduates in mind would mean excluding the tech friendly new generation of potential architects, a large group of people who are still in the process of completing their education. Therefore, it was decided to develop the MAR application for both novice and expert architects. Some of the basic operations of MimAR includes generating pre-defined 3D geometries, copying, deleting, moving, rotating, and scaling objects. Additionally, the MimAR makes it possible for its users to visually inspect the spatial layout and relationships of design mass in a fast, effective, and economical manner in a way that is motivating, fun and familiar for the users.

3.1. Graphical user interface, application components, and interaction modes of MimAR

A typical augmented reality system consists of four main components such as trackers, computational devices, display devices as outputs and lastly data input devices (Figure 1). Since it

Equation 2. HARUS Calculation

Contribution Score = CS
User Response = UR
Final Usability Score = FUS

CS (PositivelyStatedQuestions) = UR-1
CS (NegativelyStatedQuestions) = 7-UR
$FUS = ((CS_1 + CS_2 + \ldots + CS_n) / 48) \times 100$

Figure 1. *Augmented reality system components (Adapted from (Wang, 2009)).*

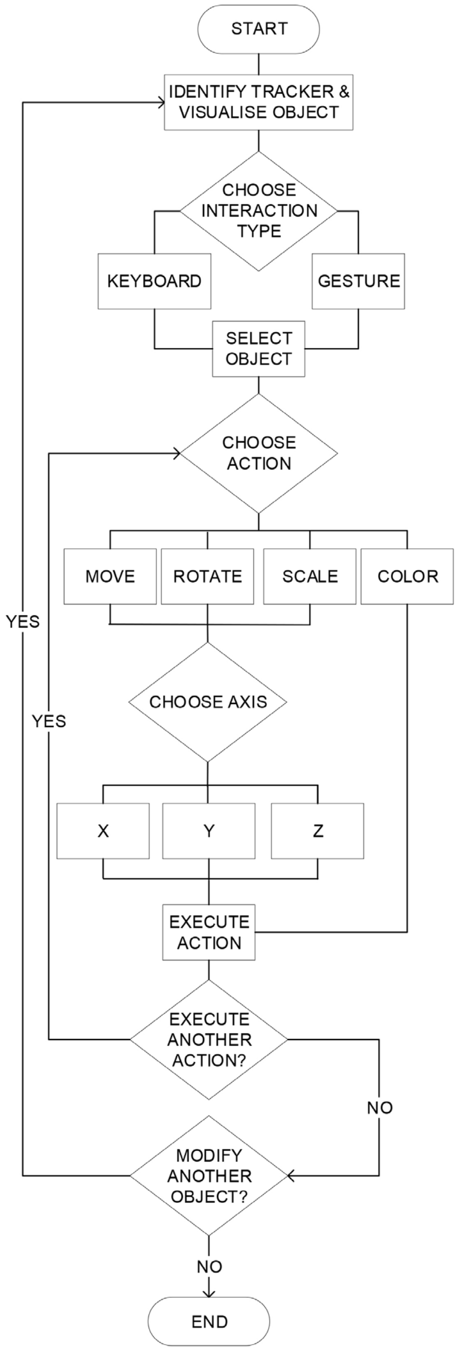


Figure 2. *Flowchart of the MimAR application.*

is developed for smart mobile devices, MimAR uses the computational, display and data input capabilities of the smart mobile devices, given that the device meets the minimum system requirements listed in the section 3.2. To be more precise, MimAR uses device's camera to identify and track virtual objects in the physical environment; the touch screen to display virtual objects and to input data via the application's Graphical User Interface (GUI). In addition, QR based trackers are used to keep the computational load low and give users a physical anchor related to the virtual objects.

The real world and the virtual objects can be seen through MimAR's interface on the smart mobile device's screen. To create geometries with MimAR, first, 4x4 cm QR trackers, which have the information required to visualize predefined 3D objects, such as the type, initial color, size, and orientation, must be scanned through the camera of the mobile device. Virtual objects that are introduced to the scene could be further manipulated after users choose an interaction type. Users could choose either the virtual keyboard or touch screen gestures to use the application. After the interaction type is chosen, the sub-menu box that houses manipulation commands such as rotate, move, and scale, becomes visible. A relevant axis that the manipulation would be implemented on should also be selected after users select a manipulation command. When users are satisfied with the results of the manipulation a new action could be implemented on the selected object or users could select another object to manipulate (Figure 2).

The application "MimAR" provides its users the ability to visualize their massing proposals in different visual viewpoints as well. Additionally, it allows users to inspect and evaluate their massing proposals in the real world with its familiar and intuitive interaction methods and interface. MimAR's GUI is designed to be visually as clear as possible in order to minimize the possible confusion and frustration that the users might feel while using the application. MimAR's GUI basically consists of an always-visible main menu ribbon

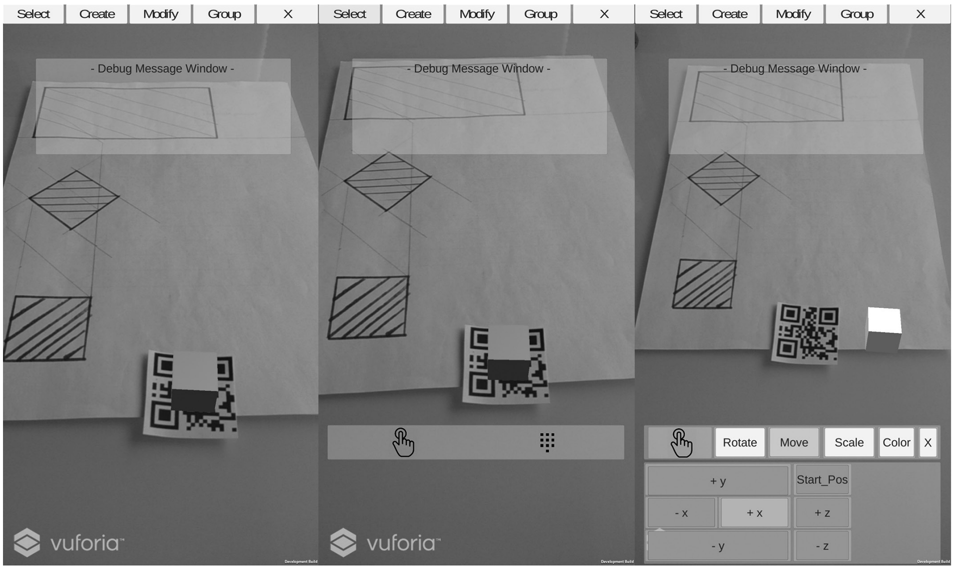


Figure 3. *Graphical user interface of the MimAR application.*

and a sub-menu box that only becomes visible when needed. Commands such as select, create, modify, group and exit could be found on the main menu ribbon located at the top of the screen (Figure 3). The sub-menu box, which is invisible in its inactive state, is located at the bottom of the screen. After users introduce a virtual object to the system and select it, the sub-menu box, in which commands that are related to the selected object such as interaction type, rotation, scale, and move could be found, becomes visible.

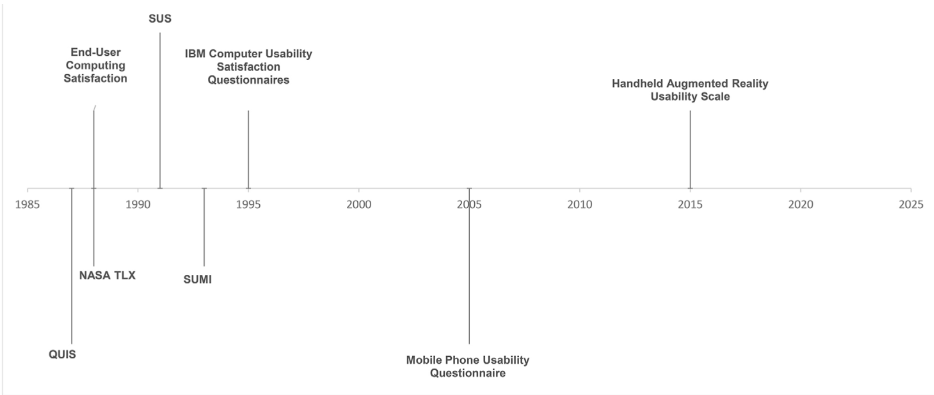
3.2. Hardware and software features of the application

"MimAR" was developed for smart mobile devices, which are known to be user friendly and economically reachable. The system requires a minimum Android 4.0 operating system with a functional back camera. During the usability studies, a smart mobile device that has a screen resolution of 1080x1920 pixels, 5.0' touch screen, a back camera with the resolution of 13.1 Megapixels, with Android 5.1.1 operating system was used. The AR application was developed by using the 2017.3.0f3 build of the Unity Game Engine with Vuforia add-on, which is widely used in the industry to develop 2D and 3D games and simulations. In addition, Visual Studio 2015 Community Edition and C# programming language was used to implement some of the behavioral features to GUI components and virtual objects.

4. The concept of usability and usability studies

The broader concept of usability is developed within the field of Human Computer Interaction. In order to evaluate systems, usability tests are generally implemented to applications to decide whether they meet certain requirements such as ease of use, user satisfaction, not having distracting components, keeping the users' motivation and performance at a certain level. Participants of the usability tests are expected to evaluate the system according to user experience and the results are usually collected via interviews or questionnaires. However, thorough observation of the participants' actions and experience during the tests could also yield supplementary data for the researchers. The results of the usability tests and the insight gained from these experiments are used to make the necessary revisions to the developed systems. To find out a reliable method to evaluate the usability of the MimAR application, several tests are utilized with the participation of users with different levels of expertise.

A detailed literature review into the field of human computer interaction shows that several evaluation methods were developed in various research projects over the years (Figure 4) such as Questionnaire for User Interface Satisfaction (Chin, Diehl, & Norman, 1988), Software Usability Measurement Inventory (Kirakowski & Corbett, 1993), System Usability Scale (Brooke,



Year	Researchers	Name of the Scale	Num. of Questions	Categories	Scale	Focus of Evaluation
1987	Chin, Diehl, Norman	Questionnaire for User Interface Satisfaction	27	5	Likert (10)	Software
1988	Doll & Torkzadeh	End-User Computing Satisfaction	12	5	Likert (5)	Software
1988	Hart & Staveland	NASA Task Load Index	6	6	Bipolar	Hardware & Software
1991	J. Brooke	System Usability Scale	10	-	Likert (5)	Software
1991	Lewis	After Scenario Questionnaire (IBM - ASQ)	3	3	Likert (7)	Software
1992	Lewis	Post Study System Usability Qestionnaire (IBM - PSSUQ)	19	4	Likert (7)	Software
1992	Lewis	The Computer System Usability Questionnire (IBM - CSUQ)	19	4	Likert (7)	Software
1993	Kirakowski	The Software Usability Measurement Inventory	50	-	Mult. Choice	Software
2005	Ryu & Smith-Jackson	Mobile Phone Usability Questionnaire	72	6	Likert (7)	Hardware
2015	Santos, Polvi & Taketomi	Handheld Augmented Reality Usability Scale	16	2	Likert (7)	Hardware & Software

Figure 4. *Timeline of usability studies.*

1996), Mobile Phone Usability Questionnaire (Ryu & Smith-Jackson, 2006), Handheld Augmented Reality Usability Scale (Santos et al., 2015; Santos et al., 2014). These evaluation methods differ from each other by several aspects. First of all, most of the evaluation methods have different number of questions ranging from 72 (Mobile Phone Usability Questionnaire (Ryu & Smith-Jackson, 2006)) to 3 (IBM's After Scenario Questionnaire (Lewis, 1995)). Second of all, apart from SUS (System Usability Scale) (1996) and The Software Usability Measurement Inventory (Kirakowski & Corbett, 1993) questionnaires, all of these questionnaires have explicitly defined categories such as; content; accuracy; ease of learning/use/task completion; efficiency; system capability etc. In addition, most of the early methods were specifically developed to evaluate only the software without any consideration given to capabilities of the hardware these systems operate on. Because around 1990's when most of these methods were first developed, the hardware that these systems operated on were cumbersome and had similar processing capacity. Users of those systems had limited options in terms of input, output, and display devices. For example, technological capabilities of the past made it possible for users to interact with only mouse or keyboards as input methods and use stationary computing and display devices.

Aforementioned methods are investigated in order to make an informed decision about choosing the right evaluation method to test MimAR and then, SUS and Handheld Augmented Reality Usability Scales (HARUS) are decided to be used within the scope of this research.

4.1. SUS and HARUS questionnaires: The why and the how?

The main goal for using these two methods to evaluate the usability of the Mobile Augmented Reality (MAR) application is to ensure the reliability and validity of the application's usability. In addition, comparing the overall results of the usability scores of the same application according to two different usability evaluation methods and reporting the differences (if any) would provide useful insight regarding usability evaluation methods for future research projects. For example, an application might get acceptable usability scores when it was tested with an evaluation method that has a generalized approach to usability. However, that result alone should not be considered as the absolute usability score of that application. As mentioned at the beginning of this section, usability depends on various aspects. Therefore, evaluating the application by using more than one method, and using methods that were able to isolate and evaluate various aspects of the application would

Table 1. *Sample questions from the SUS and HARUS questionnaires.*

System Usability Scale - Sample Questions	Handheld Augmented Reality Usability Scale - Sample Questions
● I found the various functions in this system were well integrated. ● I think that I would need the support of a technical person to be able to use this system ● I needed to learn a lot of things before I could get going with this system. ● I feel very confident using the system.	Comprehensibility Scale Sample Questions ● I think that interacting with this application requires a lot of mental effort. ● I thought that the information displayed on the screen was confusing. Manipulability Scale Sample Questions ● I found it easy to input information through the application. ● I think the operation of this application is simple and uncomplicated.

be beneficial in getting a better idea about the usability of the application. Moreover, being able to identify which aspects of the application needs further improvement would be invaluable in the development process.

One of the methods that was chosen to be used in the evaluation of the application, SUS questionnaire, was formulated in the mid 90's. Other than being the one of the earlier evaluation methods, SUS differs from the other evaluation methods by its scope and ease of application. On the other hand, HARUS questionnaire was developed specifically for evaluating mobile augmented reality applications and, has sub-scales for evaluating the comprehensibility and manipulability of the applications. Getting similar results from both questionnaires would indicate that the SUS, even though it was originally designed to evaluate software that was much more basic compared to HARUS questionnaire, could be used as an indicator of usability in research projects when time is an issue. Furthermore, utilizing an evaluation method such as HARUS, which offers valuable insight into different facets of the application i.e., comprehensibility and usability of the system, would made identifying problems easier.

The data gathered from the results of the usability studies is planned to be used in the future development cycles of the application to pinpoint the necessary areas that needs to be revised according to the user's feedback. The cyclical approach, where user's feedback is integrated, is expected to be resulted in defining novel ways to the development of the AR based design tools. Sample questions from both HARUS and SUS questionnaires are shown in Table1.

Based on the recorded responses of the participants to the SUS and HARUS questionnaires, the usability scores of the application "MimAR" were calculated as instructed in the related studies (Brooke, 1996; Santos et al., 2015). An application's usability score according to SUS questionnaire was calculated by first finding the contribution value of every item on the questionnaire. The contribution value of an item was found by subtracting 1 from the scale value of an item when the item was the odd number in the questionnaire i.e., 1, 3, or 5. When the item has an even number, i.e., 2, 4 or 6, the contribution value of that number was found by subtracting the scale value of that item from 5. The final value of usability was found by multiplying the sum of contribution values by 2.5 (see Equation 1).

MimAR's usability score according to HARUS questionnaire was calculated in a similar way. For positively stated questions, contribution score for the question was found by subtracting 1 from user response. For the negatively stated questions the contribution score for the question was found by subtracting the user response from 7. The overall HARUS score was calculated by dividing the sum of final scores from every question to the highest possible score of 48 and multiplying the result by 100 (see Equation 2).

The results of usability questionnaires could be used as a source to measure the overall usability of a system. Keeping in mind that the overall scores should not be taken as a definite sign of perfect user experience, an application's SUS score should be between 0 to 100

Contribution Score = CS
Scale Value = SV
Final Usability Score = FUS

CS (OddNumbers) = SV-1
CS (EvenNumbers) = 5-SV
$\text{FUS} = (\text{CS}_1 + \text{CS}_2 + \ldots + \text{CS}_n) \times 2.5$

Equation 1. *SUS Calculation.*

Contribution Score = CS
User Response = UR
Final Usability Score = FUS

CS (PositivelyStatedQuestions) = UR-1
CS (NegativelyStatedQuestions) = 7-UR
$$FUS = ((CS_1 + CS_2 + \ldots + CS_n) / 48) \times 100$$

Equation 2. *HARUS Calculation.*

points. Brooke (2013) also suggested that these results should not be interpreted as percentages just because SUS uses a margin between 0 and 100 points. SUS scores could be interpreted according to acceptability ranges (Bangor, Kortum, & Miller, 2008), grading scales (Lewis & Sauro, 2018) or adjective ratings (Bangor, Kortum, & Miller, 2009). Three classification types in acceptability ranges could be explained as unacceptable (below 50 points), marginally acceptable, which consists of a low (from 50 to 63 points) and a high end (from 63 to 70 points), and acceptable (from 70 points and above). SUS scores could also be interpreted by using grade scales. After analyzing data from 241 industrial usability studies, Lewis and Sauro (2018) created a curved grading score that has 11 grades in which a SUS score of 68 is at the center of the range. A SUS score of 68 points seems critical because it is also close to the threshold where an application stops being considered as marginally acceptable and becomes acceptable. Furthermore, SUS scores between 80 to 90 points could be considered as an above average (better than acceptable) while scores above 90 points are considered as the best. Therefore, even though an application with a SUS score between 50 to 70 points could be considered as marginally acceptable, that means the application still needs further improvement.

4.2. Setup of the experiment

The convenience sampling (Creswell, 2012) method was used as the sampling method for the usability studies that includes expert and novice designers (N:5). Having small number of participants are common in user studies to run preliminary tests in the early stage of product development (Nielsen, 2012; Nielsen & Landauer, 1993). The usability study was designed as having two groups of participants that differ from each other in terms of their experience to maintain the credibility of the study. Conducting the study with only expert designers would have resulted in getting high (biased) usability scores because of participants' experience and familiarity with CAD programs and concepts. Having both novice and expert designers in the study would made it possible to determine if the expert designers' expertise in design give them an advantage over novice designers in terms of handling of the application. Additionally, it would also be possible to identify any usability problems that users with different experience levels could encounter.

In the beginning of the experiment, participants were given training about how to operate the AR application based on its two distinct interaction modes: the first one is virtual buttons and the second one is the gestures performed on touch screen. Later, they were asked to design the mass of a building. Once the task completed, participants were asked to answer both HARUS and SUS questionnaires to evaluate the system in order to ensure the validity and reliability of these usability tests and to be able to compare them in the future for any inconsistency (Table 1). Additionally, the duration of the tasks completion times, error rates and the number of task resets were also noted and recorded with the audiovisual devices to obtain the quantitative data during the tests.

The first part of the usability study consisted of several modelling tasks such as, manipulating the virtual object by implementing translation, scale, and transformation operations. The modelling tasks could be classified from the easiest to hardest, such as translation, scale, and transformation operations. Later, available interaction modes and selection operations have been practiced to try out the different features of the application. There was no time limitation set during the experiment, and the participant was allowed to start, stop, resume, or finish the tasks at any given time. After this introduction, the design brief was given to the participants who were reminded the focus of the study, that is the interaction with the application's interface and not their ability to design or the quality of their proposed design solution. The participants of the study were given site plan and brief and asked to develop massing design solutions for a mix-use communal space that would be situated in an area surrounded by artificial lakes and dense foliage, single-family houses and a house complex consisting of row houses. The communal space was expected to have a restaurant, which has closed and semi-open spaces, multi-purpose studios and a gallery space that would be used for exhibitions.

Experiments were conducted in a confined space that has optimal artificial lighting. Participants used a desk as a workspace and other necessary equipment such as the site plan, the brief, QR trackers, drawing paper, rulers, and pens/pencils. Researcher was present at all times as an observer in every session, and each session was recorded with an audio-visual recording device for future reference. The setup of the experiment can be seen in Figure 5.

4.3. Experiment with an expert architect

The expert participant has a PhD degree in architecture, considerable experience with the smart mobile devices, knowledge in information technology concepts and CAD tools. The expert was asked to complete the design tasks by using both interaction methods, respectively: interacting with the virtual objects by using virtual buttons as a data input mode and interacting with the virtual objects by using touch screen gestures (Figure 6). Then, the expert architect completed the HARUS and SUS questionnaires.

The result of the SUS questionnaire with the expert user shows that the interaction mode with the virtual buttons (87.5) was scored higher than the interaction mode with the touch screen gestures (40). When the application was evaluated using the HARUS, a similar tendency in results was also recorded: the interaction mode with the virtual buttons (85.41) was scored higher than the interaction mode with the touch screen gestures (68.74) (as shown in Figure 7). The supplementary data that was collected during the experiment also showed that the expert's satisfaction declined drastically when the gesture-based interaction mode was used. The reason for the expert's dissatisfaction could be due to the several object selection problems that the expert designer encountered in touch-screen gesture mode which increased the expert's task completion time.

MimAR was rated with high scores

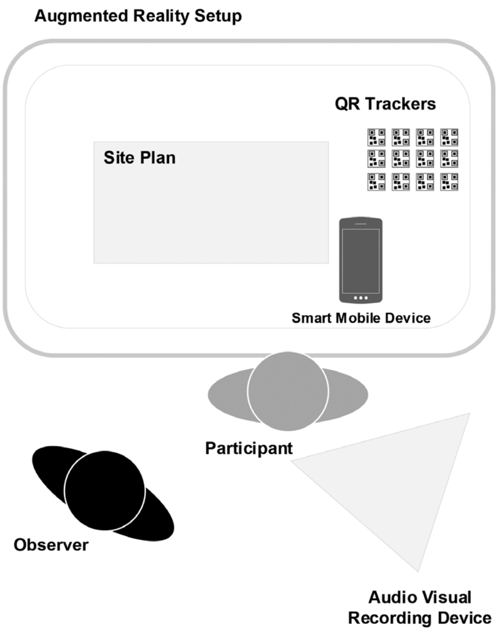


Figure 5. *Setup of the experiment.*

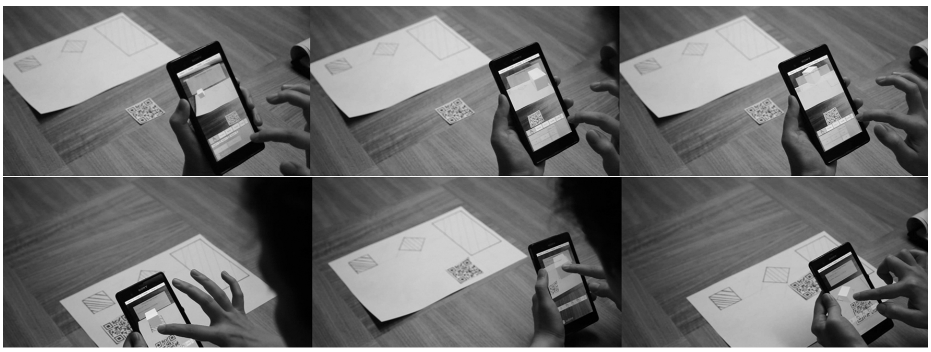
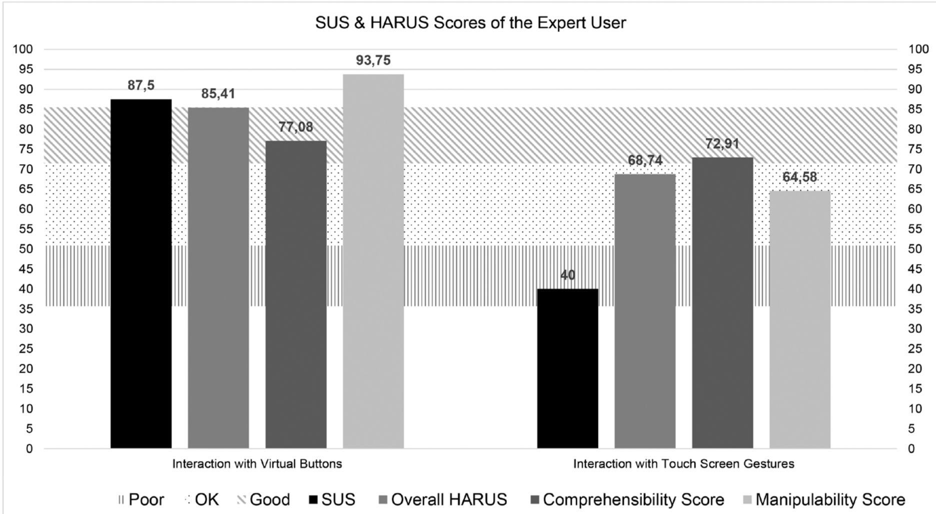


Figure 6. *Still images from the usability study conducted with the expert user.*



	Interaction with the Virtual Buttons	Interaction with the Touch Screen Gestures
SUS Score	87,5	40
Overall HARUS Score	85,41	68,74
Comprehensibility Score	77,08	72,91
Manipulability Score	93,75	64,58

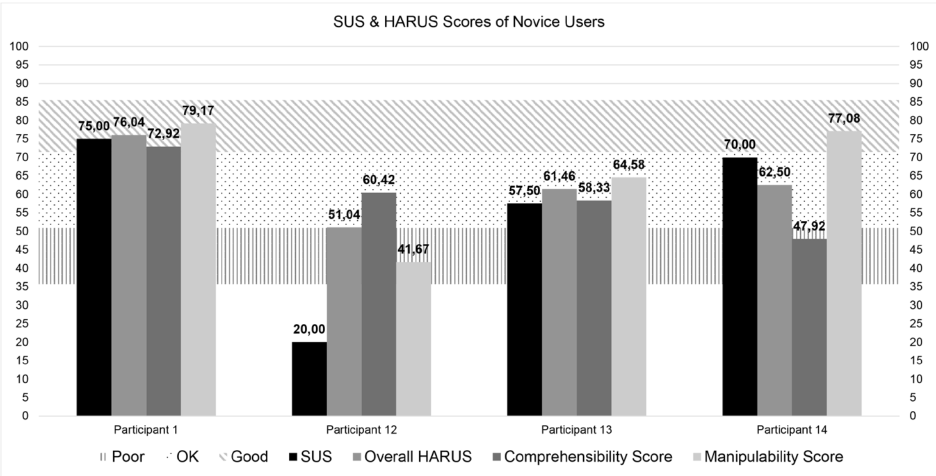
Figure 7. *SUS and HARUS scores of the expert user.*

Figure 8. *Still images from the usability study conducted with the novice users.*

on both HARUS (85.41) and the SUS (87.5) questionnaires while the virtual buttons were used as an interaction mode. The application was evaluated with more than one evaluation method in order to increase the reliability of the chosen methods by comparing those results. As the application was rated with such high scores when either of the evaluation methods was used, MimAR could be considered better than acceptable in terms of usability.

4.4. Experiment with novice architectural design students

The second usability study was carried out with the participation of the four first year architecture students at the X University, Department of Architecture. The participants, in this case consisted of novice designers, were asked to develop design alternatives to a massing design problem using the AR application. Similar to the expert study, with the completion of



Participant Nr.	SUS	Overall HARUS	Comprehensibility Score	Manipulability Score
1	75,00	76,04	72,92	79,17
12	20,00	51,04	60,42	41,67
13	57,50	61,46	58,33	64,58
14	70,00	62,50	47,92	77,08
Score Means	55,63	62,76	59,90	65,63

Figure 9. *SUS and HARUS scores of novice users.*

each design session, the participants were expected to evaluate the perceived usability of the application, as shown in Figure 8.

The result of the second user study showed that mean HARUS score was 62.8 points, and the mean SUS score was 55.6 points (Figure 9). Based on HARUS scores the results show that the mean score of the usability of the application was higher than the mean score of the comprehensibility of the application. Some of the participants reported when they first started to use the application it felt a little overwhelming to use it. They mentioned that it was because of the pressure they feel about using the application successfully. However, they concluded that the feeling dissipates after a couple of minutes of using the application. In addition, one participant reported heating and battery life problems about the hardware at the end of a session. Nevertheless, it was expected to encounter some hardware related heating problems after back-to-back experiment sessions. In order to increase the understandability of the application and to cater the needs of the novice students, several upgrades were planned for the next iteration of the application.

5. Discussion and the conclusion

In this paper, the results of an experiment that was conducted to investigate the usability of an AR application, which was developed as a part of an ongoing PhD study at the X University, is presented. SUS, the older nevertheless a faster method to evaluate usability, and HARUS, relatively recent and more device-oriented method, were chosen as the usability evaluation methods.

Bai et al. (2013) and Hurst and Van Wezel (2013) mentioned the fun factor that the participants experienced while exploring the geometry with gesture based interaction modes. This might be the reason why the current research into the interaction modes in AR applications are shifting towards interactions made with gestures. However, the usability study with the expert participant demonstrates that a more precise mode of interaction is still preferred for completing certain tasks efficiently in the AR environment.

For the intents and purposes of this study, the advantages of using gesture-based interaction do not outweigh the disadvantages such as the requirement of high computing power, high power consumption (Hegde et al., 2016), high monetary and time cost of respectively device and development. The results of the usability study with the expert user show that the usability of the application is rated with high scores in both SUS and HARUS questionnaires while the expert user is using the virtual keypad. According to the results of the HARUS questionnaire, while the comprehensibility of the system is scored similarly with both interaction modes, the usability score of the application is scored considerably lower when the touch screen gestures are used as an interaction mode. In addition, the supplementary data suggest that the expert user's usability ratings and satisfaction with the application decrease as the task completion time is increased while using gesture-based interaction mode.

Despite the gesture-based AR interfaces' rising popularity, an interaction mode that utilizes device's touch screen still seems like the better choice for an effective MAR application for design purposes. Moreover, it is evident in the usability study with the expert user that the gesture-based interaction mode has low accuracy (Hürst & Van Wezel, 2013) and does not provide the precision that the users need in a design activity as Gül (2018) also suggested. The results of the usability study also supports Henrysson et al.'s findings (2007) as the application has acceptable usability scores when evaluated with both SUS and HARUS methods while an expert is using the application with the virtual keypad as an interaction method.

The usability study conducted with the novice users resulted in confirming two points: the first one regarding the reliability and the validity of the chosen evaluation methods, and the second one regarding the possible differences between expert and novice users. Both novice and expert usability studies showed the application was scored with similar scores regardless of the evaluation method. Therefore, it can be said that the chosen evaluation methods are proved to be consistent and could be safely used to evaluate the usability of an AR application.

As mentioned before in the previous sections, potential users of the MimAR are identified as architects. Therefore, conducting this experiment with architects who have certain expertise would have been a valid choice. However, this choice would have been resulted in excluding novice users (architecture students), who could have benefitted from this study as much as the expert users. In addition, it was assumed that the expert users' experience with digital media and CAD tools and their knowledge regarding some of the concepts and terms used in the development of the MAR application might give them an advantage in using the application over novice users. Furthermore, expert users could have found the application familiar and easy to use and evaluated the application's usability accordingly. Thus, results of the study might have been biased. Therefore, it was decided to conduct the experiment with participants who have different levels of experience to increase the credibility of the study and underline the possible differences between users by comparing expert and novice users' evaluations regarding the MAR application's usability.

According to Dünser and Billinghurst (2011) knowing the potential users, their expectations and understanding of a system is one of the key factors in developing and evaluating a system as smooth as possible. Because the differences between users' characteristics are directly related to the various ways they interact with those systems (Preece, Rogers, & Sharp, 2015). Even though identifying MimAR's potential users and their unique traits significantly helped in development process, developing a MAR application for more than one group of users with different characteristics such as varying levels of experience presented its own challenges. First of all, developing MimAR in a way that even users that have limited exposure to digital media and other 3D drawing and modelling tools could easily use, was a laborious undertaking. Secondly, because MimAR's user base consists of architects with varying experience degrees, these users' understanding of the MAR application and the way they interact with it also varies. Therefore, in order to cater to the needs of every user and be able to provide an acceptable user experience across the board, even though it took more time in the development phase, MimAR is developed by considering both novice and expert users' expectations.

The results of the usability study underline the difference between the novice and expert users' evaluations of the same system. The results of the comprehensibility sub-scale of the HARUS questionnaire shows that the novice users rated the application's comprehensibility considerably lower than the expert user. That means novice users are not adequately familiar with some of the concepts and terms used in MimAR as much as expert users.

It was assumed that conducting the study with only expert users might lead to biased results. The comparison of the results of expert and novice users showed that expert's HARUS (85,4 points) and SUS scores (87,5 points) are indeed much higher than the mean HARUS (62,7 points) and SUS (55,6 points) scores of novice users. Moreover, novice users' mean comprehensibility score (59,9 points) clearly demonstrates that understanding the application had more effect on the overall usability score of MimAR than being able to use it. Based on these results, it was demonstrated that experts' experience and familiarity give them an advantage in using the application over novice users. In addition, the results show that if the experiment were conducted with just expert users, the results of the usability study would have been much higher. Furthermore, it would not be possible to report that the novice users might encounter more problems than the experts and issues related to the comprehensibility of the application were needed to be addressed. These results also supports that the users' satisfaction with the AR applications are depended on the users' knowledge as Xue Sharma and Wild (2019) suggested. While the expert user's familiarity with the computer aided design tools and the IT concepts might have a role in this result, other variables that might affect this outcome should be further investigated within a larger population of expert users. Apart from one participant's scores and the comprehensibility problems, the results of the usability study conducted with the novice users showed that the usability of the application is scored close to the acceptable threshold of 68 points. In conclusion, the results of the novice study showed although the novice users enjoyed the overall experience, the comprehensibility of the application is needed to be increased for novice users to effectively use the application.

The results of the study showed that even though the usability of the application could be considered above the acceptable threshold, revising some aspects of the application could increase the overall comprehensibility of the application. Revising the graphical user interface in terms of readability and integrating a more understandable user feedback module that could provide clear on-screen text messages, were identified as the most useful improvements that could increase the usability of the application. We believe that the conclusions drawn from the results of this study would be beneficial for researchers working on the development of the MAR applications for design activities. Especially getting insight regarding reliable evaluation methods for an application based on empirical data and user feedback might prove invaluable in the development cycle of these kind of MAR applications.

Acknowledgement

This research is part of a PhD study by the first author at the Istanbul Technical University, Institute of Science and Technology, Department of Informatics, Architectural Design Computing Graduate Program. The authors would like to thank Dr. Sema Alaçam, Dr. Ethem Gürer and Dr. Elif Sezen Yağmur Kilimci for the access they provided for the experimental studies and the Class of Fall 2018-2019 Architectural Project 1 students from Istanbul Technical University Faculty of Architecture, Department of Architecture, for their participation in this study.

References

Akin, Ö., & Weinel, E. F. (1982). *Representation and architecture.* Information Dynamics Inc.

Aliakseyeu, D., Martens, J.B., & Rauterberg, M. (2006). A computer support tool for the early stages of architectural design. *Interacting with Computers, 18*(4), 528-555. https://doi.org/10.1016/j.intcom.2005.11.010

Azuma, R. (1997). A survey of augmented reality. *Presence: Teleoperators and Virtual Environments, 6*(4), 355-385. https://doi.org/10.1162/pres.1997.6.4.355

Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S., & MacIntyre, B. (2001). Recent advances in augmented reality. *IEEE Computer Graphics* *and Applications, 21*(6), 34-47. https://doi.org/10.1109/38.963459

Bai, H., Gao, L., El-Sana, J., & Billinghurst, M. (2013). Markerless 3D gesture-based interaction for handheld augmented reality interfaces. In *IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* (pp. 1-6). IEEE. https://doi.org/10.1109/ISMAR.2013.6671841

Bangor, A., Kortum, P., & Miller, J. (2009). Determining what individual SUS scores mean: Adding an adjective rating scale. *Journal of Usability Studies, 4*(3), 114-123.

Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An empirical evaluation of the system usability scale. *International Journal of Human–Computer Interaction, 24*(6), 574-594. https://doi.org/10.1080/10447310802205776

Boeykens, S., Santana Quintero, M., & Neuckermans, H. (2008). Improving architectural design analysis using 3D modeling and visualization techniques. In *Digital Heritage: Proceedings of the 14th International Conference on Virtual Systems and Multimedia* (pp. 67-73). Archeolingua.

Brooke, J. (1996). SUS-A quick and dirty usability scale. In P. W. Jordan, B. Thomas, I. L. McClelland, B. Weerdmeeester (Eds.), *Usability evaluation in industry* (pp. 4-7). CRC Press. https://doi.org/10.1201/9781498710411

Brooke, J. (2013). SUS: a retrospective. *Journal of Usability Studies, 8*(2), 29-40.

Caudell, T. P., & Mizell, D. W. (1992). Augmented reality: An application of heads-up display technology to manual manufacturing processes. In *Proceedings of the Twenty-Fifth Hawaii International Conference on System Sciences.* Kauai, HI, USA. https://doi.org/10.1109/HICSS.1992.183317

Chen, Z. R. (2007). How to improve creativity. In A. Dong, A. V. Moere, & G. S. Gero (Eds.), *Proceedings of Computer-Aided Architectural Design Futures (CAADFutures)* (pp. 571-583). Springer. https://doi.org/10.1007/978-1-4020-6528-6

Chin, J. P., Diehl, V. A., & Norman, K. L. (1988). Development of an instrument measuring user satisfaction of the human-computer interface. In J.J. O'Hare (Eds.), *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 213-218). Association for Computing Machinery. https://doi.org/10.1145/57167.57203

Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th ed.). Pearson.

Cross, N. (2011). *Design thinking: Understanding how designers think and work.* Berg.

Dünser, A., & Billinghurst, M. (2011). Evaluating augmented reality systems. In B. Furht (Ed.), *Handbook of augmented reality* (pp. 289-307). Springer. https://doi.org/10.1007/978-1-4614-0064-6_13

Ens, B., Quigley, A., Yeo, H. S., Irani, P., Piumsomboon, T., & Billinghurst, M. (2017, November 27-30). *Exploring mixed-scale gesture interaction* [Conference Poster]. SIGGRAPH Asia 2017 Posters, Bangkok, Thailand. https://doi.org/10.1145/3145690.3145740

Funk, M., Kritzler, M., & Michahelles, F. (2017, October 22-25). *HoloLens is more than air tap: Natural and intuitive interaction with holograms* [Conference Poster]. In IOT'17: Proceedings of the Seventh International Conference on the Internet of Things, Linz, Austria. https://doi.org/10.1145/3131542.3140267

Goel, V. (1995). *Sketches of thought.* MIT Press.

Goldschmidt, G. (1991). The dialectics of sketching. C*reativity Research Journal, 4*(2), 123-143. https://doi.org/10.1080/10400419109534381

Goldschmidt, G. (2004). Design representation: Private process, public image. In G. Goldschmidt, & W. L. Porter (Eds.), *Design representation* (pp. 203-217). Springer. https://doi.org/10.1007/978-1-85233-863-3_9

Gross, M. D., & Do, E. Y.-L. (1996). Ambiguous intentions: a paper-like interface for creative design. In D. Kurlander, M. Brown, & R. Rao (Eds.), *UIST'96: Proceedings of the 9th annual ACM symposium on User interface software and technology* (pp. 183-192). Association for Computing Machinery. https://doi.org/10.1145/237091.237119

Gruenefeld, U., Ennenga, D., Ali, A. E., Heuten, W., & Boll, S. (2017). EyeSee360: Designing a visualization technique for out-of-view objects in head-mounted augmented reality. In *SUI'17: Proceedings of the 5th Symposium on Spatial User Interaction* (pp.109-118). Association for Computing Machinery. https://doi.org/10.1145/3131277.3132175

Gül, L. F. (2008). Affordances of virtual environments: Do design media change the interaction with the design representation? In W. Nakapan, E. Mahaek, K. Teeraparbwong, & P. Nilkaew (Eds.), *13th International Conference on Computer Aided Architectural Design Research in Asia, CAADRIA 2008, Beyond Computer-Aided Design* (pp. 213-220). Pimniyom Press

Gül, L. F. (2017). Studying architectural massing strategies in co-design-mobile augmented reality tool versus 3D virtual world. In A. Fioravanti , S. Cursi, S. Elahmar, S. Gargaro, G. Loffreda, G. Novembri, A. Trento (Eds.), *ShoCK! - Sharing Computational Knowledge! - Proceedings of the 35th eCAADe Conference* (pp. 703-710). Faculty of Civil and Industrial Engineering, Sapienza University of Rome.

Gül, L. F. (2018). Studying gesture-based interaction on a mobile augmented reality application for co-design activity. *Journal on Multimodal User Interfaces, 12*(2), 109-124. https://doi.org/10.1007/s12193-017-0252-0

Hegde, S., Perla, R., Hebbalaguppe, R., & Hassan, E. (2016). GestAR: Real time gesture interaction for AR with egocentric view. In *2016 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)* (pp. 262-267). IEEE. https://doi.org/10.1109/ISMAR-Adjunct.2016.0090

Henrysson, A., Marshall, J., & Billinghurst, M. (2007). Experiments in 3D interaction for mobile phone AR. In *GRAPHITE'07: Proceedings of the 5th international conference on Computer graphics and interactive techniques in Australia and Southeast Asia* (pp. 187-194). Association for Computing Machinery. https://doi.org/10.1145/1321261.1321295

Hürst, W., & van Wezel, C. (2011). Multimodal interaction concepts for mobile augmented reality applica¬tions. In K. T. Lee, W. H. Tsai, H. Y. M. Liao, T. Chen, J. W. Hsieh, & C. C. Tseng (Eds.), *Advances in Multimedia Modeling: 17th International Multi¬media Modeling Conference MMM 2011* (pp. 157-167). Springer. https://doi.org/10.1007/978-3-642-17829-0_15

Hürst, W., & Van Wezel, C. (2013). Gesture-based interaction via finger tracking for mobile augmented reality. *Multimedia Tools and Applications, 62*(1), 233-258. https://doi.org/10.1007/s11042-011-0983-y

Ibrahim, R., & Rahimian, F. P. (2010). Comparison of CAD and manual sketching tools for teaching architectural design. *Automation in Construction, 19*(8), 978-987. https://doi.org/10.1016/j.autcon.2010.09.003

Kalay, Y. E. (2006). The impact of information technology on design methods, products and practices. *Design Studies, 27*(3), 357-380. https://doi.org/10.1016/j.destud.2005.11.001

Kirakowski, J., & Corbett, M. (1993). SUMI: the software usability measurement inventory. *British Journal of Educational Technology, 24*(3), 210-212. https://doi.org/10.1111/j.1467-8535.1993.tb00076.x

Knight, M., Dokonal, W., Brown, A., & Hannibal, C. (2005). Contemporary digital techniques in the early stages of design. In B. Martens, & A. Brown (Eds.), *Computer Aided Architectural Design Futures 2005* (pp. 165-174). Springer. https://doi.org/10.1007/1-4020-3698-1_15

Lewis, J. R. (1995). IBM computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use. *International Journal of Human–Computer Interaction, 7*(1), 57-78. https://doi.org/10.1080/10447319509526110

Lewis, J. R., & Sauro, J. (2018). Item benchmarks for the system usability scale. *Journal of Usability studies, 13*(3). 158-167

Lin, C.-Y. (2001). A digital procedure of building construction: a practical project. In K. Ny, P. Tom, J. Verbeke, & J. Verleye (Eds.), *AVOCAAD - Added value of computer aided architectural design* (pp. 459-468). Hogeschool voor Wetenschap en Kunst - Departement Architectuur Sint-Lucas.

Mitchell, W. J., & McCullough, M. (1995). *Digital design media* (2nd ed.). Wiley & Sons.

Nielsen, J. (2012, June 3). How many test users in a usability study? *Nielsen*

Norman Group. https://www.nngroup.com/articles/how-many-test-users/

Nielsen, J., & Landauer, T. K. (1993). A mathematical model of the finding of usability problems. In *CHI'93: Proceedings of the INTERACT'93 and CHI'93 Conference on Human Factors in Computing Systems* (pp. 206-213). Association for Computing Machinery. https://doi.org/10.1145/169059.169166

Preece, J., Rogers, Y., & Sharp, H. (2015). *Interaction design: Beyond human-computer interaction* (4th ed.). Wiley.

Rahimian, F. P., & Ibrahim, R. (2011). Impacts of VR 3D sketching on novice designers' spatial cognition in collaborative conceptual architectural design. *Design Studies, 32*(3), 255-291. https://doi.org/10.1016/j.destud.2010.10.003

Reffat, R. (2007). Revitalizing architectural design studio teaching using ICT: Reflections on practical im-plementations. *International Journal of Education and Development us¬ing ICT, 3*(1), 39-53.

Ryu, Y. S., & Smith-Jackson, T. L. (2006). Reliability and validity of the mobile phone usability ques¬tionnaire (MPUQ). J*ournal of Usability Studies, 2*(1), 39-53.

Sachse, P., Leinert, S., & Hacker, W. (2001). Designing with computer and sketches. *Swiss Journal of Psychology / Schweizerische Zeitschrift für Psychologie / Revue Suisse de Psychologie, 60*(2), 65-72. https://doi.org/10.1024/1421-0185.60.2.65

Salman, H., Laing, R., & Conniff, A. (2008). The changing role of CAAD in the architectural design studio. T*he Built and Human Environment Review, 1*, 25-39.

Santos, M., Ericson, C. , Polvi, J., Taketomi,T., Yamamoto, G., Sandor, C., & Kato, H. (2015). Toward standard usability questionnaires for handheld augmented reality. *IEEE Computer Graphics and Applications, 35*(5), 66-75. https://doi.org/10.1109/MCG.2015.94

Santos, M., Ericson, C., Taketomi, T., Sandor, C., Polvi, J., Yamamoto, G., & Kato, H. (2014). A usability scale for handheld augmented reality. In *VRST'14: Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology* (pp. 167-176), Association for Computing Machinery. https://doi.org/10.1145/2671015.2671019

Schon, D. A. (2008). *The reflective practitioner: How professionals think in action*. Hachette.

Shih, Y. T., Sher, W. D., & Taylor, M. (2017). Using suitable design media appropriately: Understanding how designers interact with sketching and CAD modelling in design processes. *Design Studies, 53,* 47-77. https://doi.org/10.1016/j.destud.2017.06.005

Simon, H. A. (1995). Problem forming, problem finding and problem solving in design. In A.Collen, & W. W. Gasparski (Eds.), *Design & Systems: General applications of methodology* (Volume 3 of Praxiology: International annual of practical philosophy and methodology series, pp. 245-257). Transaction Publishers.

Simon, H. A. (1996). T*he sciences of the artificial* (3rd ed.). MIT Press.

Stones, C., & Cassidy, T. (2007). Comparing synthesis strategies of novice graphic designers using digital and traditional design tools. *Design Studies, 28*(1), 59-72. https://doi.org/10.1016/j.destud.2006.09.001

Sutherland, I. E. (1968). A head-mounted three dimensional display. In *AFIPS'68 (Fall, Part I) Proceedings of the Fall Joint Computer Conference, part I* (pp. 757-764). Association for Computing Machinery. https://doi.org/10.1145/1476589.1476686

Suwa, M., Purcell, T., & Gero, J. S. (1998). Macroscopic analysis of design processes based on a scheme for coding designers' cognitive actions. D*esign Studies, 19*(4), 455-483. https://doi.org/10.1016/S0142-694X(98)00016-7

Tang, H. H., Lee, Y. Y., & Gero, J. S. (2011). Comparing collaborative co-located and distributed design processes in digital and traditional sketching environments: A protocol study using the function–behaviour–structurecoding scheme. *Design Studies, 32*(1),1-29. https://doi.org/10.1016/j.destud.2010.06.004

Wang, X. (2009). Augmented reality in architecture and design: Potentials and challenges for application. *International Journal of Architectural Computing, 7*(2), 309-326. https://doi.org/10.1260/147807709788921985

Won, P. H. (2001). The comparison

between visual thinking using computer and conventional media in the concept generation stages of design. *Automation in Construction, 10*(3), 319-325. https://doi.org/10.1016/S0926-5805(00)00048-0

Xue, H., Sharma, P., & Wild, F. (2019). User satisfaction in augmented reality-based training using Microsoft HoloLens. *Computers, 8*(1), 1-23. https://doi.org/10.3390/computers8010009

Zimmer, C., Bertram, M., Büntig, F., Drochtert, D., & Geiger, C. (2017, November 27-30). *Mobile augmented reality illustrations that entertain and inform: Design and implementation issues with the HoloLens* [Conference presentation abstract]. SA'17: SIGGRAPH Asia 2017 Mobile Graphics & Interactive Applications, Bangkok, Thailand. https://doi.org/10.1145/3132787.3132804