Vlachos, A., Economides, A.A., & Perifanou, M. (2022). Smart glass application testing for augmented reality. In: ***Proceedings of EDULEARN 2022, 14th International Conference on Education and New Learning Technologies***, pp. 9569-9577. Palma de Mallorca, Spain, 4-6 July. IATED. doi: 10.21125/edulearn.2022.2310

SMART GLASS APPLICATION TESTING FOR AUGMENTED REALITY

A. Vlachos[1], A. Economides[2], M. Perifanou[2]

[1]University of the Aegean (GREECE)
[2]University of Macedonia (GREECE)

Abstract

A multitude of mobile device applications are being designed every day, a significant amount of which are Augmented Reality applications. Current smartphones and tablets can use such applications to supersede multi-modal information, such as text, images or predesigned 3D digital objects, over the camera's real-world input. This technology offers the users immersive educational, industrial, touristic and even research experiences. The main drawback of such applications is that the users have to hold their device towards their view point and look at the screen, which limits their mobility. Smart glasses are believed to be the next big thing in wearable technology, since they can offer an always-on augmented experience, without such constraints. The users are able to move and act in any way they wish, with the glasses working on certain cues, such as location beacons, image or object recognition and others. These devices are already seeing some use in industry, while pilot projects and experiments with them are also being performed in other fields. Smart glasses are expected to improve significantly and become more accessible to the public in the following few years, as technological advancements in computing and wearable technologies progress. Microsoft has already introduced the HoloLens, which is not yet a consumer device but is eventually aiming to be, while Apple is also working on a consumer headset that is expected to be complete and available in the next few years. In this paper we review the methodology used in testing mobile applications before publication, from the designer's point of view, both for Android and iOS. We then use this current methodology and initially apply it to an Augmented Reality application for smartphones. Following that, we test it on the modified version of the same application for smart glasses. That includes testing both the hardware aspects, such as power efficiency testing, GPS accuracy testing etc, and the software aspects, for example usability, readability, high contrast options or colourblind mode etc. Our goal is to determine whether this set of guidelines can work for both cases and devices, since the original criteria suggested were only with handheld devices in mind and for touch focused applications, where the user has to look at his/her smartphone and all content is generated through it. We then seek to determine what, if any, extra criteria or new steps would be necessary for testing applications for Smart Glasses and offer our own set of criteria recommendations on the subject. Smart glass applications are, in our opinion, the next big thing in educational mobile applications, be it in virtual educational tours or 3D modelled reconstructed environments or even used as testbeds, for example in experimenting with cultural heritage artifacts without physical contact. Our testing and suggestions can benefit educators, stakeholders and designers by offering a better understanding of the challenges such applications can present, and change the way they are designed and tested for publication, in preparation for the wider use of Smart glasses.

Keywords: Application Testing, Augmented Reality, Metaverse, Smart Glasses, Epson Moverio

1 INTRODUCTION

Augmented Reality (AR) applications are becoming more common every day. They can be designed and applied in many domains, from gaming and tourism, to education and industry. Such applications can offer multiple advantages over typical mobile applications, with access to multimodal information overlaid over the object they are viewing being the main benefit. Users can also interact with the objects in a non-destructive, touchless manner, in some cases allowing further experimentation. The main issue with current AR applications is the medium, which in this case is the user's smartphone or tablet. To view the desired digital overlay, users have to look at the screen of the device and, through it and the camera, the world around them, with the digital information seen on screen. Understandably, this takes away from the experience and immersion as it constantly reminds the users to maintain their device in a specific place and orientation and thus limiting their mobility and their view into a narrow field.

This is where smart glasses come in, to alleviate these issues and enhance the users' immersion and overall experience. AR smart glass (ARSG) applications are being designed as pilot projects, due to the low availability of smart glasses. This is expected to change in the next few years, with many companies looking to create consumer devices. Microsoft already has the HoloLens, which is quite impressive and, though it remains a developer and researcher-oriented device, it eventually aims to be available to a wider public. Facebook, now rebranded as Meta, has entered the Virtual Reality (VR) field with the Oculus Rift and Quest head mounted displays (HMD) and is also looking to the AR field through their MetaVerse initiative. Apple is also rumoured to be working on their own ARSG consumer headset.

The coming of so many devices will most likely be followed by a great influx of ARSG applications in their respective app stores. Many are already being developed as pilot projects, while most existing AR applications can probably be modified for smart glasses. There is, however, no set of guidelines to test such applications on these new devices. The question that arises is whether the current mobile application testing guidelines are enough or do the ARSG applications require a modified or new set of their own. To answer that, we used an application developed as a pilot project [1] and tested the smartphone version according to a set of guidelines. We then used the same set to test the ARSG version of the application on a smart glass device, and attempted to suggest modifications to better fit these upcoming devices.

2 SMART GLASSES

Smart glasses are believed to be the next big innovation in the field of wearable devices. They have many advantages over smartphones and tablets in AR use cases. They require much less focus on the user's part, since they are worn and any user input is through a simple touch controller. They offer much higher immersion, by not limiting the user's view in the specific boundaries of his/her device's screen. Finally, they do not limit the user's movement, rather encouraging him/her to move more, by not offering a zoom mechanic, motivating the user to physically walk closer to an object for a better look.

Since most smart glasses work when being paired with a smartphone, they do share some disadvantages, the main being location approximation and power consumption. While GPS sensors on smartphones can pinpoint a user's location and orientation, they have an accuracy of approximately 10 meters. Many application designers have experimented with ways around this by attempting image and feature based tracking, as well as Wi-Fi beacons and a multitude of sensors, for example gyroscopes, accelerometers, proximity sensors. Some smart glasses use the built-in camera or their own sensors, to alleviate this issue and offer more accurate user tracking. The other issue, power consumption, depends on device use and application demands. AR applications with very detailed 3D models will require much more power from the displaying device, whether it is a smartphone or glass. With glasses having, most of the time, very small capacity batteries, this can shorten experiences significantly in the case the smart glass does the processing of the application itself. In the case where the paired smartphone does the heavy lifting and feeds information to the smart glass, this issue is not as prominent. While both issues do remain, we expect them to be less pronounced in coming years, as technological advances can offset or eliminate them entirely.

At the same time, smart glasses lack an important feature of smartphones, the touch imitating interaction. In an AR application on a smart phone, the user experiences the illusion of touching an object to interact with it, while in the case of the glasses all interaction takes place via the controller. This fact already creates a design guideline for ARSG applications, that they have to be very easy to use, with simple interactions, to avoid user mistakes and fatigue [2].

With this advent of smart glasses on the horizon, some experiments on their use are already being performed. Applications are also being designed and tested, in an experimental manner, as smart glasses are not yet widely available. [3] presented the Rapid Application Development methodology for creating a complete AR application, in the context of cultural heritage tourism. [2] presented a pilot ARSG tour guide for outdoors cultural heritage sites. [1] was a pilot project for educational tourism in a cultural heritage context. [4] examined how well users respond to ARSG applications in a visit to a UK art gallery. Others look for ways to bypass the simplistic design guideline mentioned above, which could very well be possible through voice commands or eye tracking [5]. Overall, ARSG applications and their shortcoming and benefits are actively being tested.

3 MOBILE APP TESTING

Mobile applications testing is a very large field, with many types of tests possible. To name but a few, there are functionality, performance, interrupt, usability, installation, certification tests and many others. The types of testing we performed and will be referencing below are as follows:

- Outdated software: the application is tested on older smartphone operating systems. Many devices are not always up to date or cannot be updated, therefore testing on older versions is required.
- Installation: tests if the application can be installed, updated and uninstalled without any special actions on the user's part.
- Performance: the application is tested in conditions such as low battery mode, low memory or heavy processor load due to many applications working in the background.
- Functional: tests if the application works as intended, if all outlined functions are available to the users and looks for user interface (UI) problems.
- Usability: tests if the application is user friendly and that the experience is equally good across various devices and operating systems. This testing requires user participation and feedback.
- Interrupt: tests how the application responds when an interrupt, such as a phone call, SMS, or notification, occurs and whether the application inhibits such operations in any way or the interrupts cause abnormal application behaviour, such as crashes.

To effectively test an application, developers use specifically created software, such as Appium (https://appium.io), to put the app through a series of tests. There is also the option of using emulators and simulators of the various mobile operating systems. Real device clouds, on platforms such as Browserstack, are also available, where one can load his/her app on an actual device and test it. The automated options, however fast and easy, cannot effectively judge the usability or accessibility of an application as well as a human can, which is why manual testing is still relevant today. Smart glasses are not common devices, therefore emulators and simulators, as well as real device clouds, were not used for our tests. We instead performed manual testing on both versions of the application.

4 METHODOLOGY

CARAT is a pilot project and application designed for both smartphones and smart glasses. It aims to offer a touristic education experience to users in cities, where the history behind street names is revealed in multimodal form and any interconnected material, such as a famous person's residence or a museum concerning a specific era, are made available through points of interest (POIs) and dynamic tours. The application was designed in two versions, one that works only on smartphones, both iOS and Android, and one specifically for smart glasses, that only requires limited information from the paired smartphone, such as location.

We tested the CARAT mobile application on many devices, while the ARSG version on one device. The mobile application was tested on Xiaomi Redmi Note 4X (Snapdragon version) running Android 7, a Samsung Galaxy Tab S7 running Android 10 and later upgraded to 11, an iPhone 7 with iOS 14 and an iPhone 12 with iOS 14 and updated to 15, and finally an iPad Pro 2021 running iPad OS 15. All devices were 4G/5G capable, with the exception of the Xiaomi smartphone and the iPhone 7, that only work on 4G. The ARSG application was tested on the smart glasses available to the project for testing, the Epson Moverio BT-350. They come with a touch pad controller that connects to the smartphone and the smart glasses, and acts as the link between them.

We performed installation, performance, functionality and interrupt testing on all devices in turn. We need to note that while we also performed usability testing, in part, and outdated software testing, these were only possible on the smartphone version, as there was no such software for the smart glasses, therefore we will only mention them in passing.

Our flowchart for each testing process can be seen below in Fig. 1,2,3 and 4. All steps were performed in sequence, meaning we did not, for example, turn off background apps before loading the more processor and memory intensive ones during performance testing, except for the battery load check at the end of that testing phase. Between every step we tested whether the application and all its functions were working properly or if it exhibited any abnormal behavior.

Installation testing, as seen in Fig. 1, was performed by loading the appropriate application on the device via USB cable, then proceeded with installation. We then executed the application and gave all necessary permissions for each operating system. We tried updating the application, if possible, and then uninstalled it.

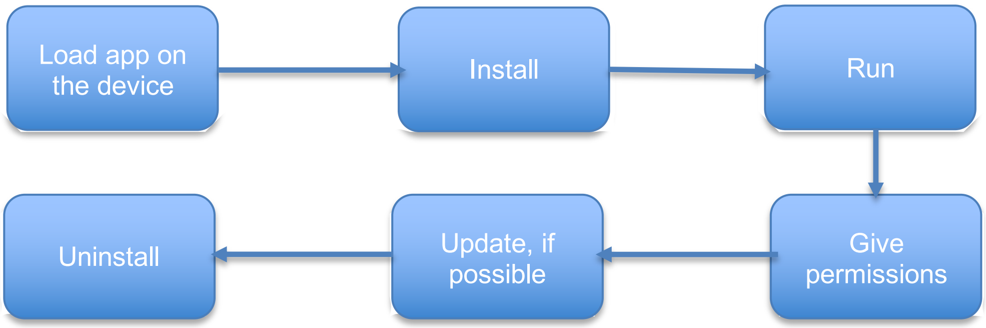

Figure 1. Installation testing.

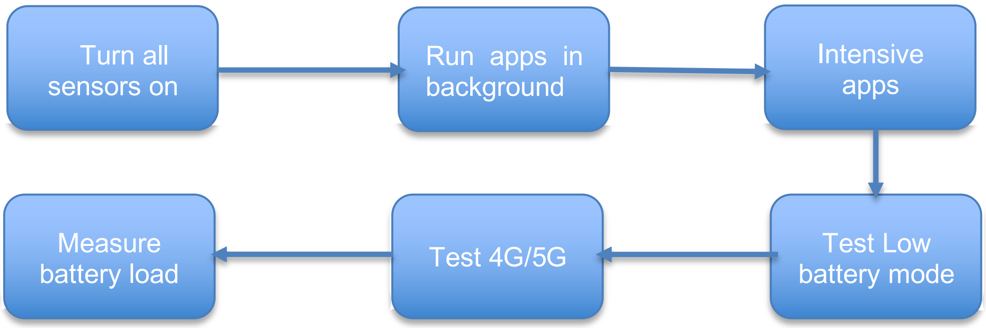

Figure 2. Performance testing.

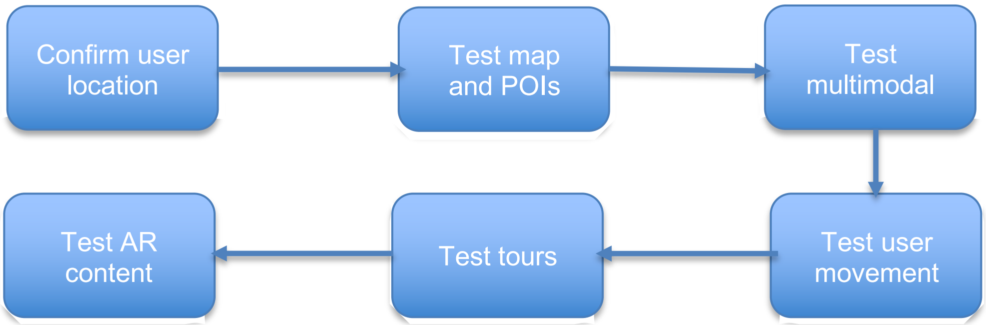

Figure 3. Functional testing.

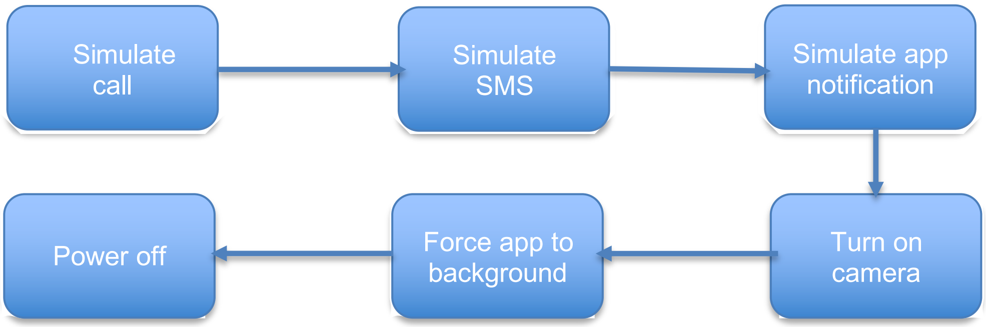

Figure 4. Interrupt testing.

Performance testing comprised of the steps in Fig. 2. We initially turned all sensors on for each device, which meant Bluetooth paired devices such as headphones, smart watches and even Airdrop devices all attempted to connect to the device being tested. We then started apps that are usually in the background, such as email clients, messaging applications, note taking applications and any one randomly selected. This was followed by simultaneously running memory and processor intensive apps, such as photo editors, AR content, video editors and players. We then observed application behavior in low battery mode and during a switch from 4G networks to 5G and vice versa. Finally, after turning off all other apps, we measured the battery load after continuous test app usage for 20 minutes.

Functional testing, seen in Fig 3., was performed in the city, beginning with confirming the user location and orientation. We then interacted with various POIs on the map, as well as moved the map to look at more distant POIs and locations. We randomly interacted with some of them and examined the multimodal information presentation, such as text, video and photos. We then moved to a different location, both by car and on foot, and examined how fast and correct the application was in both location and orientation. The tours were tested next, on accuracy in context, on location and navigation. In the end, the AR content was tested.

Interrupt testing took place last, as seen in Fig. 4. We initiated a phone call to the device running the application from another smartphone. Then we sent three SMS messages to the device, one separately and two in quick succession. Following that, we turned on the camera with gestures or buttons, depending on the device. After this test, we forced the test app to the background, again by swiping or pressing the home button. Finally, since none of these devices have removable batteries, we attempted to simulate a sudden power off by running the test app on less than 2% remaining battery and then examine whether it loads properly on the next power on.

5 TESTING RESULTS

We will look at the results of the tests mentioned above in the same sequence, presenting the results for the smart glasses at the end of each test section.

5.1 Installation testing

Installation testing was successful in almost all devices. The Xiaomi phone, while it properly installed and given all permissions, would not load the map in the application. We tried many various settings but could not get it to work. After discussing it with some colleagues, we found that the version of this phone with the MediaTek processor could run it properly. We attribute this to a possibly different GPS sensor or other hardware element on our device, or perhaps even a faulty one. We therefore omitted the Xiaomi phone from any further testing.

Since the CARAT application is a pilot project and not available in app stores, we had to install it via USB as mentioned, the exception being the iOS and iPad OS devices that could install it via TestFlight. Testing the update process was, therefore, only available on those devices, where we tested it separately at a later date. The only other difference in installation testing was in the way permissions were given, in each operating system.

On the smart glasses, the process was similar to the smartphones. We had to connect it via USB and install the app from inside the custom Android environment. We cannot say whether the process would be as simple in the case of the app store, since Epson has closed the ARSG store available to Moverio devices. It seems to have been similar to the Google Play Store, so it would possibly be just as easy. Uninstalling was the same as on any smart device.

5.2 Performance testing

Performance testing was successful on all devices, with the weaker hardware-wise iPhone 7 exhibiting slower response time when having many memory and processor intensive apps running. Battery load was approximately even in percentages across all devices, where after 20 minutes of continuous app usage, the battery indication was 1-2% less. This variation was expected since not all devices have the same capacity batteries and some were not brand new. Overall, this confirms the application was properly designed to utilize system resources as needed.

On the smart glasses, this testing phase was different, as they run the application but receive their location information from the connected smartphone. We therefore had to test the ARSG in a more limited manner, by running the application on the smart glass while trying processor and memory intensive tasks on the connected smartphone. These apps on the smartphone did not have an impact on the glass. Since the app selection on the smart glass was also limited, we could not run many apps to test their behaviour. With the ones available, the device did exhibit slower response times overall, as well as in the apps currently running. This is, we believe, a hardware issue, since the device has fairly weak specifications, compared to a recent smartphone. We also tested if running apps that use the GPS sensor on the smartphone, such as Google Maps, would delay the location update on the smart glass, which they did not. 4G/5G and low battery mode were not tested on the glasses as they did not have that functionality. The overall battery usage was measured to be approximately 5% during a 20 min continuous usage but cannot be compared to another smart glass device.

5.3 Functional testing

Functional testing was overall successful on all devices. There were some omissions in the initial versions of the application, such as a “back” or “home” button that could help the user return to the main screen quickly. Another problem was the orientation of the device, in that the application was designed for portrait mode and when the device was turned to landscape, the app would remain in portrait. The one important problem we discovered was that when the user was in an area with multiple POIs close together and walked close to them, the location indicator would be effectively hidden by the POI markers. All these and other minor issues we found, were fixed in the latter versions of the application. AR content was displayed properly, though not entirely on the real-world object. In the case of a sculpture, the AR content covered it entirely, though it missed some small corners and was therefore slightly skewed to one direction. This was possibly due to the photogrammetry performed and accuracy of the process.

We did, in part, perform a usability test simultaneously, since having this many devices allowed us to see whether the experience was similar across all platforms. We all seem to prefer the iOS – iPad OS look and feel of the app, since the image appeared sharper, but that is subjective. Other than this preference, there was no other difference across the various devices and operating systems. We also tested the app with some of the accessibility options on the devices. Voice over did not work properly, but that was expected since the app is in Greek. Text zooming, colour filters, touch controls and others all worked properly within the app.

The ARSG version of the app worked properly, though selecting POIs with the touch controller was difficult at first. They also exhibited some slow response times, when loading the larger files of some of the POIs photos. This, again, is probably a hardware issue. In some cases, the multimodal information could take up the entire viewing area, e.g., a large picture or text. In devices that use only a small area of the glass, usually the lower part, this is not an issue. It can be a serious problem in devices that use the entire area, if the displayed information is too dark and therefore hinders the users’ vision. AR content that is placed incorrectly, such as a statue or sculpture that is erroneously placed 1 meter to the left of its proper location, could also significantly hamper user vision and perhaps even endanger him.

5.4 Interrupt testing

The Interrupt testing was performed last. The call function is not available in the tablets, though we simulated it in the iPad by connecting it to the same Wi-Fi as the iPhone and answering the call on the iPad. All devices were 4G/5G capable, so could and did receive the SMS messages. The app performed properly in most cases of interrupts. On calls, it retreated in the background and once the call ended was again in the foreground. We also tested moving between calls, which cause a very slight, less than 1 second, delay in updating the map. Effectively, this appears as if the user is at the location the call started but is then almost instantly updated to the new location. This could be graphics or memory speed related, as it seems the user location is updated live when moving while on a call, with the app on the foreground. SMS messages and other app notifications did not affect app usage, even when multiple notifications or messages appeared simultaneously. Power loss had the same effect in every device, that is the application had to be restarted and reconnect to the GPS sensor to get location data.

On the ARSG version, we again had to test the smartphone under these circumstances rather than the glasses, since they have no 4G/5G capability, nor any preinstalled apps that could possibly display notifications. We did get some due to low battery, but were non invasive and did not affect the overall app performance. Other than that, calls, SMS messages and notifications on the connected smartphone did not impact the device or app running on it in any way. Power loss affected the ARSG app the same way it affected the smartphone version.

5.5 Other testing

As mentioned, we did perform usability testing in part, by having a multitude of devices and operating systems available. Apart from the slight preference for the visuals in Apple devices, there was no real difference in any versions of the application.

At the same time, we did test the application for outdated software, on iOS 14 and 15, iPad OS 14 and 15, Android 10 and 11. It never worked for us on Android 7, probably due to the hardware issue with the Xiaomi phone. We did not have any devices with Android 8 or 9, nor any iOS version before 14 to test further.

6 DISCUSSION

From all the testing phases above, we can discern some issues with testing on ARSG.

The first is that performance testing and interrupt testing were, in our smart glass model, not entirely possible on the device itself. We rather had to test if such conditions on the connected smartphone impacted the smart glass. We believe this to be a purely technological issue. If future smart glass models can function independently, these tests then become possible. 5G connectivity on such a device is quite likely, seeing how wearables such as smart watches already possess eSIM capabilities. These watches are tethered to a phone, but 4G/5G capable ones can function independently, and are therefore available for interrupt testing, in the same way any eSIM device is. Performance testing would also become possible. In our case, our smart glass model was fairly weak, hardware-wise. With processors becoming smaller, more powerful and energy efficient every year, it seems possible that consumer grade models would be just as capable as smart phones. A non-tethered smart glass would need to perform all processing on-board, therefore all current performance testing becomes available.

The second issue that bears discussion is about controlling the app. The touch pad was indeed difficult, at first, and sometimes slow to respond to user input. This can be frustrating for the end users and discourage them from further using an app. Voice control could remedy this situation, though that is when language barriers might become an issue. We cannot recall any current smartphone app that is available in every language, and seeing how that has not yet happened with text-based apps, it could take even longer with audio. There is also the option of using eye tracker technology, that lets the user select options based on looking at them. As with most new technologies, ease of use could significantly impact overall acceptance. As an example, the iPod Classic, Apple’s phenomenally successful MP3 player, was one of the easiest devices to use, as it only incorporated a mechanical scroll wheel and one main button, with four others around it. The Moverio model we tested did have a similar pad, though it included a touch surface, which was difficult to use and had the response issues. Overall, we feel there needs to be an extra testing guideline for usability, dealing with how easy to use the app is. This could branch into different testing types, for touch pads, voice activation and eye tracking, but it seems necessary for any ARSG app. The testing steps of this phase could test for various age groups, mobility and response speed on user input and of that input to the glass.

A test that seems unnecessary for AR on smartphones but could be vital in ARSG, is AR placement. From the tests we performed, it was clear that functional testing guidelines were created with a smart phone in mind, where blocking the users’ vision does not matter, as they still have peripheral vision and can always just move the phone. As said previously, a large AR object that is not placed properly could endanger the ARSG user by obstructing their vision. This could also occur if the object is placed properly but the orientation of the user detected by the application is incorrect, whether that is done via image and feature tracking or is sensor based. We suggest a placement accuracy phase, where the location - orientation detection aspect of the app and the proper placement of AR components is tested thoroughly.

The same could be said for the viewing area and what and how much of that area is devoted to any ARSG app. In the case of a smart glass where the entire area is used by the apps, any very bright, dark or opaque element could hinder the user. As mentioned previously, one of the main advantages of smart glasses is that the users can see the world around them directly, without having to look at a screen. Blocking off that view works against the technology’s strengths. An opacity testing could be important, and could be combined with the placement accuracy testing, perhaps into an overall safety testing category.

One thing we feel we should mention is app availability. Similar to every case of using outdated hardware, our experience with a practically non-existent app store was disheartening. This was understandable, due to our smart glass model being older and now discontinued, but we do believe that app stores and preinstalled apps for all necessary functions should be included on all models, especially in consumer grade products. Therefore, this could create another testing phase, app - software availability. It could be included in outdated software testing or be an entirely new phase, but it seems needed so that smart glasses do not end up like the multitudes of older and now unusable smartphones that are, effectively, e-waste.

In Table 1 below, we summarize our suggestions for testing on AR for smartphones and ARSG, both tethered and independent.

Table 1. Caption for the table.

	Smartphone AR	Tethered ARSG	Independent ARSG
Installation	Yes	Yes	Yes
Outdated Software	Yes	Yes	Yes
Performance	Yes	-	Yes
Functional	Yes	Yes	Yes
Usability	Yes	Yes	Yes
Interrupt	Yes	-	Yes
AR Placement	-	Yes	Yes
Viewing Area	-	Yes	Yes
Opacity	-	Yes	Yes

7 CONCLUSIONS

The tests performed in both types of devices were not exhaustive. We did try to test for as many categories we believe to be useful in the context of this paper. As seen above, some testing phases could be directly adapted to smart glasses, if not now, probably in the future, with the advent of consumer grade devices. We believe that performance and interrupt testing will be, in the near future, just as available to smart glasses as they currently are for smart phones and, in part, smart watches.

Other testing phases need to be modifier or added. Smart glass models, at this time, have larger frames than regular glasses, since they need to include all their hardware in the frame. This can slightly reduce the users' field of view. We also need to keep in mind that, while a smartphone's screen can go all black and cause only mild annoyance, the same thing happening to an ARSG app that uses the entire viewing area, can cause extreme discomfort, if not endanger the user.

With the above in mind, we do suggest adding a safety testing category specifically for the smart glasses. This should test for accurate orientation and location detection, correct AR element placement and opacity below a set limit.

Future research could look at how high or low opacity should be limited at and whether users should be able to override that option. There are also many other application testing phases that would be interesting to try on smart glasses, such as load testing, and others that will be necessary once these devices see wider use, such as security and certification testing. If the Oculus Quest, which requires a Facebook account to use, is any indication of the direction the Metaverse AR could go, then security testing should move to the top of this list, to protect the early adopters.

ACKNOWLEDGEMENTS

This study was partially supported by the project History in Cities: Augmented Reality tools and applications (CARAT) under award number T1ΕΔΚ-04136, ΕΣΠΑ 2014-2020.

REFERENCES

[1] A. Vlachos, A. A. Economides, M. Perifanou, and T. Chatzidaki, “CARAT: History in Cities Street Names,” *Press*.

[2] E. Litvak and T. Kuflik, “Enhancing cultural heritage outdoor experience with augmented-reality smart glasses,” *Pers. Ubiquitous Comput.*, vol. 24, no. 6, pp. 873–886, Dec. 2020, doi: 10/ghgm35.

[3] W. K. Obeidy, H. Arshad, and J. Y. Huang, “TouristicAR: A Smart Glass Augmented Reality Application for UNESCO World Heritage Sites in Malaysia,” vol. 10, no. 3, p. 8.

[4] D.-I. D. Han, M. C. Tom Dieck, and T. Jung, “Augmented Reality Smart Glasses (ARSG) visitor adoption in cultural tourism,” *Leis. Stud.*, vol. 38, no. 5, pp. 618–633, Sep. 2019, doi: 10/gh64ms.

[5] M. Mokatren, T. Kuflik, and I. Shimshoni, “Exploring the potential of a mobile eye tracker as an intuitive indoor pointing device: A case study in cultural heritage,” *Future Gener. Comput. Syst.*, vol. 81, pp. 528–541, Apr. 2018, doi: 10/gf59r2.