Article

User Satisfaction in Augmented Reality-based Training using Microsoft HoloLens

Hui Xue [1,*], Puneet Sharma [1] and Fridolin Wild [2]

[1] UiT The Arctic University of Norway; hui.xue@uit.no; puneet.sharma@uit.no
[2] Oxford Brookes University, United Kingdom; wild@brookes.ac.uk
* Correspondence: UiT The Arctic University of Norway; Tel.: +47-776-46-625

Abstract: With the recent developments in augmented reality (AR) technologies comes an increased interest in the use of smart glasses for hands-on training. Whether this interest is turned into market success or not depends at the least on whether the interaction with smart AR glasses satisfies users, an aspect of AR use that so far has received little attention. With this contribution, we seek to change this. The objective of the article, therefore, is to investigate user satisfaction in AR applied to three cases of practical use. User satisfaction of AR can be broken down into satisfaction with the interaction and satisfaction with the delivery device. A total of 142 participants from three different industrial sectors contributed to this study, namely, aeronautics, medicine, and astronautics. In our analysis, we investigated the influence of different factors, such as age, gender, level of education, level of Internet knowledge, and the roles of the participants in the different sectors. Even though users were not familiar with the smart glasses, results show that general computer knowledge has a positive effect on user satisfaction. Further analysis using two-factor interactions shows that there is no significant interaction between the different factors and user satisfaction. The results of the study affirm that the questionnaires developed for user satisfaction of smart glasses and the AR application performed well, but leave room for improvement.

Keywords: augmented reality; Microsoft HoloLens; AR application; user experience; user satisfaction

1. Introduction

Augmented Reality (AR) means enhancing the user's perception "with additional, artificially generated sensory input to create a new experience including, but not restricted to, enhancing human vision by combining natural with digital offers" (Wild et al., 2018). Augmented Reality typically has three characteristics [1]: first, AR combines the virtual with the real world; second, objects are registered from both the real and virtual world in one coordinate system; third, the interaction between the objects of both worlds is possible in real time.

Hands-on training is important for many disciplines and professions, such as medical workers, mechanics, technicians, electricians, engineers, sailors, pilots, and firefighters. In the past decade, AR has been increasingly employed for a number of training applications, such as medical education [2], rehabilitation engineering [3], automotive safety [4], task assistance [5], and manufacturing [6].

For the successful adoption of AR-based training across different domains, one of the key factors is user satisfaction. User satisfaction is defined as a combination of different factors associated with the usage of the AR application and the associated delivery device[7]. These factors include: a feeling of powerfulness and achievement; an efficient use of time, effort, and other resources; meaningful content; a better insight to the training environment; a natural interaction; a feeling of amazement; performance that exceeds expectations; playfulness; the invoking of positive feelings and pleasing memories; immersion and engagement; a transparent interaction; the feeling of participation in a community; a sense of privacy of the user's content; inspiration, encouragement, and motivation; and, finally, artistic creativity [7].

The rest of this paper is organized as follows. First, we turn to the state of the art, summarizing what the research has found so far with respect to AR user interaction, AR user satisfaction, and questionnaires used for evaluating user satisfaction. Next, the AR app used in the trials is described. Subsequently presented are the research methodology and a summary of the information of the participants, devices, design of trial tasks, and evaluation methods. Finally, findings and results are illustrated, and the discussion and conclusion are given at the end.

The main objective of this study is to test and observe user satisfaction in using AR applications and using AR glasses. The method for evaluating includes questionnaires and interviews. The AR app used in this evaluation, therefore, has two parts: one is the expert recording the experience in the workplace, and the other part is the novices training on work-related procedures using said recordings. In this study, we evaluated the following research hypotheses: to find if experts and students are satisfied with the prototype application, to see if the application can increase interest in learning new skills, and to evaluate if the users find the application easy to use.

2. State of the art

2.1. AR user interaction

AR technologies provide a different user experience than that of, for example, mobile phone apps. The user interacts with the surrounding real world, combining inputs from the environment with digital augmentations. Popular examples include PokemonGO and SnapChat. These type of apps certainly brought the term "augmented reality" into the spotlight [8]. With the advent of consumer-grade AR glasses, different types of AR user interactions are becoming necessary. For example, a user who is wearing Microsoft's HoloLens can communicate diagrams and other types of graphics directly embedded into the environment to a different, remote user (see Figure 1).

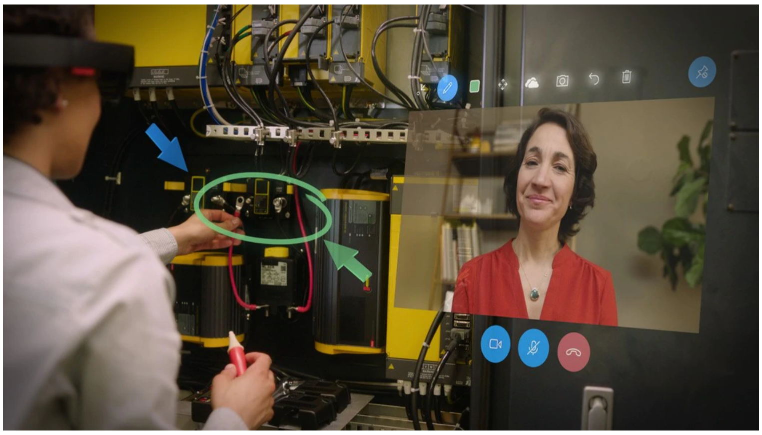


Figure 1. With Microsoft HoloLens, a user connects the wires with remote assist [9].

2.2. AR user satisfaction and questionnaires for evaluating user satisfaction

AR Technology has evolved from offline to online, from static devices to mobile devices, and from desktop and mobile to wearable devices [10]. Consequently, with AR development over the past decade or so, special attention has been drawn to the maximization of AR user satisfaction. AR user satisfaction is dependent on both the design of the user interface (UI) and the choice of the AR hardware. Personalization of AR glasses can lead to greater AR user satisfaction [11]. AR apps designed for a good user experience result in a more overall satisfied AR user. This applies to AR navigation apps, AR health apps, AR education apps, and certain AR smart glasses games [12].

There are several concepts and subjective measures for evaluating the user experience of AR services. With regards to the user, satisfaction questionnaires are common tools used to evaluate a user's experience. One such tool—the Questionnaire for User Interaction Satisfaction (QUIS)—is

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 November 2018 doi:10.20944/preprints201810.0594.v2

Peer-reviewed version available at Computers 2019, 8, 9; doi:10.3390/computers8010009

Figure 2. User Interface of the recording mode. Image from the WEKIT consortium in 2017.

designed to assess users' subjective satisfaction with specific aspects of the human–computer interface [13]. The results of QUIS facilitate new developments by addressing reliability and validity problems found using its satisfaction measurements. Therefore, the measure is highly reliable across many types of interfaces.

QUIS consists of a demographic questionnaire, a six-scale measure of overall system satisfaction, and hierarchically organized measures. The measures include the following specific interface factors [13]: screen factors, terminology and system feedback, learning factors, system capabilities, technical manuals, online tutorials, multimedia, teleconferencing, and software installation. Each area is measured by a 7-point scale according to the user's overall satisfaction with the interface and the above factors [13].

3. The AR application

The AR application consists of two modes: recorder and player.

The *recorder* is designed for capturing an expert's workplace experience and combining it with technical documentation associated with a given scenario. The player is used to reenact the scenario to verify the recordings and usually employed to train a novice for the scenario.

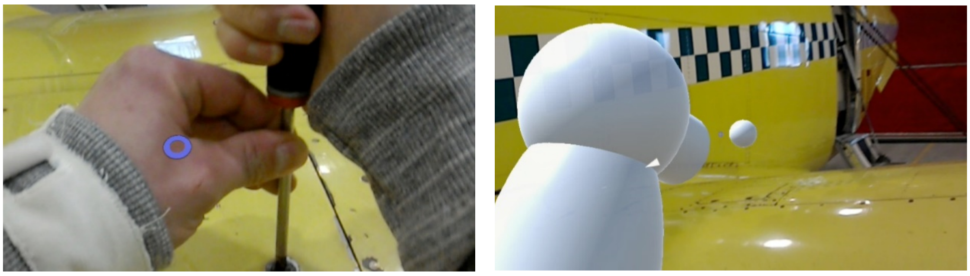
In order to capture an expert's experience, a set of transfer mechanisms were defined in [15]. The so-called transfer mechanisms allow us to map the key aspects of an expert's performance to low level data and subsequent sensors. For more details on the different sensor components and their integration please see [16]. The recorder (as shown in Figure 2) consists of a radial menu that allows us to select different options for capturing diverse annotations such as: pictures, videos, text annotations (for adding text information to different objects in the environment), audio, ghost hands (to capture the locations and movements of user's hands) and 3D models (useful for performing the task).

Trainers can use a so-called 'ghost track' to record their own position and indoor movement, while at the same time recording voice explanations. When replaying such recording to the trainees, the holographic 'ghost' representation of the expert provides more intuitive guidance on where to be, where to focus, and what to do – than merely reading about the task to be learned in a manual using text and illustration. Figure 3 shows an example of such ghost track recording and replay for an aircraft maintenance task.

The *player* is the mode designed for trainees to learn how to do procedural operations (kind of 'do-torial' mode). The app executes AR learning experience models (IEEE standard association, working group p1589), so allows to load different learning and training activities. Activities can

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 November 2018 doi:10.20944/preprints201810.0594.v2

Peer-reviewed version available at Computers 2019, 8, 9; doi:10.3390/computers8010009



(a) (b)

Figure 3. A ghost track in WEKIT Recorder mode: (a) recording a ghost track; (b) ghost track replay. Image from [19].

be transferred from device to device as well as from place to place, using a calibration marker to recalculate the relative positions of all points of interest, while utilising 3D environmental mapping to provide stable projections.

The WEKIT player starting screen is shown in Figure 4. Once the task starts, the first action step and its associated augmentations are shown on the smart glasses display. From the perspective of the users, this typically means that the visual annotations overlay onto their unimpeded realworld view (optical see-through). Step by step, they guide the user through the learning task at hand. Gesture commands, voice commands, and the hardware clicker are all available when using the app. Figure 5 shows an example of the WEKIT player in action. When the sensors on the HoloLens detect the particular tangible object, the virtual button is displayed in front of the trainee, while instruction on handling and movement are given at the same time.

Figure 4. Starting screen in WEKIT Player mode. Image from [17].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 November 2018 doi:10.20944/preprints201810.0594.v2

Peer-reviewed version available at Computers 2019, 8, 9; doi:10.3390/computers8010009

Figure 5. Example of user interface of WEKIT Player mode. Image from [17].

4. Research Design/Experiment Methodology

4.1. Participants

In order to evaluate the satisfaction of the user interaction and the smart glasses user experience, the WEKIT application was deployed in three different pilot testing scenarios: aviation, medical imaging, and space. Moreover, in the experiments, the test population was divided into two main groups, experts and students, respectively. A total of 47 experts (8 females; 39 males) with a high level of technological competency in their respective fields were recruited. A total of 95 learners (23 females; 72 males) from the three different fields voluntarily participated in the trials. The majority of the participants (68) were in the 18–24 age group, followed by 48 of the participants in the range between 25 and 34. Most of the participants had moderate or better computer knowledge and internet knowledge, expressed on a 5-point Likert scale ranging from very poor, poor, moderate, good, to very good. All participants gave written consent for their participation in the trials.

4.2. Material and Apparatus

The trial used the Microsoft HoloLens as wearable AR glasses for assessing the user's satisfaction with AR training. There are two parts in the WEKIT technology platform [18] deployed on HoloLens. One is a recorder for capturing expert experience and the other one is a player for presenting the expert's experience to the trainees. During the trial, all interactions and manipulations were done by using gesture and voice command only.

4.3. Trial design/task

The trial used Microsoft HoloLens as wearable AR glasses for assessing the user's satisfaction with AR training. There are two parts in the WEKIT technology platform [18] deployed on HoloLens. One is a recorder for capturing expert experience and the other one is a player for presenting the expert's experience to the trainees. During the trial, all interactions and manipulations were done by using gesture and voice command only.

4.4. Trial design/task

The trial tasks were separated into three different areas, as mentioned in section 4.1. Tasks in the Aeronautics use case were performed at Lufttransport, Norway. The scenario used for the aeronautics use case was a pre-flight inspection consisting of checking and securing different items such as baggage, exits, locks, and checking the status of components such as landing gears, brakes, engine switches, battery, and fuel. The experts comprised of maintenance apprentices, skilled workers (mechanics), and technicians working on base maintenance at Lufttransport. The novice group comprised of student

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 November 2018 doi:10.20944/preprints201810.0594.v2

Peer-reviewed version available at Computers 2019, 8, 9; doi:10.3390/computers8010009

volunteers from UiT The Arctic University of Norway [19]. Figure 6 shows a novice engaging in the pre-flight inspection task. Experts had been using the different types of annotations to create the required instruction for the training procedure, which then was provided to the trainee in the player mode of the AR app. The novice followed the instructions in order to complete the task in the cockpit.

The pre-flight inspection scenario consisted of following steps (see Table 1)

Table 1. Steps of the pre flight inspection scenario for Beechcraft B200.[19]

No.	Cabin/cockpit	Action	Content
1.	Baggage	Secure	Ensure that the baggage compartment and net is secured.
2.	Emergency Exit	Secure and unlocked	Emergency exit handle must be in the secured position and the lock must be in the unlocked position.
3.	Control locks	Remove and stowed	The control locks must be removed and stowed.
4.	Trim Tabs Exit	Set to "0"	Including elevator trim tab, aileron trim tab, elevator trim tab.
5.	Condition levers	Fuel cut-off	Must be set to the fuel cut-off position.
6.	Landing gear control	Down	Must be in down position.
7.	Parking brake	Set	If required, ensure that the parking brake is set on.
8.	Ignition and engine start switches	Ensure off	Must be in the off position.
9.	Battery	Check for minimum 23V	Turn on the battery master switch. Check for minimum 23V on the voltmeters by pushing the push-to-test knobs on the voltmeters.
10.	Fuel quantity	Check	Check the fuel quantity in main fuel tanks. Move and hold the “fuel quantity”-switch to auxiliary position and check the fuel quantity in auxiliary fuel tanks.

Figure 6. Maintenance Engineer in the cockpit of a Beechcraft B200 King Air model. Image: Mikhail Fominykh, WEKIT consortium (2017).

The medical task involved imaging and diagnostic workers and was conducted at EBIT (Esaote’s Healthcare IT Company) in Genoa, Italy [20]. This task was for training medical students and radiologist apprentices on using MyLab8, an ultrasound machine produced by ESAOTE [21]. Similar to the trial at Lufttransport, the users executed the steps of the procedure using the player mode of the application. The scenario for the medical use case was to perform a particular ultrasound

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 November 2018 doi:10.20944/preprints201810.0594.v2

Peer-reviewed version available at Computers 2019, 8, 9; doi:10.3390/computers8010009

examination to analyse a patient's condition. The patient was a paid actor. During the task, the novice doctors needed to combine data from different sources in order to arrive at the correct diagnosis. As for the holographic training instructiton, the guidance was set up for the player mode again using experts, adding the step by step description needed to guide the trainee through the full scanning and assessment procedure. The novice doctors then wore the HoloLenses and tired to perform the examination. The tasks are shown in the Table 2. In Figure 7, we can see a novice performing a task by positioning the probe in the target direction and taking measurements using the player application.

Table 2. Steps of diagnostic training of radiology students performing an ultrasound examination.[21]

No.	Ultrasound equipment	Action	Content
1.	Probe	Choose	Choose the proper probe. Point to the linear probe and listen to the audio annotation explaining which probe to select and why, and how to hold it (with a raised edge).
2.	Button	Select the mode	Point to the "B/M" button to select the correct mode.
3.	Probe	Transversal position	Position the probe in a transverse direction
4.	Probe	Longitudinal direction	Position the probe in a longitudinal direction.
5.	Button	Change the mode	Change the mode to Color Mode.
6.	Button	Choose button	Position the center line in the middle of the artery.
7.	Button	Change the mode	Change the mode to Doppler Mode. If required, ensure that the parking brake is set on.
8.	Circle button	Pointing	Point to Circle button highlighted in the following figure.
9.	Measure button	Measure	Choose correct button to start reading.
10.	Trackball button	Measure	Position the cursor over the highest peak in the curve, then click the left trackball button to set the first data point. Repeat for the lowest point in the graph.
11.	Image button	Snapshot	Take a snapshot with the measure.

Figure 7. A radiologist conducting ultrasound training. Image: WEKIT consortium (2017).

The space task that was conducted at the facilities of ALTEC in Turin and it involved training astronauts on how to install a Temporary Stowage Rack (TSR). The TSR installation is a procedure that astronauts have to perform on the International Space Station (ISS) [22]. Similar to the trials at the other two organizations, experts designed the training scenario, while a larger number of trainees then executed the scenario on the player application. The evaluation of the expert's experience was

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 November 2018 doi:10.20944/preprints201810.0594.v2

Peer-reviewed version available at Computers 2019, 8, 9; doi:10.3390/computers8010009

conducted using the recorder mode of the app as well as the player, while the trainees used only the player mode. The steps for this procedure are as follow. First, trainees were asked to scan the working area to create the 3D model of the environment, to then identify the six seat track studs location on the structure, the position of the ball bearing and brackets. Next, they were instructed to fix the six studs in specific locations. Finally, they were asked to extract TSR, deploy it, and fix it to the correct places. The novices performed the task based on the recorded content. In Figure 8, we can see a participant of the trials performing a task in a replica module of the International Space Station.

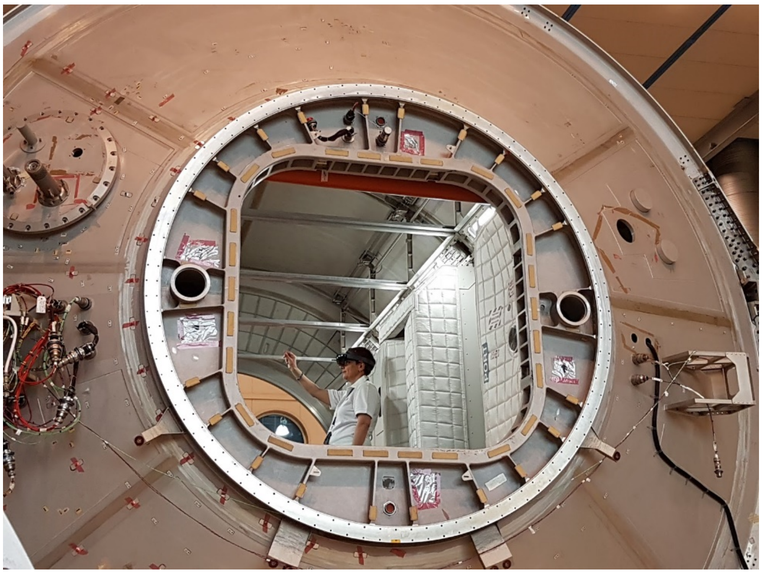


Figure 8. Astronaut trainer in a replica training module of the international space station. Image: WEKIT consortium (2017).

4.5. Smart Glasses User Satisfaction (SGUS)

The Smart Glasses User Satisfaction (SGUS) questionnaire was created for the WEKIT trials. It is a tool designed to assess users' subjective satisfaction with smart glasses. SGUS is a method and measure to scrutinize aspects, such as an enhanced perception of the environment, interaction with the augmented environment, implications of location and object awareness, the user-created AR content, and the new AR features that users typically use [7]. The general objective of the questionnaire is to understand the potential end users' central expectations of AR services with smart glasses, especially from an experiential point of view [7]. In this study, the smart glasses used for the different use cases were Microsoft HoloLens. SGUS measures subjective satisfaction on the basis of different features associated with user satisfaction, such as the content and interaction with the content. SGUS is based on evaluation criteria for web-based learning [14] and statements evaluating the user experience of mobile augmented reality services [7]. SGUS consists of 11 items (statements) on a 7-point Likert scale (1–7) [19]. The 11 statements include three categories of evaluation criteria, which are general interface usability criteria, AR interaction-specific criteria for an educational AR app, and learner-centered effective learning [14].

4.6. Questionnaire for User Interface Satisfaction (QUIS)

The Questionnaire for User Interaction Satisfaction (QUIS) measures subjective satisfaction with specific aspects of the interface and interaction between the user and the AR application [23]. In this study, QUIS was modified for AR glasses, i.e., HoloLens. Hence, a questionnaire with 15 items was used. In order to maintain consistency with the survey in other sections, each item was mapped to a numeric value of 1–7 instead of the 9-point scale [23].

Peer-reviewed version available at *Computers* **2019**, *8*, 9; doi:10.3390/computers8010009

4.7. Procedure

As most participants had no experience with AR glasses, at the beginning of the trial, they were asked to familiarize themselves with the AR glasses, i.e., HoloLens. In order to do this, gesture training with HoloLens was done before they started using the application. The application comprised a scenario that the participants had to complete in a particular use case setting. The content of the application was generated by experts in that specific use. After the participants completed all the tasks, they were provided with the QUIS and SGUS questionnaires to complete.

5. Results/Findings

5.1. Descriptive statistics

In this section, we report on descriptive statistics for the smart glasses user interaction and the interaction satisfaction. We organise the findings alongside the investigation of eight hypotheses, with the summary of these shown in Table 3.

*H1: Does gender matter? In Science and Engineering, gender is not balanced and there are fewer women than men[**?**]. Gender stereotypes can affect use of established technologies. We therefore investigate whether the influence on user satisfaction of these new media will be moderated by gender.*

H2: Does age matter? Studies imply that younger people embrace new technologies more easily [24]. Since we are using AR glasses and applications for training, we would like to know whether age affects user satisfaction.

H3: Are experts more tech savvy? It is likely that experts have more experience with technology applications, as in general they also have more domain-specific knowledge and skills. We assume that they would be more able to grasp the app concept, thus being more satisfied with the interaction. The novices, however, may have less knowledge and skills, hence, may find the app difficult to use.

H4: Does education matter? Higher levels of education go hand in hand with higher levels of ICT skills. It is justified to hypothesize that the educational level predicts satisfaction.

H5: Does computer knowledge matter? Higher levels of ICT and media skills typically involve transfer skills. The AR smart glasses headset used, Microsoft Hololens, is a stand-alone device. We need basic computer knowledge to use it. Those with better computer knowledge might find it easy to use, and hence, give a higher score in terms of user satisfaction.

H6: Does internet knowledge matter? In analogy to computer skills, one can expect Internet skills to influence the user satisfaction levels in a positive manner.

H7: Are there differences in satisfaction levels between the participants of the three test-beds? The trials involved three different learning tasks, in three different environments, with three different groups of participants. As all three trials are about training a particular procedure, that there are no differences identified across test-beds.

H8: Is there any interaction between the above mentioned factors?

Table 3. Summary of the hypotheses

#	Description	Expectation
H1	Gender	Men are more satisfied with the user interaction than women.
H2	Age	Younger participants give a higher score.
H3	Experts vs. novices	Experts have higher satisfaction levels.
H4	Education level	Higher education users have higher satisfaction levels.
H5	Computer knowledge level	Users with better computer knowledge might be more satisfied.
H6	Internet knowledge level	Might have influence.
H7	Three different test-beds	Might have different results.
H8	above 7 factors	There might be interactions between factors.

5.1.1. SGUS

As mentioned before, SGUS has 11 items. The summation of the score for the 11 items is the SGUS score. As shown in Table 4, we provide data such as: n (number of participants), mean, standard deviation, minimum value, Q1 (the first quartile: "middle" value in the first half of the rank-ordered data set), median, Q3 (the third quartile: "middle" value in the second half of the rank-ordered data set), and maximum value for the following variables: gender, education level, roles, and organizations. Based on these results, it is clear that the mean scores are similar across the different levels associated with the variables.

Table 4. Descriptive statistics of the Questionnaire for Smart Glasses User Satisfaction (SGUS).

Variable	Level	n	Mean	St.Dev	Min	Q1	Median	Q3	Max
Gender	Female	31	58.74	7.96	43	54.5	58	64.5	72
	Male	111	58.49	8.45	20	54	60	64.5	72
Role	Experts	47	56.98	8.83	33	49.5	58	64	72
	Students	95	59.32	7.99	20	55	60	65	74
Education level	Upper secondary school or lower	45	57.98	7.92	33	54	57	64	70
	Bachelor's or higher	97	58.8	8.52	20	55	60	65	74
Organization	Space(1)	39	59.54	9.46	20	57	61	65.5	71
	Medicine(2)	48	58.69	7.43	38	54.75	59	64	72
	Engineering(3)	55	57.71	8.26	33	52	57	64	74

5.1.2. QUIS

Similarly, the overall Questionnaire for User Interface Satisfaction (QUIS) score was calculated by summation of the score for the 15 QUIS items. Summary data for all questions in QUIS are presented in Table 5. The 15 items were designed independently from each other. These items aim to investigate the satisfaction of users with different aspects of the interface, including usability and user experience in using AR applications.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 November 2018 doi:10.20944/preprints201810.0594.v2

Peer-reviewed version available at Computers 2019, 8, 9; doi:10.3390/computers8010009

Table 5. Descriptive statistic of the Questionnaire for Smart Glasses User Satisfaction (QUIS).

Variable	Level	n	Mean	St.Dev	Min	Q1	Median	Q3	Max
Gender	Female	30	75.94	11.44	55	69	76.5	82.75	98
	Male	103	76.99	13.23	18	71	78	86	103
Role	Experts	43	76.28	12.06	49	69	77	85.5	97
	Students	90	77.01	13.21	18	72	78	85	103
Education level	Upper secondary school or lower	43	75.14	13.82	18	69.5	75	85	95
	Bachelor's or higher	90	77.56	12.30	33	71	78	85.75	103
Organization	Space(1)	39	76.67	12.44	33	72	77	86	96
	Medicine(2)	42	80.50	9.71	55	75	80	85.75	97
	Engineering(3)	52	73.85	14.61	18	66	74.5	83.5	103

5.2. Correlation

In this section, we discuss correlation for SGUS and correlation for QUIS.

5.2.1. Correlation of SGUS

Spearman's correlation coefficient, ρ, measures the strength and direction of association between two ranked variables in the range [-1, 1]. Based on the 11 items, the results of Spearman's rank correlation are shown in Table 6: the first value of each row represents Spearman's correlation coefficient, and the second value of each row represents the p value. It can be seen that almost all items are statistically significant ($p < 0.05$) and have a low positive correlation. This implies that all the items are independent.

In the study of SGUS, each of the items investigates a different aspect of the user experience. For the analysis, the overall averages for all items were calculated. Figure 9 shows the plot of the average score from individual items. The box in the plot depicts the answer of 50% of the participants, with the line in the middle indicating the median. The dotted lines span the 95% confidence interval. Outliers are depicted with black dots. The connected red dots indicate the medians. The results imply that most of the participants had a good conception of what is real and what is augmented when using AR-glasses (GL5). The participants indicated that the system and content helped them to accomplish the task quite well (GL7) and their attention was captivated in a positive way (GL6). The provided content was also seen as contextually meaningful (GL2). However, performing the task with AR glasses was experienced as less natural (GL9, GL4), and following and understanding the task phases (GL8, GL10–11) was not very easy. The results were very much in line across the three.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 November 2018 doi:10.20944/preprints201810.0594.v2

Peer-reviewed version available at Computers 2019, 8, 9; doi:10.3390/computers8010009

Table 6. Spearman's rank coefficient of correlation for SGUS: the first value of each row represents Spearman's correlation coefficient, and the second value of each row represents the p value.

	GL1	GL2	GL3	GL4	GL5	GL6	GL7	GL8	GL9	GL10	GL11
GL1	1	0.316 0.000**	0.209 0.013	0.269 0.001	0.164 0.053	0.301 0.000**	0.270 0.001	0.323 0.000**	0.285 0.001	0.376 0.000**	0.336 0.000**
GL2		1	0.335 0.000**	0.371 0.000**	0.239 0.005	0.308 0.000	0.345 0.000**	0.227 0.007	0.287 0.001	0.354 0.000**	0.398 0.000**
GL3			1	0.487 0.000**	0.172 0.041	0.270 0.001	0.444 0.000**	0.312 0.000**	0.320 0.000**	0.265 0.002	0.289 0.001
GL4				1	0.121 0.154	0.293 0.000**	0.376 0.000**	0.337 0.000**	0.492 0.000**	0.226 0.008	0.243 0.004
GL5					1	0.260 0.002	0.178 0.036	0.170 0.046	0.026 0.763	0.062 0.468	0.166 0.052
GL6						1	0.416 0.000**	0.387 0.000**	0.453 0.000**	0.317 0.000	0.403 0.000**
GL7							1	0.490 0.000**	0.500 0.000**	0.453 0.000**	0.435 0.000**
GL8								1	0.563 0.000**	0.442 0.000**	0.390 0.000**
GL9									1	0.455 0.000**	0.364 0.000**
GL10										1	0.558 0.000**
GL11											1

Signif. codes: *p < 0.05 ; **p < 0.01; ***p < 0.001

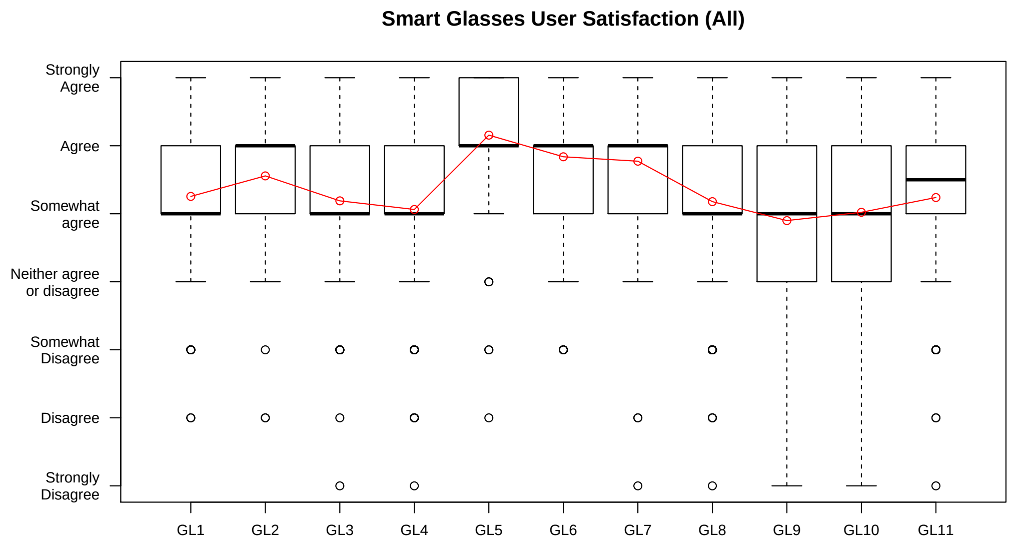


Figure 9. Plot of SGUS score for each item.

262 5.2.2. Correlation of QUIS

263 The correlation for QUIS is based on 15 items. The results of Spearman's rank correlation are
264 shown in Table A1 (see Appendix). The values in the table have the same meaning as in Table 6. The
265 results are similar to those of SGUS; most of the items are statistically significant ($p < 0.05$) and have a
266 low positive correlation. This implies that most of the items are independent.

267 In the study of QUIS, each of the items investigated different aspects of the user experience. For
268 the analysis, the overall average from all items was calculated. Figure 10 shows the plot of the average
269 score from individual items, and the description of the plot is the same as that of the SGUS plot. The
270 results imply that most of the participants agree that learning to operate the AR glasses (QS13) seemed

to be rather easy, and the overall enthusiasm towards the system seemed (QS1, QS5) to be very positive. The characters on the screen were relatively easy to read (QS9). The means of QS3, 4, 6, 7, and 8 indicate that the system was experienced as rigid, unreliable, and slow, which may cause frustration [19].

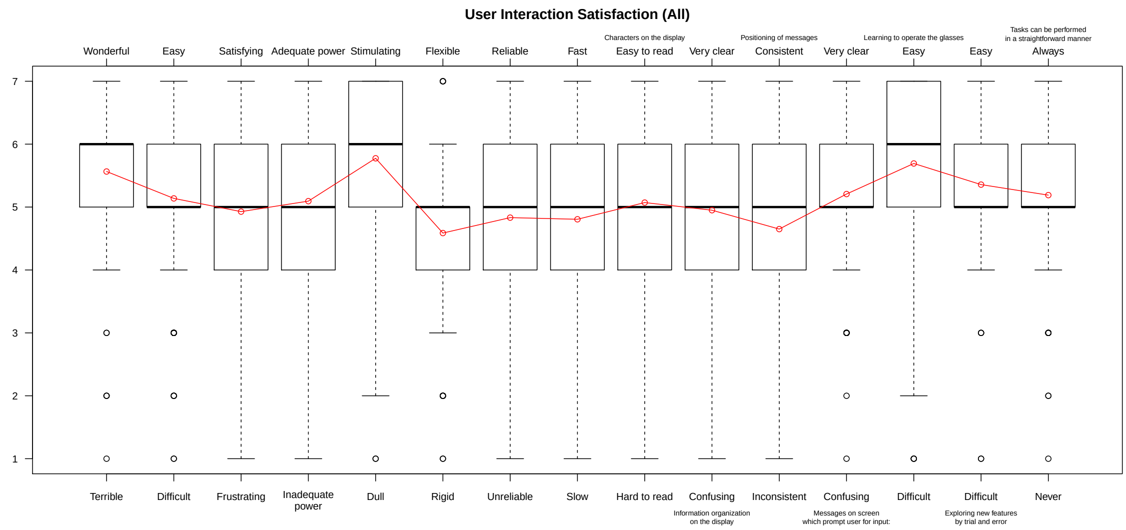


Figure 10. Plot of QUIS score for each item.

5.3. *Analysis of variance and Interaction plots*

The participants are described by seven factors: gender, age, role, education skill level, computer knowledge level, internet knowledge level, and organization. Each factor is divided by two levels, except for organizations, which are in three levels. Please note that none of the participants claimed that they have a poor or very poor internet knowledge level. The following section discusses the analysis of variance (ANOVA) of QUIS and of SGUS. In this ANOVA study, SGUS and QUIS scores were investigated for using the application on the AR glasses with six independent variables, i.e., the relationships between: age distribution, gender, roles, highest level of education, organization, and computer knowledge. Therefore, there are 6 main effects and 57 interactions. We are interested in whether there is a relationship between the satisfaction levels (measured by the questionnaire) and these factors.

5.3.1. ANOVA of SGUS

In this study, we investigated whether the age, gender, roles, computer knowledge level, or different organizations have an effect on the satisfaction of using AR glasses. To determine this, we needed to look at the simple main effects: the main effect of one independent variable (e.g., age) at each level of another independent variable (e.g., for students and for experts).

Figure 11 shows the main effects of the six factors. Participants with different computer knowledge levels have the greatest differences in the SGUS results. This means that the participants with good computer knowledge and poor computer knowledge gave different scores for user satisfaction. The results show that participants with good or very good computer knowledge were, in general, more satisfied with the smart glasses application, and there is a significant effect from computer knowledge levels (F value = 8.87, p = 0.003). The result implies that the SGUS score was affected by the effects of good computer knowledge.

Table 7 shows the summary results of the linear model of the independent variables. The estimate for the model intercept is 54.688 and the coefficient measuring the slope of the relationship with computer knowledge level is 4.324. There is strong evidence that the significance of the model

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 November 2018 doi:10.20944/preprints201810.0594.v2

Peer-reviewed version available at Computers 2019, 8, 9; doi:10.3390/computers8010009

coefficient is significantly different from zero: as the computer skill level increases, so does the satisfaction. The information about the standard errors of these estimates is also provided in the Coefficients table. In the result of the multiple regression model, only 8.8% of the variance in the SGUS scores is explained by each of the factors (Multiple R-squared is 0.088). There is no statistically significant factor that explains the variation in the SGUS scores (overall p value is 0.08).

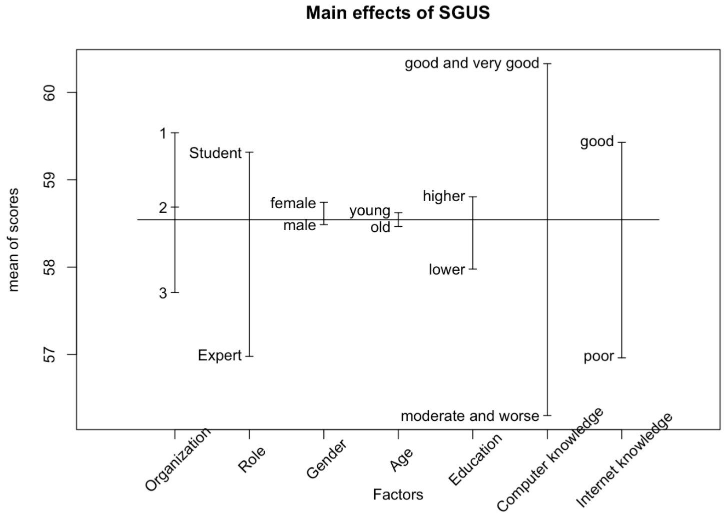


Figure 11. Main effects of SGUS.

Table 7. Results of the linear model of the independent variables.

Source of Variation	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	54.688	2.652	20.622	<2e-16***
Medicine	-1.604	1.835	-0.874	0.384
Engineering	-0.996	1.906	-0.523	0.602
Role	2.862	1.624	1.762	0.080
Gender	1.250	1.756	0.712	0.478
Age	0.563	1.634	0.344	0.731
Education level	-0.147	1.716	-0.086	0.932
Computer skill	4.324	1.452	2.978	0.003

Signif. codes: $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

To investigate the interaction, it is interesting to find out whether the SGUS score depends on an interaction between good computer knowledge and the other factors. The two-factor interaction plot is shown in Figure 12. The following are the findings from the plot:

- Female participants with good computer knowledge have a higher SGUS score than males with good computer knowledge; both females and males with moderate and worse computer knowledge have nearly the same, lower SGUS score (Figure 12a).
- Participants from medicine with good computer knowledge tended toward a higher SGUS score than participants from engineering, and there is no significant difference between them and the participants with good computer knowledge from astronautics and medicine (Figure 12b).
- There is no significant interaction between participants with different computer knowledge levels from astronautics and engineering (Figure 12b).
- There is no significant interaction between students and experts with different computer knowledge levels (Figure 12c).
- Participants younger than 25 years old with good computer knowledge tended toward a higher SGUS than participants older than 25 years old; however, participants younger

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 November 2018 doi:10.20944/preprints201810.0594.v2

Peer-reviewed version available at Computers 2019, 8, 9; doi:10.3390/computers8010009

than 25 years old with a moderate and worse computer knowledge level tended toward a lower SGUS score (Figure 12d).

- Participants with secondary school or lower education level and good computer knowledge tended toward a higher SGUS score than participants with a bachelor's or higher education level and good computer knowledge level. However, participants with secondary school or lower education level and moderate and worse computer knowledge tended toward a lower SGUS score than participants with a bachelor's or higher education level and moderate and worse computer knowledge level (Figure 12e).

From the result of the ANOVA table (Table 8), there is insufficient evidence of statistical significance for two-factor interactions, since all p values are higher than 0.05.

Table 8. ANOVA results for SGUS with regard to organization, role, and computer knowledge level (reducing factors).

Source of Variation	Df	Sum Sq	Mean Sq	F value	Pr(> |F|)
Organization	2	77·9	38·95	0·576	0·563
Role	1	184·4	184·39	2·729	0·101
Gender	1	0·2	0·19	0·003	0·958
Age	1	4·2	4·16	0·062	0·805
Education level	1	0·0	0·02	0·000	0·988
Computer knowledge	1	589·3	589·31	8·723	0·004 **
Education level : Computer knowledge	1	65·0	64·98	0·962	0·329
Gender : Computer skill	1	121·5	121·49	1·798	0·182
Organization : Computer knowledge	2	28·9	14·47	0·214	0·807
Age : Computer knowledge	1	11·6	11·60	0·172	0·679
Roles : Computer knowledge	1	28·6	28·55	0·423	0·517
Residuals	128	8647·7	67·56		

Signif. codes: *p < 0.05 ; **p < 0.01; ***p < 0.001.

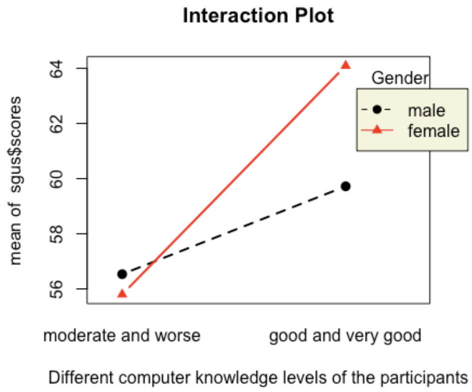
5.3.2. ANOVA of QUIS

In this section, the effect of the six independent variables (age, gender, roles, computer knowledge level, and different organizations) on user interaction satisfaction is reported. Satisfaction includes specific aspects of the interface, usability, and user experience of the AR application.

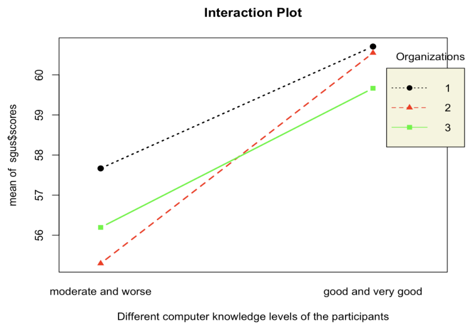
A total of 133 participants were used for this part of the study and completed the questionnaire. The simple main effects are shown in Figure 13. The results obtained by using the ANOVA in Table 9 indicate that the significance of the two-factor interaction of computer knowledge levels and organizations is not supported since all p values are more than 0.05. Table 9 also shows that the computer knowledge levels and different organizations have a significant effect on QUIS (p value is 0.008 for computer knowledge levels and 0.041 for different organizations).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 November 2018 doi:10.20944/preprints201810.0594.v2

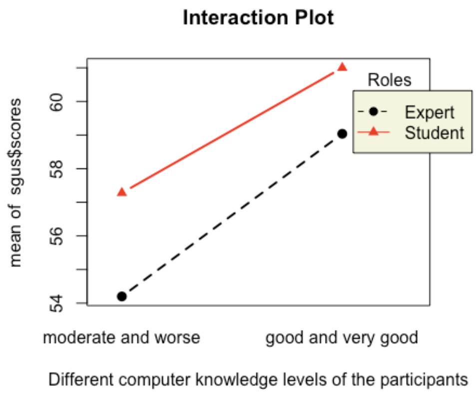
Peer-reviewed version available at Computers 2019, 8, 9; doi:10.3390/computers8010009



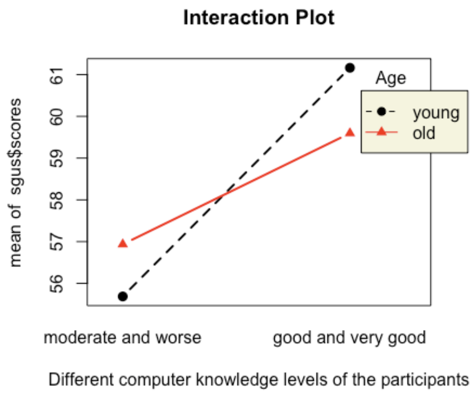
(a)



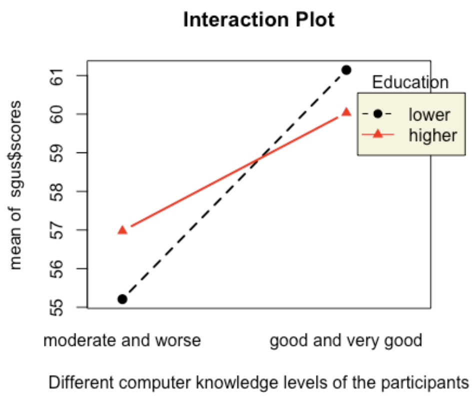
(b)



(c)



(d)



(e)

Figure 12. Interaction effects plots for SGUS: (a) Different computer knowledge levels with different genders of the participants; (b) Different computer knowledge levels with different organizations of the participants; (c) Different computer knowledge levels with different roles of the participants; (d) Different computer knowledge levels with different age groups of the participants; (e) Different computer knowledge levels with different education levels of the participants.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 November 2018 doi:10.20944/preprints201810.0594.v2

Peer-reviewed version available at Computers 2019, 8, 9; doi:10.3390/computers8010009

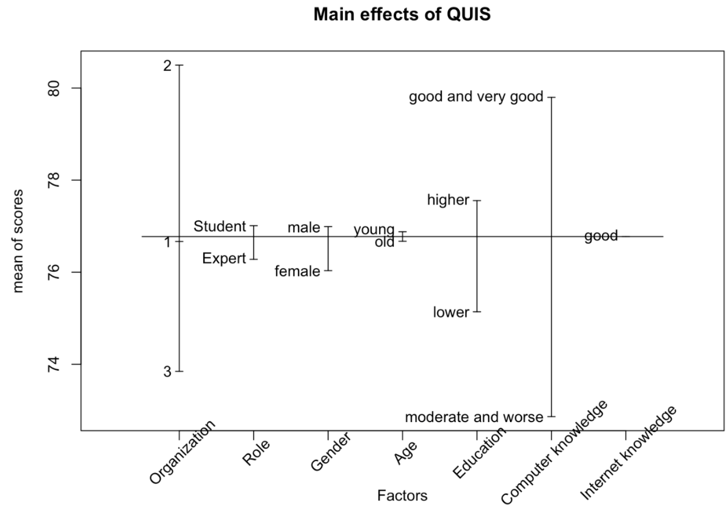


Figure 13. Main effects of QUIS.

Table 9. ANOVA results for QUIS with regard to organization, role, and computer knowledge level (reducing factors).

Source of Variation	Df	Sum Sq	Mean Sq	F value	Pr(> | F |)
Organization	2	1029·3	514·65	3·279	0·041∗
Role	1	10·4	10·37	0·066	0·798
Gender	1	90·3	90·31	0·575	0·450
Age	1	5·8	5·79	0·037	0·848
Education level	1	32·0	32·02	0·204	0·652
Computer knowledge	1	1138·1	1138·14	7·251	0·008∗∗
Education level : Computer knowledge	1	165·5	165·55	1·055	0·307
Gender : Computer skill	1	449·7	449·74	2·865	0·093
Organization : Computer knowledge	2	0·9	0·46	0·003	0·997
Age : Computer knowledge	1	28·2	28·18	0·180	0·673
Roles : Computer knowledge	1	31·8	31·84	0·203	0·653
Residuals	119	18679·1	156·97		

Signif. codes: *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

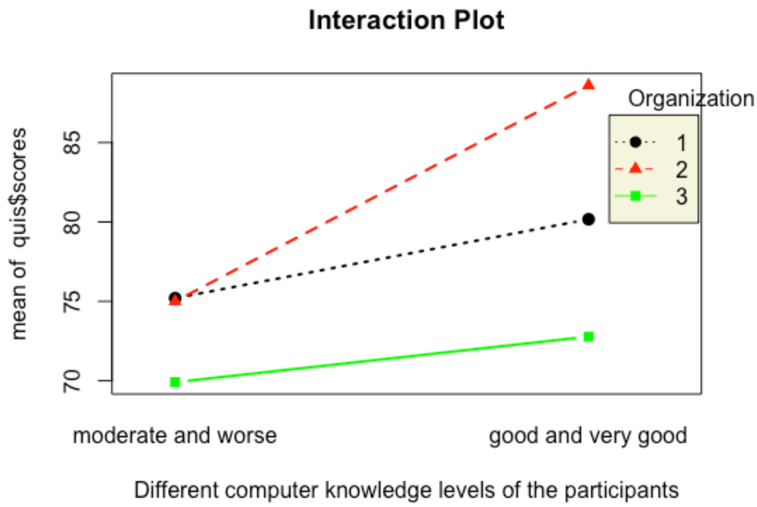


Figure 14. Interaction plot of different computer knowledge levels and the different organizations for QUIS

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 November 2018 doi:10.20944/preprints201810.0594.v2

Peer-reviewed version available at Computers 2019, 8, 9; doi:10.3390/computers8010009

Figure 14 shows that in all three organizations, participants with moderate or worse computer levels were given lower scores than participants with good and very good computer levels. There are no significant interactions between them.

We selected the factors of organization and computer knowledge level to investigate the interaction between them, and the summary results of the linear model regression (see Table 10) shows that the estimate for the model intercept is 73.533, while there is no significant interaction between them. The information about the standard errors of these estimates is also provided in the coefficients table (Table 10). From the result of the multiple regression model, 10.6% of the variance in QUIS scores is explained by each of the factors (Multiple R-squared is 0.106). There is a statistically significant factor to explain the variation in the QUIS scores (overall p value is 0.0133).

Table 10. Summary results of the linear model of the independent variables for QUIS.

Source of Variation	Estimate	Std. Error	t value	Pr(> |t|)
(Intercept)	73.533	3.188	23.063	<2e-16***
Medicine	2.533	4.509	0.562	0.575
Engineering	-2.748	3.951	-0.695	0.488
Computer knowledge	5.092	4.064	1.253	0.213
Medicine: Computer knowledge	1.805	5.686	0.317	0.751
Engineering: Computer knowledge	1.539	5.322	0.289	0.773

Signif. codes: *p < 0.05 ; **p < 0.01; ***p < 0.001.

6. Discussion

This study established a set of norms to be used for the evaluation of satisfaction of using AR glasses and AR applications. The relationship between each questionnaire item shows weak correlation, both in SGUS and in QUIS. Each questionnaire item is designed for evaluating a specific aspect of satisfaction of the smart glasses and AR applications. From the mean score of both questionnaires, we observe that most of the participants are satisfied with the AR glasses and the AR applications. It was found that the system and content helped the participants to accomplish the task quite well and their attention was captivated in a positive way. In other words, the result shows that the user interface is well designed. The user sees "useful information" displayed next to each part.

The main factors age, gender, education level, roles of the participants, and organizations do not have significant effects on the satisfaction of using smart glasses and AR applications. However, computer/internet knowledge level does influence user satisfaction. Participants who have better computer/internet knowledge are more satisfied with the smart glasses and AR applications. There is no significant interaction between all these factors. Since most participants have a moderate level or better than moderate level of knowledge using computers and the internet, it can be predicted that most educated people can easily accept smart glasses and AR applications.

Table 11. Summary of findings

Hypothesis number	Description	Accepted/rejected
H1.	Gender matters	Rejected
H2.	Age matters	Rejected
H3.	Experts and novices will have different level of user satisfaction	Accepted
H4.	Education level matters	Rejected
H5.	Computer knowledge level matters	Accepted
H6.	Internet knowledge level matters	N/A
H7.	Three different test-beds might give different results	Accepted in QUIS, Rejected in SGUS
H8.	There is interaction effects among all these factors	Accepted in SGUS, Rejected in QUIS

Based on the results associated with the seven hypotheses, we outline the following statements:

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 November 2018 doi:10.20944/preprints201810.0594.v2

Peer-reviewed version available at Computers 2019, 8, 9; doi:10.3390/computers8010009

H1: It is commendable, but surprising that we could not identify any gender differences in satisfaction. One possible explanation for this is that our experiment setup asked for volunteers, so we ended up with only people who were interested in the technology, thus not giving us the option to explore, whether there would be any gender differences in the general population with respect to AR training satisfaction.

H2: A possible explanation for not finding any differences in satisfaction levels by age could be that the target group had no prior exposure to AR smart glasses, hence age effects of younger people typically being more open to experimentation of emerging technologies in their home context could not yet affect the picture.

H3: It has probably to do with our applications. The recorder is a more complicated application, challenging experts in their interaction. Even if usually, experts would be more technical savvy, in this case, the findings probably reflect more the differences in user friendliness of the applications.

H4: Only the space case had people in higher education. Most participants in the aviation test bed come from upper secondary backgrounds. Still, the were no differences found in the impact of education level on user satisfaction. The differences may not be obvious in satisfaction levels, but - judging from observation during trials - there were differences across test-beds with respect to how long it took to explain the applications and their use. The application and the use cases enabled everyone, regardless of whether secondary and tertiary education to use the app.

H5: Computer knowledge matters: Better computer knowledge will drive satisfaction with holographic applications. It seems that existing knowledge is still relevant. At the same time, this also clearly indicates that the required support and assistance needs to be provided in order to make the introduction of AR applications on smart glasses a success. Not everyone is a digital native.

H6: Internet knowledge matters: All participants in the trial claimed that they have good internet knowledge and very few people claimed that they have poor internet knowledge, so there was no chance to observe any differences.

H7: There is no difference between the three test-beds in SUGS: We did not find significant differences between the three test-beds. This indicates that occupation does not have direct influence on satisfaction of the AR glasses. Procedure oriented trainings seem to be covered well. There are some difference between the three test-bed in QUIS. The medicine test-bed have the highest satisfaction of the AR app, while the engineering test-bed gave the lowest scores. The procedures of the tasks might effect the satisfaction of the AR app.

H8: There are no interaction effects for QUIS results but some interaction effects amongst the SGUS results. Young people with good computer knowledge are more satisfied the AR glasses. People with lower education and good computer knowledge are more satisfied with the AR glasses than the others.

7. Conclusions

This study was started by noting the scarcity of AR applications for hands-on training. As a first step toward incorporating the recorded teaching activities into learning procedures, the AR application was developed on AR glasses. In this work, the Questionnaire for Smart Glasses User Satisfaction (SGUS) and Questionnaire for User Interaction Satisfaction (QUIS) were investigated for augmented reality applications using Microsoft HoloLens.

The results of this study show that the approach is feasible. The experts wore the AR glasses to show the process, and the activities were recorded. The AR applications can facilitate the students to learn the process. The results show that the satisfaction of both teaching and learning are acceptable. The results indicate that satisfaction does increase when participants have higher computer knowledge levels. It also shows that gender, age, education level, and roles of students or experts do not have any effect on user satisfaction.

Author Contributions: H.X., P.S., F.W. contributed in the conceptualizing, writing and methodology. H.X. and F.W. performed the analysis. H.X. did validation and visualization. P.S. and F.W. helped in the review.

Funding: This study is part of the WEKIT (Wearable Experience for Knowledge Intensive Training) project. It is a European project supported under Horizon 2020 and is aiming to enable smart wearable technology for industrial training.

Acknowledgments: We express sincere appreciation to Kaj Helin, Principal Scientist, Project Manager of VTT Technical Research Centre of Finland, and Timo Kuula, Research Scientist of VTT Technical Research Centre of Finland for their modification of questionnaire QUIS and their design of questionnaire SGUS. We thank all the participants and companies for their valuable contributions.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AR	Augmented Reality
UI	user interface
QUIS	Questionnaire for User Interaction Satisfaction
SGUS	Smart Glasses User Satisfaction
WEKIT	Wearable Experience for Knowledge Intensive Training

Appendix

Table A1. Spearman's rank coefficient of correlation of QUIS: the first value of each row represents Spearman's correlation coefficient, and the second value of each row represents the p value.

	QS1	QS2	QS3	QS4	QS5	QS6	QS7	QS8	QS9	QS10	QS11	QS12	QS13	QS14	QS15
QS1	1	0.39 0.000**	0.53 0.000**	0.50 0.000**	0.49 0.000**	0.34 0.000**	0.53 0.000**	0.47 0.000**	0.17 0.05	0.37 0.000**	0.37 0.000**	0.28 0.001	0.31 0.000**	0.32 0.000**	0.44 0.000**
QS2		1	0.53 0.000**	0.41 0.000**	0.34 0.000**	0.30 0.000**	0.37 0.000**	0.43 0.000**	0.16 0.07	0.35 0.000**	0.25 0.003	0.33 0.000**	0.58 0.000**	0.50 0.000**	0.52 0.000**
QS3			1	0.56 0.000**	0.55 0.000**	0.39 0.000**	0.49 0.000**	0.45 0.000**	0.16 0.06	0.33 0.000**	0.27 0.001	0.22 0.009	0.35 0.000**	0.37 0.000**	0.40 0.000**
QS4				1	0.49 0.000**	0.23 0.008	0.42 0.000**	0.41 0.000**	0.18 0.04	0.35 0.000**	0.38 0.000**	0.30 0.000**	0.27 0.001	0.27 0.001	0.40 0.000**
QS5					1	0.22 0.01	0.41 0.000**	0.45 0.000**	0.14 0.11	0.22 0.01	0.13 0.12	0.14 0.10	0.24 0.005	0.36 0.000**	0.34 0.000**
QS6						1	0.36 0.000**	0.26 0.002	0.26 0.001	0.18 0.03	0.25 0.003	0.09 0.28	0.28 0.000**	0.33 0.000**	0.33 0.000*
QS7							1	0.54 0.000**	0.17 0.05	0.38 0.000**	0.39 0.000**	0.28 0.001	0.24 0.004	0.35 0.000**	0.44 0.000**
QS8								1	0.23 0.006	0.40 0.000**	0.26 0.002	0.33 0.000**	0.26 0.002	0.40 0.000**	0.43 0.000**
QS9									1	0.35 0.000**	0.31 0.000**	0.31 0.000	0.19 0.024	0.32 0.000**	0.24 0.005
QS10										1	0.57 0.000**	0.45 0.000**	0.27 0.001	0.29 0.001	0.44 0.000**
QS11											1	0.43 0.000**	0.25 0.003	0.33 0.000**	0.38 0.000**
QS12												1	0.34 0.000**	0.30 0.000**	0.42 0.000**
QS13													1	0.57 0.000**	0.48 0.000**
QS14														1	0.47 0.000**
QS15															1

Signif. codes: *p < 0.05 ; **p < 0.01; ***p < 0.001

References

1. Azuma, Ronald T. "A survey of augmented reality." In *Teleoperators & Virtual Environments* 6.4 (1997): 355-385.
2. Sherstyuk A, Vincent D, Berg B, Treskunov A. Mixed reality manikins for Medical Education. In *Handbook of Augmented reality*; Springer: New York, NY, 2011; pp. 479-500, ISBN.
3. Ong SK, Shen Y, Zhang J, Nee AYC. Augmented Reality in Assistive Technology and Rehabilitation Engineering. In *Handbook of Augmented Reality*; Springer: New York, NY, 2011; pp. 603-630, ISBN.
4. Nilsson J, Ödblom ACE, Fredriksson J, Zafar A. Using Augmentation Techniques for Performance Evaluation in Automotive Safety. In *Handbook of Augmented Reality*; Springer: New York, NY, 2011; pp. 631-649, ISBN.
5. Ras E, Wild F, Stahl C, Baudet A. Bridging the Skills Gap of Workers in Industry 4.0 by Human Performance Augmentation Tools. *Proceedings of the 10th International Conference on PErvasive Technologies Related to Assistive Environments - PETRA '17* **2017**, doi:10.1145/3056540.3076192.
6. Perey C, Wild F, Helin K, Janak M, Davies P, Ryan P. Advanced manufacturing with augmented reality. *2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* **2014**, doi:10.1109/ismar.2014.6948518.
7. Olsson T. Concepts and Subjective Measures for Evaluating User Experience of Mobile Augmented Reality Services. In *Human Factors in Augmented Reality Environments*; Springer: New York, NY, 2013; pp. 203-232, ISBN.
8. Augmented Reality: What Does It Mean for UX? In: Nielsen Norman Group. Available online: https://www.nngroup.com/articles/augmented-reality-ux/ (accessed on 5 Jul 2018).
9. Microsoft. Microsoft HoloLens. In: Microsoft HoloLens. Available online: https://www.microsoft.com/en-us/hololens/commercial-overview (accessed on 5 Jul 2018).
10. Rauschnabel, Philipp A.; Brem, Alexander; Ro, Young K. Augmented Reality Smart Glasses: Definition, Conceptual Insights, and Managerial Importance.Unpublished Working Paper, The University of Michigan-Dearborn, College of Business (2015). Available online: https://www.researchgate.net/profile/Alexander_Brem/publication/279942768_Augmented_Reality_Smart_Glasses_Definition_Conceptual_Insights_and_Managerial_Importance/links/5721ec2e08aee857c3b5dd6c/Augmented-Reality-Smart-Glasses-Definition-Conceptual-Insights-and-Managerial-Importance.pdf (accessed on 12 Jul 2018)
11. Rauschnabel PA, Brem A, Ivens BS. Who will buy smart glasses? Empirical results of two pre-market-entry studies on the role of personality in individual awareness and intended adoption of Google Glass wearables. *Computers in Human Behavior* **2015**, *49*, 635-647, https://doi.org/10.1016/j.chb.2015.03.003
12. Wiederhold, Brenda K. "Time to port augmented reality health apps to smart glasses?." *Cyberpsychology, Behavior, and Social Networking* **2013**, *Volume: 16 Issue 3*, 157-158, http://doi.org/10.1089/cyber.2013.1503
13. Questionnaire For User Interaction Satisfaction. Available online: http://lap.umd.edu/quis/ (accessed on 29 Jan 2018).
14. Ssemugabi S, de Villiers R. A comparative study of two usability evaluation methods using a web-based e-learning application. *Proceedings of the 2007 annual research conference of the South African institute of computer scientists and information technologists on IT research in developing countries - SAICSIT '07* **2007**, doi:10.1145/1292491.1292507.
15. Limbu B., Fominykh M., Klemke R., Specht M., Wild F. (2018) Supporting Training of Expertise with Wearable Technologies: The WEKIT Reference Framework. In: Yu S., Ally M., Tsinakos A. (eds) Mobile and Ubiquitous Learning. Perspectives on Rethinking and Reforming Education. Springer, Singapore
16. Puneet Sharma, Roland Klemke, Fridolin Wild, Experience Capturing with Wearable Technology in the WEKIT project, SIGWELL 2018
17. Fominykh M. D2.4 First Prototype. Available online: http://wekit.eu/d2-4first-prototype/ (accessed on 2 Sep 2018).
18. WEKIT D1.4 Requirements for Scenarios and Prototypes. Available online: http://wekit.studiohangloose.it/wp-content/uploads/2017/06/WEKIT_D1.4.pdf (accessed on 7 Nov 2017).
19. WEKIT D6.4 Implementation of Evaluation Trials in Aeronautics. Available online: http://wekit.eu/wp-content/uploads/2017/09/WEKIT_D6.4 (accessed on 7 Nov 2017).
20. WEKIT D6.2 Annex 1 Training scenario and Evaluation Plan for Engineering. Available online: http://wekit.studiohangloose.it/wp-content/uploads/2017/06/WEKIT_D6.2.pdf (accessed on 9 Nov 2017).
21. WEKIT D6.5 Implementation of Evaluation Trials in Engineering. Available online: http://wekit.eu/wp-content/uploads/2017/09/WEKIT_D6.5 (accessed on 9 Nov 2017).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 November 2018 doi:10.20944/preprints201810.0594.v2

Peer-reviewed version available at Computers 2019, 8, 9; doi:10.3390/computers8010009

22. WEKIT D6.6 Implementation of Evaluation Trials in Space. Available online: http://wekit.eu/wp-content/uploads/2017/09/WEKIT_D6.6 (accessed on 5 Jun 2018).
23. Chin, John P., Virginia A. Diehl, and Kent L. Norman. In *Proceedings of the SIGCHI conference on Human factors in computing systems*; ACM: New York, USA, 1988; pp. 213-218, ISBN:0-201-14237-6.
24. Hall, Douglas T., and Philip H. Mirvis. "The new career contract: Developing the whole person at midlife and beyond." Journal of vocational behavior 47.3 (1995): 269-289.