EXPLORING BENEFITS OF USING AUGMENTED REALITY FOR USABILITY TESTING

Choi, Young Mi; Mittal, Sanchit
Georgia Institute of Technology, United States of America

Abstract

This study explores the use of augmented reality (AR) in product usability testing compared to traditional methods. AR is being used in various fields, but this technology's use for usability testing has been very limited. This study specifically explored if it could be used for product usability testing. For this study, a product that is already in the market was used. First, the usability of the product was tested using traditional methods. Then the same product was modeled for augmented reality environment and subsequently a different set of users were asked to accomplish same tasks in AR environment. Same questionnaire was provided to all users for feedback. The feedback received using the two methods was compared. It was hypothesized that similar usability feedback could be obtained through AR as compared to traditional usability testing. The results showed that the feedback gathered using the two methods were the same, showing that evaluation of the AR representation of the product's usability is comparable to usability evaluation of the actual product.

Keywords: New product development, Augmented Reality, Usability Testing

Contact:
Dr. Young Mi Choi
Georgia Institute of Technology
School of Industrial Design
United States of America
christina.choi@gatech.edu

Please cite this paper as:
Surnames, Initials: *Title of paper*. In: Proceedings of the 20th International Conference on Engineering Design (ICED15), Vol. nn: Title of Volume, Milan, Italy, 27.-30.07.2015

1 INTRODUCTION

Augmented Reality (AR) refers to a view of real or physical world in which certain elements of the environment are computer generated. These virtual elements could be a modification of a current element in the real world or could be an entirely new element. Augmented reality technology is being used in a variety of industries including product development and healthcare. This paper will present the current state of the art elated to augmented reality in various fields and discuss how it can be used in improving usability testing process.

Every year companies devote significant resources to the development of new products (Donahue, 1999; Ehrlich, 1994). Companies like Google and Microsoft spend 5-6 billion US dollars every year on research and development (Krantz, 2012). A new product development (NPD) cycle involves five steps (Exploration and screening, Business analysis, Product development and Testing, Commercialization). Product development and0 Testing uses 54 % of the resources (R. G. Cooper, 1988). These products could be interactive or non-interactive. Before bring a product to market, companies develop the entire product with all interactions integrated into the product and then release models for usability testing. If they find some glaring issues during testing, they have to revisit the development process. Once a product is in development, it costs 10 times more to fix a problem than during design and once the product has been released it costs 100 times more to fix a problem than in design (Donahue, 1999). The cost to fix design defects during design might be lowered if the cost of building an actual product can be removed and reliable usability testing can be done on a virtual product. In this case if there are issues with the usability of a product, all that needs to be done is changing the virtual rather than a physical model so that another round of testing can be performed.

Most studies discuss techniques in which augmented reality can be used in product development, but they do not talk about using this technique for gathering user input. The purpose of this study is to explore if augmented reality can be beneficial in the usability testing process to gather similar usability input as compared to traditional method of usability testing where a more fully finished physical version of a product is built and tested. This research might provide results that would provide evidence that AR can be used for usability testing and help product developers understand the benefits of using AR in product design.

Augmented reality works as an assistive technology where it helps the user to visualize objects in a real Physical environment. Early studies in the field of AR include the research done at The Media Laboratory at Massachusetts Institute of Technology (Starner, 1997). The researchers developed text based AR where text could be overlaid onto a physical object in a real environment which keeps changing as user interacts with different objects in the real world. The applications of this technology included detecting existing real objects and overlaying them with 2D AR graphics which provided instructions on how to interact with that object. This was achieved using wearable computers which are currently known as head mounted devices (HMD) for AR. HMD is a device that displays content in front of your eyes and is worn on your head which can be part of a helmet. HMDs work in a fashion that they enhance the real world environment instead of replacing real environments (Rolland, Holloway, & Fuchs, 1994).

Another emerging application of AR is to use this technology in product development process. Once a product is developed, it need to be tested. This means if some problems are found with the product, the developers have to go back to development board and go through the development process again which requires extra time and money (Griffin, 1993). Pedro and colleagues (Santos, Graf, Fleisch, & Stork, 2003) suggest that AR should be used in early stages of product development so that they can be modified or redesigned interactively. They developed a method of using AR in their design process with a set up that included a tracking system, a head mounted device (HMD), PC-based graphical station and tangible objects for interaction. The process is explained by an example of car design in virtual environment where a 3D model of a car is overlaid on a physical car model and then identifying the areas to modify by the use of markers. A designer can then use hand gestures to redesign elements of the car.

Designers now have the ability to use tangible augmented reality during the product development process. Tangible augmented reality combines visualization properties of AR with tangible interactions of physical objects (Billinghurst, Kato, & Poupyrev, 2008). Lee and Kim (G. A. Lee & Kim, 2009) talk about how developers and designers can create better target content using tangible augmented reality, which is, being able to do usability interactions before being used or felt by the target audience. The implementation process included making markers on cards which were placed in front of user. These markers are detected by an AR device like HMD and a 3d virtual object is placed on the markers. Users can then interact with the cards by moving them around in space as they would normally do with a physical object in real environment to interact with virtual objects. Such a tool will be ideal for communication between designers and programmers.

A similar tool is described by Shen and colleagues (Shen, Ong, & Nee, 2010) which gave the ability to designers to interact with a virtual panel overlaid on real surfaces as conventional button systems like a keypad or keyboard. As the designer interacts with these buttons they perform different functions. There were two types of interaction cursors used, a physical stylus (could also be a finger) and a virtual stylus. They use occlusion based interaction where if the tip of the stylus is in the area of the button, it interacts with the button. Lee and colleagues (G. A. Lee, Billinghurst, & Kim, 2004) conducted experiments based on occlusion based interactive methods with tangible AR and found that this kind of interaction enhanced the natural way of interacting with physical objects. The users were able to get instant visual feedback while they interacted with virtual buttons.

Lee and Park (W. Lee & Park, 2005) also describe that rapid prototyping techniques have some limitations such as being expensive and laborious, not easily modifiable and the difficulty of representing material properties such as color and textures.

Augmented reality is still a developing technology, but it has reached a point where it might be employed as a useful tool in many aspects of product development. The goal of this study is to evaluate the validity of employing AR for usability testing. It will compare the results of usability tests collected from AR representations of a product with usability test results collected from the fully functional, physical product. The AR representation of the product is an analogue to a digital prototype which may be available early in the development process where the physical version is an analogue to a functional prototype which would be available much later. By mocking up an existing product in AR we compare the same product, only represented differently. In order to be a potentially useful tool on which to base early design decisions, the usability results between the AR and physical versions of the product should show no significant differences.

2 METHOD

The Sony Walkman NWZ-E463 (Electronics, 2013) was used for usability testing. It was chosen for its combination of tangible buttons and touch screen interface. It was also chosen as it is a relatively less popular/commonly used product (compared to more common players available on Apple iOS or Google Android based devices, for example) but it was also easily available. As it was less commonly used, it was much easier to select participants who had never specifically encountered the device before.

Figure 1. Sony Walkman NWZ-E463

Before testing, the Walkman was first modeled for AR. A 3D model of the product and interface was recreated in Solidworks (Figure 2). After modeling, an augmented reality application called Layar (Layar, 2013) was used to place the virtual model in the real environment. Layar is an IOS application which lets you view 3D models and HTML in augmented reality view. Marker detection technology was used to detect the marker (Figure 3) and replace marker with the digital 3D model in the real environment when viewed through an iPad. While the users viewed the virtual model through an iPad, they had the ability to interact with the object through augmented reality app on the iPad. Screen graphics matching the product interface were created in Adobe Photoshop. The interactivity was added by creating screen interactions in adobe Edge Animate. The Layar then overlaid HTML on top of the 3D model allowing users to look at the 3D model and also interact with it.

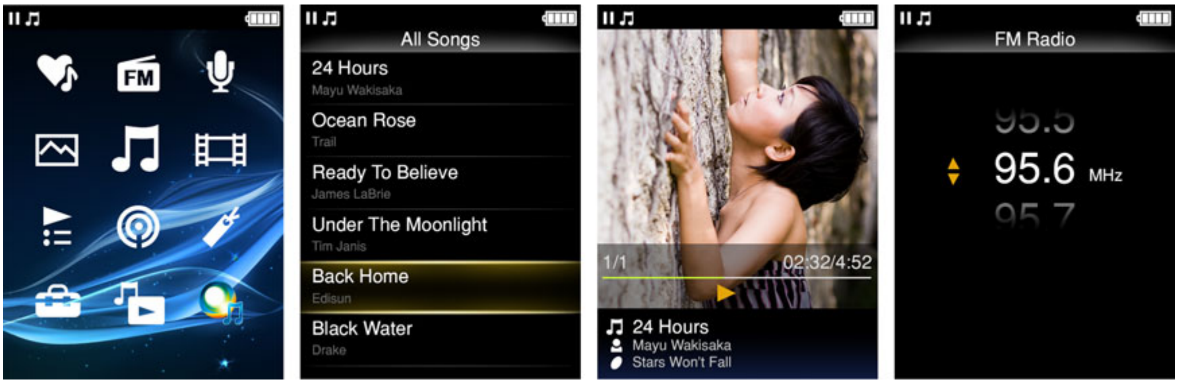

Figure 2.Four of many screens created in Adobe Photoshop

Figure 3. Graphic that was used as marker

A physical model with the same basic size and weight of the Sony Walkman was 3D printed (Figure 4). This gave users something tangible to hold in their hand while interacting with the virtual product. The graphic that was used as marker was pasted on to the 3D printed model so that as users moved it around they could see the virtual product move around as well.

Figure 4. Front, side and back view of 3D printed model with the marker pasted on it.

Figure 5. The Layar Application detecting the marker graphic and placing the 3D model and HTML/Java Script interaction on top of 3D model.

Figure 5 shows the final AR representation of the Sony Walkman viewed through the Layar application. The Layar application was active on the iPad so that the user could look at the virtual product through the tablet's screen. A user would then interact with the Walkman interface by touching the interface elements on the iPad.

Usability testing required users to perform four specific tasks with one of the representations of the product. Users were told that none of the tasks were timed. The tasks were:

- Task 1: The user was asked to find a playlist named 'Running' and play a song of their choice from that playlist. The user was then asked to return back to the home screen.
- Task 2: The user was told to find a video named 'Fifa World Cup', play that video and then return back to the home screen once the video had finished playing.

- Task 3: The user was asked to find a genre named ‘Rock’, play any song in the list, and then asked to return back to the home screen.
- Task 4: The user was told to tune into FM frequency 99.7 and then return to the home screen.

Usability Testing was done by 60 users. 20 of these users used the actual product for testing, 20 users used the 3D printed model with a marker card pasted on it and 20 users used just the marker card to view AR interactions (Figure 6). Each user did the testing independently.

Figure 6. Usability testing using actual product (Non-AR, Left), 3D printed model with a marker card pasted on it (AR1, Center) and marker card (AR2, Right).

Users testing the physical product were simply instructed to hold the device in their hand and perform each of the 4 tasks. Users testing the AR representations of the Walkman were instructed to hold the marker in front of the iPad and to perform the tasks by interacting with the Walkman interface via the iPad touch screen. The iPad was mounted on a stand during the study as shown in Figure 7.

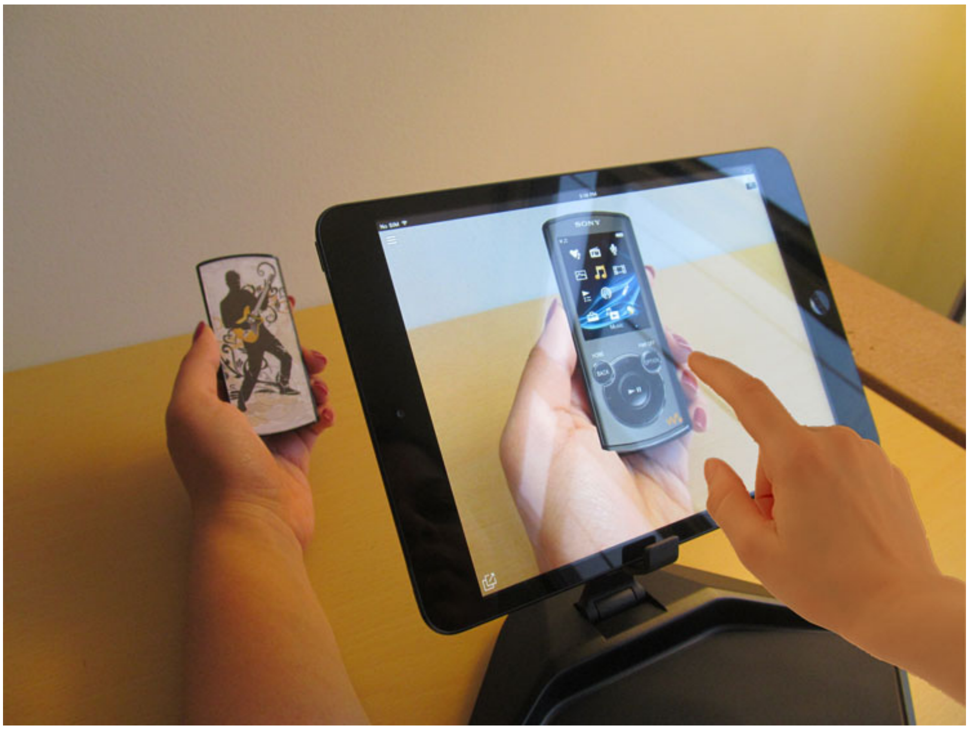

Figure 7. Usability testing using AR method.

After completing all of the tasks, each user provided usability feedback by completing the, USE Questionnaire (Lund, 2008-09a, 2008-09b). The same questionnaire was provided to all the users to compare results from all three groups.

3 RESULTS

Three sets of groupings were made, one for the 20 users who tested the actual product, second for the 20 users using augmented reality for usability testing which use a marker card pasted on a 3D printed model and third for 20 users that used augmented reality for usability testing just using the marker card. Responses for each question were grouped separately. Grouping the responses for each question provided with the frequency of each response and gave the answers users were tending towards.

Kruskal-Wallis tests (Kruskal, 1952; Lowry, 1999-2013) performed to evaluate significant differences between the groups. This is a non-parametric equivalent of a one way analysis of variance and most appropriate for the ordinal nature of the data and relatively small sample size. The responses to each question from each of the three groups: AR1 (marker only), AR2 (marker on 3D printed model) and Non-AR (the actual product) were compared. No significant differences were found between them.

The 30 questions on the USE questionnaire are divided into four categories: (1) usefulness (eight questions), (2) ease of use (eleven questions), (3) ease of learning (four questions) and (4) satisfaction (seven questions) (Lund, 2008-09b). The average score for each category were calculated for each user. The resulting scores for usefulness, ease of use, ease of learning and satisfaction were each compared to respective scores from the other groups using Kruskal Wallis. Again, no significant differences were found between any of the groups.

4 DISCUSSION AND CONCLUSION

The data from this study indicate that there were no differences usability testing performed via AR and testing performed with a functional product. This study has potentially important results that can be applied to the development and testing of electronic products. Early in the design process (before there is any actual product) it is common to seek input based on product concepts or mockups. In these cases the designer is interested in knowing what the user thinks about the idea as if it were in its finished form because this is what the user would a user would actually (eventually) use. In order to make informed design decisions, it is important to know that the usability collected on an early concept will reliably predict that of the finished product. If it does not, this can lead to poor design decisions based on early data and bigger problems later.

These results indicate that an AR mockup of a product can generate usability test results that are equivalent to a fully functional product. A new product development (NPD) cycle involves five steps (Exploration and screening, Business analysis, Product development and Testing, Commercialization). Product Development and testing could be further divided in to the following phases(Network, 2000):

1. Using an Iterative, Cyclical Design Process
2. Envisioning Phase
3. Planning Phase
4. Developing Phase
5. Stabilizing Phase
6. Preparing for the Next Version

The last two stages could in particular benefit from AR. The Stabilizing phase refers to the point in time the when development ends and small issues are fixed to create a final product for shipment. The usability testing in this phase is done to fine-tune the final product. This is a big step forward coming from the development phase. If there are major issues in this phase then it would cost the company 10 times more to fix the issue than during the development stage. The companies could use the AR product in the stabilizing phase rather than using the actual product. That way, if they find some problems, major or minor, could be fixed by just tweaking the 3D digital model and HTML interactions.

In the last stage companies prepare for the next version of the product where the same development and testing process is repeated. This process could really benefit from the AR technology if the next version is a mere update or a minor adjustment to the original product. The developers could tweak the AR model according to the needs of the newer version and test it again.

The early design process may also benefit. AR can provide a way of evaluating many different ideas or features in the early conceptual phase of design. A product idea may be fully conceived in a digital environment and reliably tested while avoiding the time and expense required to generate a functional prototype. This can allow many different ideas to be tested and better decisions to be made early in development when larger changes are possible.

While the study produced successful results, it had some limitations. The results of this study potentially only apply to products with interactive screens and controls. The described AR method might not be a good approach for form development or for products which do not have interactive screen elements. However, further study would be necessary to confirm the usefulness of AR for other types of products. Also, using an iPad to view augmented reality interactions, somewhat limits the size of the product that can be tested. For example, it would not be a good method for testing something like a vending machine.

The users recruited for testing in this study were all college students, which is not a diverse demographic. This study might produce different results with different participants.

Performing usability tasks with a portable music player might be relatively easy for users who regularly use products like these. They have an inherent knowledge how these type of devices work. While, for users who have never or rarely used a portable music player, it might more difficult to complete the task. This would mean that the data gathered is either incorrect or insufficient. Conversely, if a user has extensive experience with other portable music players, he would compare the product he is testing with the ones used before. In such instances if the product being testing is being compared with an excellent product, the usability feedback would always be negative. While, for someone who has never used a music player, the product being tested might be the best product ever. Eliminating these types of biases is very difficult.

The described AR method uses touch screen of an iPad to perform interaction, which might be a hindrance. Users do not get to know how the product will feel in their hands while using the product.

In next steps for this research, rather than using the touch screen of an iPad to perform the interactions, it would be much better if the users were able to use the 3D printed model itself to do the interactions and use the iPad simply to view those interactions. This could be achieved by making the user wear colored markers on their fingers and have software that has the ability to detect those markers and react to the positions of the markers. In this study, interactions with the AR representation and actual product were basically the same: touch screen and buttons. There were no interface elements that could not be represented through the touch interface. If the product required grasping and turning a knob for example, there would be no way for the user to actually do this through a touch screen. The way of interaction would be much different. The ability to actually interact with a physical model would provide more natural interaction and open the possibility of evaluating the use of AR for testing/evaluating a wider array of products.

REFERENCES

Billinghurst, M., Kato, H., & Poupyrev, I. (2008). Tangible augmented reality. *ACM SIGGRAPH ASIA*, 1-10.

Donahue, G. M., Weinschenk, S., & Nowicki, J. (1999). Usability is good business.

Ehrlich, K., & Rohn, J. (1994). Cost justification of usability engineering: A vendor's perspective. *Cost-justifying usability*, 73-110.

Electronics, S. (2013). NWZ-E463 Walkman® MP3 player - Discontinued Portable MP3 Players & Docks Sony Store - Sony US. Retrieved October 10, 2013, from http://store.sony.com/nwz-e463-walkman-mp3-player-zid27-NWZE463PNK/cat-27-catid-EOL-Portable-MP3-Players-Docks%20%20

Griffin, A. (1993). Metrics for measuring product development cycle time. *Journal of Product Innovation Management, 10*(2), 112-125.

Krantz, M. (2012). Microsoft, Intel, Google outspend Apple on R&D. Retrieved December 14, 2013, from http://usatoday30.usatoday.com/money/perfi/columnist/krantz/story/2012-03-20/apple-marketing--research-and-development-spending/53673126/1

Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American statistical Association, 47(260)*, 583-621.

Layar. (2013). Home | Augmented Reality | Interactive Print | Layar. Retrieved September 30, 2013, from https://www.layar.com/

Lee, G. A., Billinghurst, M., & Kim, G. J. (2004). *Occlusion based interaction methods for tangible augmented reality environments.* Paper presented at the Proceedings of the 2004 ACM SIGGRAPH international conference on Virtual Reality continuum and its applications in industry.

Lee, G. A., & Kim, G. J. (2009). Immersive authoring of tangible augmented reality content: a user study. *Journal of Visual Languages & Computing, 20*(2), 61-79.

Lee, W., & Park, J. (2005). *Augmented foam: a tangible augmented reality for product design.* Paper presented at the Mixed and Augmented Reality, 2005. Proceedings. Fourth IEEE and ACM International Symposium on.

Lowry, R. (1999-2013). Kruskal-Wallis Test. *VassarStats: Website for Statistical Computation.* Retrieved February 10, 2014, from http://vassarstats.net/textbook/ch14a.html

Lund, A. A. M. (2008-09a). Measuring usability with the USE questionnaire. *Usability interface, 8*(2).

Lund, A. A. M. (2008-09b). Questionnaire for User Interface Satisfaction. Retrieved October 10, 2013, from http://hcibib.org/perlman/question.cgi?form=USE

Medcalc. (1993-2014). Chi-squared distribution table. Retrieved February 20, 2014, from http://www.medcalc.org/manual/chi-square-table.php

Network, M. D. (2000). Usability in the Development Process. Retrieved January 16, 2014, from http://msdn.microsoft.com/en-us/library/ms997575.aspx

R. G. Cooper, E. J. K. (1988). Resource Allocation in the New Product Process. *Industrial Marketing Management, 17*, 249-262.

Rolland, J. P., Holloway, R. L., & Fuchs, H. (1994). *A comparison of optical and video see-through head-mounted displays.* Paper presented at the SPIE Telemanipulator and Telepresence Technologies.

Santos, P., Graf, H., Fleisch, T., & Stork, A. (2003). *3D Interactive Augmented Reality in Early Stages of Product Design.* Paper presented at the HCI International 2003, 10th Conference on Human-Computer Interaction.

Shen, Y., Ong, S. K., & Nee, A. Y. C. (2010). Augmented reality for collaborative product design and development. *Design Studies, 31*(2), 118-145. doi: http://dx.doi.org/10.1016/j.destud.2009.11.001

Starner, T. S. (1997). Augmented reality through wearable computing. *Presence: Teleoperators & Virtual Environments, 6*(4), 386.