

User evaluation of mobile augmented reality scenarios

Thomas Olsson^{a,*}, Tuula Kärkkäinen^a, Else Lagerstam^a and Leena Ventä-Olkkonen^b

^a *Unit of Human-Centered Technology, Tampere University of Technology, Korkeakoulunkatu 6, 33720 Tampere, Finland*

^b *Nokia Research Center, Tampere, Visiokatu 1, 33720 Tampere, Finland*

Abstract. Augmented reality (AR) technologies are becoming state-of-the-art in the mobile technology domain. However, developing successful end user services around AR is still in its infancy, which is partially resulting from the lack of user research regarding potential users' expectations and user acceptance in the particular area. In order to identify the most potential use cases and contexts, in which mobile AR could be utilized, it is important to evaluate already the early service concepts with potential end users. This paper reports an online survey for evaluating the user acceptance of five different mobile AR scenarios, as well as for understanding the potential and risks of mobile AR in general. The results show ambivalent attitudes towards mobile AR services but also imply that the scenarios demonstrating pragmatic relevance were valued over pleasure oriented ones. AR was seen to make contextually relevant information easily available and allow novel interaction possibilities with the physical world. Distrust arose with fears of information flood, users' loss of autonomy, and virtual replacing the real. The respondents' level of technological orientation was found to highly affect the overall acceptance of the scenarios. The results help in considering key acceptance issues and potential users' expectations in the development of future mobile AR services.

Keywords: Augmented reality, mobile services, user research, user acceptance, user expectations, online survey

1. Introduction

Augmented reality (AR) as an emerging interactive technology has increasingly gained public interest during the last few years. The concept of augmented reality refers to combining real and computer-generated digital information into the user's view of the physical real world in such a way that they appear as one environment [17,39]. Furthermore, AR relates to a broader concept of mixed reality (MR), which means the integration and merging of the real and virtual worlds in general [26]. AR is becoming state-of-the-art in the mobile domain, thanks to viable solutions for the interplay of mobile processing, image recognition, and sensor technologies. Not only have its technical enablers and respective prototypes been researched extensively for decades but also the first publicly available mobile applications have lately been introduced.

Augmented reality efficiently incorporates everyday objects and locations as part of a user interface through which the interaction with digital content takes place. This new reality of digital information aligned to the physical world can be interacted with in real time. Hence, it serves as a versatile user interface for services of ambient intelligence and smart environments [32]. In fact, AR can be considered a central technology in the paradigm shift from desktop-based interaction towards ubiquitous computing, enabling smart services "anytime, anywhere" [36,42]. AR integrates various aspects of ubiquitous computing with a goal to enable people to take advantage of both their own skills and the power of networked computing while naturally interacting in the everyday physical world [25]. Overall, the new realm of augmenting the physical world with digital artifacts and information holds the potential to revolutionize the way in which information is accessed and presented

* Corresponding author. E-mail: thomas.olsson@tut.fi.

to people: “the world becomes the user interface” [17,43].

This research is based on the user-centered design approach (UCD) [18]. This approach emphasizes that the users’ expectations and needs must be enthroned to increase the opportunities of developing successful and engaging services. The services should be based on true or latent needs and expectations of the potential end user group. Prior expectations influence how people will experience and value the actual services, and what kind of implications the services might have on behavior and conventions [2]. In addition, the concepts and solutions should be evaluated as early as possible and iterated in several rounds to ensure their acceptance. Results of evaluation can further raise new user requirements that have failed to become identified in user needs studies. The UCD approach becomes especially important with new technologies that have few demonstrators or services available, and when there is a high risk of unsuccessful development investment.

Furthermore, important related concepts are user acceptance and user experience. *User acceptance* (UA) refers to users’ willingness (or intention) to accept and use available systems and services [21,40]. The technology acceptance model for mobile services (TAMM) by Kaasinen [21] states that user acceptance is mainly affected by the perceived ease of use, perceived value, and trust in the given technology. User acceptance in general relates to the theory of *diffusion of innovations* by Rogers [33]. The research in this paper represents the phases of *knowledge* and *persuasion* where potential users are exposed to an innovation (AR) and their perceptions of it might have influence on their later use of the technology.

User experience (UX) is regarded as a subjective and holistic concept describing the experience resulted from the interaction with a technological product or service. Both instrumental (e.g. utility and usability) and non-instrumental (e.g. joy, appeal, aesthetics) elements are often discussed in UX literature [12]. A recent ISO standard [18] defines UX as “A person’s perceptions and responses that result from the use or anticipated use of a product, system or service”. With regard to our approach, it is worthwhile to highlight the relevance of *anticipated* (or expected) experience. Accordingly, the expectations of potential users can also vary from utilitarian and pragmatic aspects to emotional, aesthetic and social elements in the use of a product or service.

In this paper, *AR service* is understood as a comprehensive entity that includes not only the applica-

tion and its functionalities but also the AR content, interaction with it and the required technological devices. This further implies that user evaluation can touch the user acceptance and expected user experience from multiple viewpoints.

1.1. The unexploited potential of mobile augmented reality

Current mobile devices, such as mobile phones, digital cameras, and navigators are becoming a powerful platform for augmented reality. Development of sensor (e.g. cameras, location, orientation) and communication technologies enables mobile devices to gradually become context-aware [37]. This means that the users’ physical, social, and task context, as well as the technological resources nearby can be extracted, interpreted, and utilized in applications [6,16]. Context awareness often manifests as location-awareness, where the system measures the user’s location and provides the user with information bound to that very location and the physical objects in the surroundings [31].

Mobile augmented reality (MAR) expands AR services to cover an extensive set of use cases and scenarios in the immensely rich diversity of mobile environments [41]. AR can potentially be applied in various practical day-to-day use cases, in entertainment and gaming, as well as in tourism and navigation. A mobile AR service could provide the user with context-related information that is very relevant, useful or entertaining in the very moment and location, thus supporting various context-dependent, ad hoc needs [29]. Examples of sources of information are surrounding places, near-by objects, products, or events, as well as moving objects like people or transportation.

The aforementioned elements create great opportunities and space for design of mobile AR services, which, however, has so far been very little utilized. Recently, this principle of aligning layers of digital and real world information on top of each other has been demonstrated in publicly available mobile applications like Layar (www.layar.com), Junaio (www.junaio.com), and Wikitude (www.wikitude.org). However, because of issues in sensor inaccuracy and lack of true utilitarian value, such first consumer-level demonstrators of mobile AR leave much space for utilizing the possibilities of AR in creating a pleasurable and rich user experience [30].

1.2. Related work

Most of the research on mixed and augmented reality has focused on constructive development of various types of displays and other output devices (see e.g. [26]), as well as algorithms for identifying and tracking real world objects and rendering the virtual (see e.g. [1,3,40]). User expectations, user experience, and acceptance issues of augmented reality have been studied very little, and the whole approach of user-centered development has been underutilized in AR [7,9,38]. Furthermore, most of the user research that exist has focused on evaluating early technical demonstrators in specific contexts, especially looking into perception and cognition issues, user task performance, or other usability-related aspects [10,22,28].

User experience and acceptance of such AR demonstrators have been often dismissed, especially when considering emotional aspects of the experience. Gabbard and Swan [9] and Nilsson [28] have proposed well-justified methods and processes for user-based studies for AR, however with a focus on usability studies of already developed solutions. User research with concept-level descriptions of holistic AR *services* is virtually non-existent, despite its usefulness already in the early phases of development.

In the following, earlier research on AR is discussed from a user-centered point of view, focusing on research papers that present applications and prototypes that utilize AR and other closely related technologies, and possibly include user evaluation.

To start with, Morrison et al. [27] present their findings from field trials of MapLens, which is an AR map using the magic lens metaphor over a paper map. They found that AR features facilitated place-making by allowing referencing to the physical world. The strongest potential of AR maps was seen to be their use as a collaborative tool. Another example of a mobile AR application where user tests has been conducted is the ARCHEOGUIDE, presented by Gleue and Dähne [11]. It provides augmented archaeological information related to cultural heritage sites. With a small mobile computer and display, the user is able to experience the real site while visualizations of virtual reconstructions are integrated seamlessly into the natural field of view. They present early user experiences from the first trials, however, mostly focusing on the user interface and other technical equipment and their usability, whereas the actual experiences or user acceptance aspects are discussed very little.

Other interesting AR concepts are presented, for example by Sato et al. [35] who demonstrate a novel Mixed Reality display system called MR-mirror. It merges real visual information reflected on a real mirror and a virtual one displayed on an electronic monitor. No actual user evaluations were carried out. Jain et al. [19] conducted a week-long pilot study to understand how a MR prototype for creating shopping lists is used and how its usability is evaluated. Herbst et al. [14] present a mobile outdoor mixed reality game for exploring the history of a city in the spatial and the temporal dimension. There are also various gaming-related AR solutions, for example those presented in [5,34].

Some location based systems are also worth mentioning. Ludford et al. [24] have developed a location based reminder system called PlaceMail. They demonstrated its utility in supporting everyday tasks through a field study, and found out that PlaceMail supports useful location-based reminders and functional place-based lists. Burrell et al. [4] implemented and evaluated a location-sensitive college campus tour guide called Campus Aware, which allows users to annotate physical spaces with text notes. A field-study based user evaluation was conducted, and it pointed out that users indeed provided unique content interesting and useful to others. Finally, based on interviews about people's expectations concerning context- and location-aware services, Kaasinen [20] states that the most potential application areas for mobile location-based services can be found in areas such as travel and tourist information, shopping, entertainment and event information.

In our own earlier research [29] we have studied user acceptance and users' expectations of AR with mostly qualitative methods. These have elicited a good amount of user requirements and other aspects to consider in design. However, with a qualitative approach the low number of participants has not made it possible to assess the user acceptance and expectations of AR with quantitative means.

1.3. Research rationale and approach

Despite there being a large body of related research and development of AR technology, the research community is still exploring what could be the most potential use cases for successful AR services. Little is known about what potential users expect from the multifaceted nature of AR, e.g. regarding interaction, used devices, context of use, smart functionalities like context awareness and proactivity, as

well as the ways of creating, sharing, browsing, and searching the AR content.

Technically, AR in the mobile domain is becoming mature enough for commercial, high-quality end user services to be introduced. AR gradually moves towards such a level of maturity of innovation that the services are used for their commodity value instead of the technological novelty. Therefore, it has become increasingly important to assess the service demonstrators early on to understand what people expect from them. Such understanding throws light on the overall potential of AR as a solution for providing location-based, intelligent services.

User evaluation and gathering user expectations is challenging without proper stimuli. If user research is conducted in laboratory environments, and not tied to real life tasks, it is hard for the participants to assess the overall experience and value of the service at hand. Eliciting understanding on holistic and implicit issues, such as overall user experience, or social and societal implications, is challenging without introducing the service concept in an authentic contexts of use and with the intended user group.

Taken all this into consideration, this study utilized scenarios in an online survey to evaluate AR service concepts already in an early development phase. The scenarios were materialized as language-based descriptions of use of AR services, aiming to evoke a *metaphorical* experience of the envisioned services. With relatively small data gathering effort, online scenarios can elicit valuable and multifaceted data about the perceptions of the novel technology or AR from a large group of potential users. This paper presents five service scenarios and their evaluation results from two viewpoints: both from scenario specific perspective and from the viewpoint of general acceptance of augmented reality. Both qualitative and quantitative evaluation data is used in the analysis.

2. Evaluating AR scenarios with an online survey

2.1. Research objectives and focus

The goal was to evaluate scenarios representing augmented reality already in an early phase of development, before enforcing any service implementation efforts. Introducing five different AR scenarios enabled the examination of potential end users' pre-use views on the benefits and values, as well as the drawbacks and risks of mobile AR services in general. Instead of looking into usability and perceptual

issues of AR, this study focused on the experiential level and user acceptance issues. A special focus was set to identifying potential use cases for future mobile AR services and understanding what is their overall value to the user. End user evaluation further enabled gathering user expectations and requirements for multifaceted AR *services* – not merely requirements for a certain technological solution or design.

As a foundation for the online survey, five scenarios were composed. The scenarios exhibited futuristic services in various everyday life situations and contexts, mostly demonstrating archetypical aspects of augmented reality, context awareness (esp. location-awareness), and 'smartness' of environment in general (e.g. through adaptation of the AR service to the preferences of the user). Regarding the context of use, the scenarios described people in everyday urban life using fictional, intelligent AR services to reach their goals. They described the context of use of the service, the content of the service, the content providers and how the user would interact with the service. Furthermore, to avoid social context (e.g. collaboration, other people present) in the scenarios to induce additional complexity in the acceptance evaluation, the focus was merely on personal use from a single-device perspective.

In order to evaluate the acceptance and expected value of the scenarios on a larger scale, the online survey was conducted in Finnish and in English (identical versions apart from the language). The rationale behind this was not to carry out an analysis of the differences between the country-specific data, but simply to expand the subset of people able to respond to the survey. As the main recruitment channels reached mostly Finnish people, they were given a chance to answer in their own language. The Finnish survey was open for 14 days and the English for 10 days, both during late spring 2009. 5 Amazon-vouchers worth 30eur and small product prizes were raffled among the respondents. Before launching, the survey was piloted and iterated in two rounds with potential respondents to ensure the understandability of the scenarios, questions, and instructions.

2.2. Survey structure

The online survey consisted of four parts: (1) a short introduction to mixed and augmented reality, simply to clarify the concepts, (2) the five scenarios with Likert statements and questions related to each scenario, identically for each one, (3) questions about AR services in general and, (4) questions about par-

ticipant background. The survey consisted of both quantitative and qualitative items.

Regarding part 2, the statements were applied from earlier research [21]: “The services would be appropriate in the described situation” (*perceived value*), “The services would help me to reach my goals” (*perceived value*), and “I would like to try out the services” (*intention to use*). The set of statements in this phase was kept short and simple to avoid response fatigue. Additionally, the construct of *perceived ease of use* was regarded as too challenging to be evaluated based on such abstract and textual scenarios. The Likert-scale statements had a 7-step scale (as recommended in [8]) from 1 (completely disagree) to 7 (completely agree), accompanied with “not applicable”. Furthermore, two open questions were asked after each scenario to gather qualitative data: “Which of the services mentioned in the story you consider the most useful or the most entertaining? In what way?” and “Which of the services mentioned in the story you consider useless or disturbing? Why?”

Part 3 took a more general level approach to AR, based on all the five scenarios. 18 statements regarding, for example, usefulness and user acceptance of AR overall were presented (reported in 3.3). The statements were partly based on earlier research on user acceptance [21,40]. However, new items were also created and statements from our earlier studies [29] were used as we wanted to include specific items about the UX of AR. In addition, three open questions were asked: “Which of the preceding situations you found most interesting? Why?”, “What would affect most on your willingness to try out the described-like services?”, and “In what kind of situations you would like to get the information from your environment automatically? What would the information be related to?”

Part 4 included questions and statements concerning demographic background, experience on existing mobile technologies, and attitudes and expertise of technology in general (most reported in Section 2.5).

2.3. Scenarios

The scenarios described people in various mundane situations using fictional MR services to reach their goals. The scenarios aimed at communicating potential MAR services from various viewpoints; exemplifying various potential use cases, contexts of use, types of AR information content, devices, and ways of interacting with the MAR services that could be technically feasible until 2015 or so. The five sce-

Maria has started jogging some time ago and is now leaving for a run. Maria picks up her keys, mobile phone and sports glasses that are wirelessly connected to the mobile and work as a display. In the elevator on her way down Maria uses the mobile and the glasses to browse neighboring area jogging routes that others have shared. She selects a 6 kilometer route with an easy land profile that goes around a forest.

While Maria is running the glasses guide her with real-time pointers to the selected route and with staying on it. With the help of the signals that her glasses and mobile give her, Maria sticks well to the previously unknown route. At first Maria runs a bit slowly, then speeding up her pace until the glasses indicate that Maria has reached her optimal heart rate.

As the evening starts to get darker the glasses use night vision amplifier to help Maria to see better in the dark forest. Maria sees also some comments left by other runners along the route. Maria has run faster than she expected so she decides to run a bit longer route and turns to an optional route indicated by the glasses. Soon the glasses show Maria that her fellow student is running ahead. Maria decides to speed up her pace to catch her up. It's always more fun to run in a company after all.

Fig. 1. Scenario 2: Jogging.

narios aimed at covering various types of services: from practical benefits to fun and self-expression, from personal to social information content, and with various target user groups and contexts of use. For simplicity and understandability, the use cases were mostly rather general and approachable – instead of demonstrating specific needs, for example related to geocaching or other highly location-based activities. Such could be potential application areas for AR but at the same time very uncommon activities among the population. Furthermore, the scenarios did not deal with the price of the introduced services as such elements were suspected to evoke iniquitous evaluations for our academic goals. With five scenarios the survey took approximately 15–20 minutes to complete. To exemplify, Fig. 1 describes scenario #2 as it was presented in the online survey. The contents of the four other scenarios are summarized below.

2.3.1. Scenario 1: On the bus

The scenario described MAR services offering everyday practical and location-dependent information and advertisements. Specifically it related to public transportation: assistance in planning the trip and offering both informative and entertaining content to pass time. The services included also paying the bus fare with the mobile and receiving an automatic AR notification when to get off the bus. The content and the functions of the services were tied to physical objects: screen at the bus stop and the reader and stickers on the bus, accessed through the camera view of the mobile.

2.3.2. Scenario 3: Shopping furniture

This scenario focused on additional product information and AR-based visualization of furniture models. The users have had their living room 3D-modeled and now use the virtual model in finding new suitable furniture in the room. In a furniture store the users can collect related information like the virtual model, price information, the stock situation and the color options to their mobiles by touching the price tags of the furniture on display. They can use a separate room in the store to project the model of the room with the new furniture on a wall in actual size.

2.3.3. Scenario 4: Virtual mirror

The fourth scenario described a virtual makeover service located in a clothing store. The service consisted of a virtual mirror and data glasses connected to it. By wearing the glasses the users could see themselves in the mirror with different styles and appearances. They could try on different clothes, hair styles, make-up, accessories and piercings. The virtual look could be shared with others using the same mirror with glasses on.

2.3.4. Scenario 5: Street art

This scenario described a service focusing on artistic and amusing content, i.e. user generated virtual street art attached to real world objects. The users can view the graffiti with their mobile, add comments to them or create new art themselves and share them with others. The creation in situ happens by painting in the air with the phone or by compiling existing images and text previously saved to the mobile. The service could be set to notify about nearby pieces of AR art that represent selected topics or are created by selected friends.

2.4. Analysis of the data

Regarding the analysis of open questions (qualitative), themes mentioned in the answers were categorized and the frequency of the mentions of each category was counted. The identification of categories and the quantification were carried out by two researchers in turns, reflecting both user acceptance and user experience elements. The percentages of the quantifications are meant to be inferred merely as indicative as the contribution of the results is seen to be in the diversity of the identified themes and categories instead of the generalization. Accordingly, no analysis of inter-rater reliability between the researchers was seen useful to be conducted. The origi-

nally Finnish answers presented in this paper have been translated by the authors but the user quotes from the English survey are verbatim.

The tables in this paper about descriptive statistics report also means and standard deviations of the 7-step Likert-scale ratings. However, in statistical tests for significant differences (Sections 3.1 and 3.3), the Likert scale was treated as an ordinal level scale for methodological hygiene. In addition, not all the distributions of the ratings were Gaussian, which led to using non-parametric Friedman's rank tests to compare the subjective evaluations of the scenarios for significant differences. Wilcoxon's matched pairs signed ranks tests were used in pair wise comparisons of scenarios. Bonferroni corrected significance levels are reported in all pair wise comparisons. Similarly, Mann-Whitney U-test for independent samples was used to analyze the effect of various background variables on the evaluations (in Sections 3.1 and 3.3). "Not applicable" data points were omitted from statistical analysis (most statements had 3–4 "N/A" selections, except for one with 13).

2.5. Participants

Altogether, the survey gained 262 responses: 182 of them in the Finnish and 80 in the English version. The answers of 2 respondents were omitted because of probable mischief (improper or totally lacking answers to open questions), resulting in 260 responses. 127 of those were male and 133 female. Ages varied between 17 and 64 years, however 80% of participants being young adults aged between 21 and 36. Over 80% of the Finnish version and over 70% of the English version respondents were either studying in or graduated from a university.

69% of the participants reported to at least sometimes browse the web with their mobile phone, 85% to use camera, 48% to use maps or navigation, and 51% to use e-mail. Table 1 further indicates that most participants can be considered highly or moderately technologically oriented, which, on the other hand, was also intended.

In order to get respondents who can understand the principles of AR and hence be able to express their opinions of the scenarios, they were expected to represent potential early adopters of new technology. The participants were recruited via various mailing lists targeted to people interested in user testing, and students and staff of a technical university including various faculties related to science, technology, economy, and architecture.

Table 1

Means (M), standard deviations (SD), and medians (Mdn) of technology orientation and attitude (statements in original format). Scale: 1–7 (completely disagree – completely agree)

| N = 260 | M | SD | Mdn |
|---|-----|-----|-----|
| I think that technology is necessary in my everyday life | 6.0 | 1.3 | 6 |
| I am among the first ones in my circle of friends who get new devices | 4.2 | 1.8 | 5 |
| I help my friends and relatives using technical devices | 5.4 | 1.6 | 6 |
| I like to share information about myself in the web, for example in Facebook or MySpace | 4.0 | 1.9 | 4 |
| Spreading and abusing of my personal data in the web worries me | 5.2 | 1.9 | 6 |

3. Results

As an introduction to the scenario comparisons, Table 2 describes the averages of the three evaluative statements presented after each scenario. Some of the means of the statements deviated rather little from the center point of the scale (4 in 1–7), slightly slanting towards agreeing with the statements. In addition, the standard deviations were rather high. Looking more closely in the distributions shows that they were far from a Gaussian curve. In some statement-scenario combinations the distribution was more or less straight; all the values in the scale used equally. In others, the opinions were mostly polarized in two groups; those strongly agreeing (6–7) and those strongly disagreeing (1–2), while the center values (3–5) were used relatively little. This indicates that the scenarios and the statements as such served well in eliciting people’s opinions (little mid-values and N/A’s) but at the same time gained very contrary evaluations.

Overall, the most highly rated scenario seems to be *Shopping furniture* with the highest averages (all means being over 5). The standard deviations were also slightly smaller in this case, indicating that this scenario in general was relatively well accepted. *On the bus* can be seen as a runner-up, introducing services that most would like to try out. *Jogging* and *Virtual mirror* were also evaluated somewhat positively but then again the variation in the distribution was also higher, meaning that these polarized opinions more. Generally the lowest rated scenario was *Street art* with lowest ratings in all of the statements.

3.1. Statistical analysis of the scenario-specific statements

Some of the differences between scenario ratings proved to be statistically significant. With regard to

Table 2

Means (M), standard deviations (SD), and medians (Mdn) of scenario specific statements. Scale: 1–7 (completely disagree – completely agree)

| N = 236 | The services would be appropriate in the described situation | | | The services would help me to reach my goals | | | I would like to try out the services | | |
|---------------------------|--|-----|-----|--|-----|-----|--------------------------------------|-----|-----|
| | M | SD | Mdn | M | SD | Mdn | M | SD | Mdn |
| <i>On the bus</i> | 5.0 | 1.5 | 5 | 4.4 | 1.7 | 5 | 5.5 | 1.7 | 6 |
| <i>Jogging</i> | 4.8 | 1.9 | 5 | 4.5 | 1.9 | 5 | 5.3 | 2.0 | 6 |
| <i>Shopping furniture</i> | 5.6 | 1.5 | 6 | 5.4 | 1.5 | 6 | 5.3 | 1.7 | 6 |
| <i>Virtual mirror</i> | 4.7 | 1.8 | 5 | 4.4 | 1.9 | 5 | 4.8 | 2.0 | 5 |
| <i>Street art</i> | 4.1 | 1.9 | 4 | 3.0 | 1.8 | 3 | 3.3 | 2.2 | 4 |

the statement about appropriateness of the services in the described situation, the Friedman test showed a significant effect of the scenario: $\chi^2_{F} = 124.4$, $p < 0.001$. The pair wise tests showed that the respondents rated the *Shopping furniture* scenario significantly higher than all the other scenarios considering the services’ appropriateness in the situation. In addition, the *Street art* was rated as lower in the same respect than all the other scenarios (see Table 3).

Regarding also the willingness to try out the services, the analysis showed a significant effect of the scenario on the rating: $\chi^2_{F} = 137.3$, $p < 0.001$. The pair wise tests showed that the respondents rated *On the bus*, *Jogging*, and *Shopping furniture* significantly higher than *Virtual mirror* and *Street art* scenarios (see Table 4). In addition, *Virtual mirror* was significantly higher than *Street Art* ($z = 4.998$, $N = 173$, $p < 0.001$).

In Section 3.2, the participants’ reasoning behind these numerical evaluations is contemplated with a qualitative analysis of responses to the scenario-specific open questions.

We analyzed also the effect of various background variables on the statement responses within each scenario. More specifically, we looked into the effects of gender and technology orientation, which are reported below. Other background variables (e.g. age, willingness to share information in web and experience in using different mobile technologies like camera and navigation) were also analyzed for this purpose. Surprisingly, these showed only occasional effects in some scenarios. For example regarding four different age groups (17–28, 29–40, 41–52, 53–64), there were sporadic differences between, for example, the perceived appropriateness of each scenario and in the overall UX statements in the end. However, the differences did not seem to follow any

Table 3

“The services would be appropriate in the described situation”: The significant pairwise test results (Bonferroni corrected significance levels)

| Row vs. column | Shopping furniture | Street art |
|-----------------------|--|--|
| <i>On the bus</i> | $z = 5.427$, $N = 160$, $p < 0.001$ | $z = 6.517$, $N = 195$, $p < 0.001$ |
| <i>Jogging</i> | $z = 5.505$, $N = 168$, $p < 0.001$ | $z = 5.097$, $N = 183$, $p < 0.001$ |
| <i>Virtual mirror</i> | $z = 6.706$, $N = 169$, $p < 0.001$ | $z = 4.276$, $N = 184$, $p < 0.001$ |
| <i>Street art</i> | $z = 9.396$, $N = 191$, $p < 0.001$ | |

Table 4

“I would like to try out the services”: The significant pairwise test results (Bonferroni corrected significance levels)

| Row vs. column | Virtual mirror | Street art |
|---------------------------|--|--|
| <i>On the bus</i> | $z = 5.228$, $N = 164$, $p < 0.001$ | $z = 9.059$, $N = 188$, $p < 0.001$ |
| <i>Jogging</i> | $z = 3.423$, $N = 155$, $p < 0.01$ | $z = 7.753$, $N = 178$, $p < 0.001$ |
| <i>Shopping furniture</i> | $z = 4.231$, $N = 155$, $p < 0.001$ | $z = 7.869$, $N = 184$, $p < 0.001$ |

trend; in some cases younger people agreed to the statements more than older people but in some cases it was the other way around. Hence, more detailed results regarding these background variables are left unreported.

Gender showed a slightly more convincing influence (Mann-Whitney U-test for independent samples). Overall, men ($N = 127$) were slightly more positive towards AR services than women ($N = 133$). Regarding *Jogging* men agreed more to “The services would be appropriate in the described situation” than women (Means 5.1/4.8, $p = 0.020$), and “I would like to try out the services” (5.7/5.3, $p < 0.001$). Regarding *Virtual mirror* men agreed more to “The services would be appropriate in the described situation” (5.0/4.4, $p = 0.025$), and for *Street art* “The services would be appropriate in the described situation” (4.3/3.7, $p = 0.013$), “The services would help me to reach my goals” (3.3/2.8, $p = 0.015$), and “I would like to try out the services” (4.2/3.3, $p = 0.001$).

The participant’s level of technological orientation was constructed from their agreement level to the following three statements: “I think that technology is necessary in my everyday life”, “I am among the first ones in my circle of friends who get new devices”, and “I help my friends and relatives using technical devices”. A sum variable with equal relative weights was created and participants were grouped

into the three groups: (1) Highly technologically oriented ($N = 120$; mean of the three ratings ≥ 5.5), (2) Moderately technologically oriented ($N = 105$; $3.5 \leq \text{Mean} < 5.5$), and (3) Slightly technologically oriented ($N = 31$; Mean < 3.5). Mann-Whitney analysis was conducted one to one between the groups, that is, 1 vs. 2, 2 vs. 3 and 3 vs. 1.

Table 5 summarizes the results of the comparisons between groups for each statement and for each scenario. Means are reported instead of medians to highlight the differences. U-values are omitted because the table is meant to give an overall summary of the analysis. Overall, a clear trend in the means can be seen: the more technologically oriented, the more the participant agreed with all the three acceptance statements. The Mann-Whitney tests also show significant differences in rather many pair wise comparisons. This indicates that the reasonably high technological orientation of the participants had a generally positive effect on the acceptance of the scenarios in this data. This is probably a result of their generally higher readiness to take new technologies in use, but perhaps also because of the less technologically oriented participants not being able to comprehend or identify with the idea behind AR.

It is worthwhile to note the relatively weaker effect of technological orientation in *Shopping furniture* (no significant differences), and partially also in *Virtual mirror* and *Street art*. *Shopping furniture* was overall the best rated scenario, and it might be that either the added value of AR in it or the overall usefulness and novelty of the scenario in general were seen greater than in others. Therefore, the less technologically oriented participants could also identify with the scenario and understand its usefulness or the specific value of AR in it. This explanation, however, is not as credible with *Virtual mirror* and *Street art* as they were less positively evaluated. *Jogging* demonstrated perhaps the most pervasive and distinct technological novelty (data glasses) that the most technologically oriented participants might have become fascinated about. This might explain why the technological orientation had the greatest influence on its acceptance ratings. Nevertheless, the discussion of the scenario-specific differences overall remains at a speculative level and therefore too far-flung conclusions should not be made.

3.2. Qualitative evaluation of scenario use cases

After the evaluative statements, two open questions were asked: “Which of the services mentioned

Table 5

Means (M1, M2, M3) and significance levels of the technology orientation comparisons. M1: highly oriented (N = 120), M2: moderately oriented (N = 105), M3: slightly oriented (N = 31)

| | On the bus | | | Jogging | | | Shopping furniture | | | Virtual mirror | | | Street art | | |
|--|------------|------------|------------|------------|------------|------------|--------------------|------------|------------|----------------|------------|------------|------------|------------|------------|
| | M1 1vs2 | M2 2vs3 | M3 3vs1 | M1 1vs2 | M2 2vs3 | M3 3vs1 | M1 1vs2 | M2 2vs3 | M3 3vs1 | M1 1vs2 | M2 2vs3 | M3 3vs1 | M1 1vs2 | M2 2vs3 | M3 3vs1 |
| The services would be appropriate in the described situation | 4.1 * | 4.9 * | 5.3 ** | 3.5 ** | 4.8 * | 5.2 *** | 5.0 | 5.7 | 5.7 | 4.2 | 4.7 | 4.8 | 3.4 | 3.8 * | 4.4 * |
| The services would help me to reach my goals | 3.5 * | 4.3 * | 4.8 ** | 3.3 ** | 4.5 * | 5.0 *** | 4.8 | 5.5 | 5.5 | 4.0 | 4.3 | 4.7 | 2.7 | 2.9 | 3.3 |
| I would like to try out the services | 4.6 | 4.8 *** | 5.2 *** | 4.2 | 5.0 *** | 5.8 *** | 4.7 | 5.4 | 5.5 | 4.5 | 4.5 ** | 5.1 | 3.2 | 3.5 * | 4.1 * |

*: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$

in the story you consider the most useful or the most entertaining? In what way?” and “Which of the services mentioned in the story you consider useless or disturbing? Why?” In the following, the percentages represent the proportion of answers that include the element in question, i.e. aspects that respondents explicitly brought up and were identified by the analyzers. Furthermore, one must bear in mind that several aspects could be mentioned in each answer, so the reported percentages do not add up to 100%.

Overall, the responses and arguments for evaluation opinions varied extensively. As typical in such large survey data, both risks and potential were seen in all aspects: the service features praised by some were at the same time considered pointless or risky by others.

3.2.1. On the bus

The most useful or entertaining features and content were related to the planning of the bus trip itself: getting information on routes (mentioned by 33%), timetables (10%) and current traffic jams and construction works (16%). “*The most useful are those focusing on the matter at hand: route, schedules, possible delays, and when a certain stop will come*” (Finnish female, 24). In addition, respondents appreciated the more general mobile and context-aware services: using the mobile to pay (23%) and being notified about the right bus stop (22%). “*Paying for the fare using the phone would save time (less queuing) and would be convenient if you didn’t have any change. It would also help people, especially children get home safely if they had forgotten or lost their money*” (British female, 22). “*A get off the bus notification, especially when in a strange place*” (Finnish female, 30). Most of these features were considered beneficial in everyday life but especially valuable while travelling in unfamiliar places.

With regard to the useless and disturbing aspects, the content offered in the bus was not that desired. Quite opposite, it was considered excessive information by many. The advertisements (29%), mobile games (25%) and bus specific services (13%) were considered somewhat obtrusive. “*The back-of-the-seat sticker stuff is annoying. I only have a short time outside the home/office every day, and I want to watch the weather, people, cars... the REAL reality*” (German male, 28). The ones that liked the idea of the content wanted to make the time on the bus of use and preferred informative content over entertainment: news were appreciated by 10%, games by 5% and event ads by 5%. On longer travels entertainment was however considered a pleasant addition. Some also thought that if the advertisements would lower the bus fares they could accept them. “*Ads. Although, if there was a way for me to get my fare refunded for watching the ads I might:)*” (American female, 40).

3.2.2. Jogging

In the jogging scenario, route signals (mentioned by 27%) and route information (16%) were seen useful especially in unfamiliar surroundings. “*Super helpful ‘coaching’ so I can concentrate on running, not on worrying about where I am. Even for a route that I know well it is helpful to know if some hindrance or improvement has been done since I last ran there.*” (American female, 62). Together with the night vision amplifier feature (valued by 14%) the route signals were thought to increase the sense of security by preventing from getting lost. Information about the available routes and suggestions shown along the way could inspire to take a new route also in a familiar environment. “*I would love to run on trails I’ve never run before and figure out the best path for me before/during my run*” (American female, 23). “*...the guidance through the glasses would make*

me feel confident of knowing where I'm going to" (Finnish female, 34). The physiological information like the heart rate was also regarded as valuable (25%).

The social features were the most commonly mentioned among the unnecessary and obtrusive things. Using the service to find a friend was considered handy by only 9% and reading and leaving comments was considered to add variety to running by 5%. Locating others was considered disturbing or useless by 21% and commenting by 20% of the respondents. In locating people respondents were worried mostly about anyone being able to locate them, and thus violate privacy. Furthermore, many of the respondents argued heavily against spoiling the exercise with any technology at all. Running was considered to be the moment when people intentionally want to be in peace and 'off-line'. "In jogging the pleasure is in finding new routes by yourself. That's what creates the joy of exploring and it cannot happen if you yield up the thinking to a machine" (Finnish male, 28). Respondents were also worried about the amount of information shown visually on the glasses and how this would disturb the running. Especially viewing others' comments on the glasses while running was opposed. "Jogging with glasses on could be a problem – I would rather glance down at a handheld object" (American female, 40).

3.2.3. Shopping furniture

The service was considered useful and valuable overall: 11% mentioned to value everything in the scenario. Especially modeling one's own home and utilizing the model in spatial conceptualization when buying new furniture were seen useful. "Checking the colors and the overall look seemed the most useful as they are so hard to picture in mind" (Finnish male, 26). Being able to virtually fit the new piece of furniture in the room was mentioned by 33%. "Living room in natural size: Wow! Visual feedback at its best!" (American female, 62). In addition, easily receiving the product information like price (8%), availability (6%) and color options (6%) with the mobile service were valued. "Getting up to date information and possibly customer ratings etc. is very valuable information that is nowadays hard to get" (German male, 30). Regardless, seeing the model in larger size was not considered that important as only 4% of the respondents specifically mentioned the projecting possibility.

The most remarkable negative aspect in this scenario seemed to be the complexity and toilsomeness of

use considering the separate projection room (mentioned by 11% of the respondents). "Projecting the image of the living room – cannot see why anyone would want to do that. Phone displays are good enough for that purpose" (Finnish female, 41). In addition, there were a variety of sporadic concerns, including that the projection of the model would not correspond to reality (2%), that the service could not replace the expertise of personnel needed in special cases, and the ecological concern that the service supports consuming instead of recycling. "Shopping furniture is very physical. In addition to visual looks, criteria include also the feel and quality of the surface material. A projector and small display will never be the same as feeling the sofa and listening to the sound generated when sitting" (Finnish male, 28). Some were also concerned about their privacy: if users can collect information about the products, then what information is collected about them at the same time by the service.

3.2.4. Virtual mirror

Although the service was not generally seen especially useful, many liked the idea and considered it entertaining. "Virtual make-over is a nice thought as a play. [...] Maybe an attraction in an amusement park? This could potentially be quite a laugh in a playful setting!" (Finnish female, 34). Especially the possibility to try out hairdos (14%) and different outfits (12%) were appreciated. Trying different hair styles is not possible in reality but only by the means of mixed reality and virtual fitting of clothes might speed up the buying process. "A possibility to virtually try on before deciding. Currently I do this e.g. with Photoshop, and it's rather troublesome;" (Finnish female, 27). Possibility to share the view (5%) was seen not only as an entertaining addition but also practical as it is common to ask for an opinion, for example from one's companion.

Most of the respondents' doubts concerned the lifelikeness of the virtual models (19%): how well the virtual model of one's figure corresponds to the real one, how well the fit of the clothes could be modeled, and how well the virtual hairdos would take into account the real hair quality of the user. "I would use this as a first step in order to cut down all the candidates that wouldn't fit / look good. Then I would try on the final candidates 'manually', in the traditional way. I'd assume that I have to double-check the fit anyway – I wouldn't trust the virtual service all the way" (Finnish female, 31). It was also seen that the physical fitting of clothes is an impor-

tant part of the shopping experience (12%). Some mentioned that they want to try how the clothes feel on them and some said that it would be easier and more pleasant to try them on for real. *“I doubt it would be the same as trying on real clothes, both visually and mentally. For some, the doing is important as well, even more important than reaching the goal”* (Finnish male, 29). Overall, virtual fitting was seen more useful in replacing catalogue shopping at home where you do not have the possibility to try the clothes on for real.

3.2.5. Street art

The service in the scenario was seen as entertaining and providing a way for self-expression (6%). *“They would bring a new experience in the subjective cityscape, and might entertain e.g. during a longer walk or while waiting for a bus”* (Finnish female, 24). *“The street art seemed an intriguing way of sharing artistic ideas. It has the elements of chance and discovery, which to me are important in art”* (Finnish male, 41). For some, the best things about the scenario were that the service might increase the tidiness of public spaces (11%) and the limited visibility of the graffiti (7%). *“Virtual art is a good concept in itself because it allows one to personalize the world around them without actually disturbing the environment for others. I like the idea of being able to see things that are placed only for you”* (American female, 23). There was, however, also interest in street art as the notifications about fine art works was considered a valuable feature. The social interactivity (valued by 15%), such as leaving comments and enriching the art were seen useful in other contexts. Some respondents did not see themselves using the service to express themselves with graffiti but to leave more informative messages for friends.

15% of the respondents considered the scenario trifling altogether. Some were not especially bothered by anything in the scenario but did not consider it interesting either. The physical interaction was again preferred by 8%. *“The real nature and hand-made things (such as real graffiti) are good as they are, mixing everything with digital world kind of destroys the thing”* (Finnish male, 28). Easy creation of virtual graffiti was seen as a threat by 8%, as it could lead to excessive information flood. If anyone could create graffiti most of them would most likely be of poor quality and the possibility to continue other’s art work could lead to ruining it. In addition, the limited visibility could also be misused, for example by leaving insulting messages or instructions how to break in to buildings. *“Possibility to secretly mess up oth-*

ers’ property and attach e.g. insulting comments” (Finnish male, 29).

3.2.6. Scenario evaluations overall

Overall, the open questions revealed that the most interesting scenarios both in Finnish and English survey data were considered *Shopping furniture*, *On the bus*, and *Jogging*. This is well in line with the quantitative results, and the responses to the qualitative questions well explain the respondents’ reasoning behind the numerical evaluations. The services in the above mentioned scenarios were regarded as the most relevant and practically useful in the given situation. *Virtual mirror* and *Street art* were considered less useful but to some degree more entertaining.

Furthermore, the open question at the end of the survey *“Which of preceding situations you found most interesting? Why?”* received somewhat congruent responds with the scenario-specific questions. Again, the first three scenarios gained the most positive reception. *On the bus* was mentioned in 35% of the responses. *Jogging* and *Shopping furniture* closely followed with 34% and 30%, respectively. *“Shopping furniture had a clear task and goals, so the information would be actually useful and not just passing-the-time type of junk just for technology’s sake”* (Finnish female, 30). *“In sports I think it’s easier to incorporate technological enhancements than in day-to-day situations, because running is an activity with a given duration. It feels more controlled somehow”* (Finnish male, 41). Expectedly, many respondents based their choice on the relevance of the scenario for their own needs and daily routines. *“On the bus, because I usually forget the bus card and it’d be suitable if I can pay through my mobile”* (Pakistani male, 25).

Virtual mirror was mentioned only by 12% and *Street art* by 10% of the respondents. *“Art is about exploration. It provides a soft landing to technology and can help in getting the feel of the new technology”* (Finnish female, 49). *“Virtual art is a good concept of itself because it allows one to personalize the world around them without actually disturbing the environment for others”* (American female, 23). *“The virtual mirror reached such things that are hard or impossible to do nowadays”* (Finnish female, 22).

Finally, 5 respondents mentioned *“none of the scenarios”* and 2 *“all of them”*, which well exemplifies the polarized opinions.

3.3. Evaluating augmented reality in general

After the scenario specific evaluations, a set of statements about the usefulness, user acceptance, and expected UX of AR were presented (Table 6). Overall, the effect of increasing the extent of available information from real life objects was highly desirable. This would create new possibilities for services and new experiences, and increase people's understanding of the nearby environment. Overall, the expectations of real AR services seem to be rather high.

However, most of the means deviate rather little from the center point of the scale here as well and the standard deviations are large. The distributions were more or less Gaussian but also the extreme values (1–2, 6–7) of the scale were much used. The differences between distributions of different statements were rather small as well. To speculate, this might signify that for some it was hard to form a clear opinion on AR in general – based on five such different scenarios and not yet having had actual experiences of functional AR services. Thus, those who had overall a strongly negative or positive attitude, might have reflected that attitude in each statement, and those who did not have a strong opinion to either direction, probably used mostly middle values in each statement. The probable reasons behind these evaluations are further illuminated in Sections 3.3.1 and 3.3.2.

Again, the effects of gender and technical orientation were analyzed. The other background variables showed none or only sporadic effects on the ratings, and therefore these results are not reported in more detail. Gender, however, showed a similar effect as in scenario-specific statements: men were slightly more positive towards the presented scenarios than women. They were more prone to think that services of AR like described in the stories would offer them a suitable way to express themselves (Means 3.9/3.5, $p=0.032$), support communication with their friends (4.3/3.8, $p=0.029$), increase communication with people they do not know (4.6/3.9, $p=0.001$), support their activities in different communities (4.5/4.0, $p=0.027$), and inspire them to produce and share information or art to others (4.4/3.9, $p=0.017$). They agreed more also with “I expect a lot from augmented reality services” (4.9/4.3, $p=0.010$). At the same time, women were more prone to think that services of augmented reality like described in the stories would require too much of their attention (4.1/4.6, $p=0.024$), threaten their privacy (4.0/4.6, $p=0.014$), and interfere with their understanding of what is real and what is virtual (3.1/3.8, $p=0.008$).

Table 6

Overall Means and Standard deviations of statements regarding AR in general; Means and significance levels of the technology orientation comparisons. M1: highly oriented (N = 120), M2: moderately oriented (N = 105), M3: slightly oriented (N = 31). Scale: 1–7 (completely disagree – completely agree)

| I think that services of augmented reality like described in the stories... | M / SD | M1 1vs2 | M2 2vs3 | M3 3vs1 |
|---|-----------|------------|------------|------------|
| would attach completely new meanings to places, services, products etc. | 5.3 / 1.6 | 4.3 * | 5.2 ** | 5.7 ** |
| would offer me experiences | 5.1 / 1.5 | 4.2 | 4.9 *** | 5.5 *** |
| would increase my understandability of my environment and its objects | 4.9 / 1.6 | 3.7 ** | 4.8 * | 5.3 *** |
| would bring significant benefits to my everyday life | 4.7 / 1.6 | 3.5 ** | 4.4 *** | 5.2 *** |
| would be captivating | 4.7 / 1.6 | 4.2 | 4.6 * | 5.0 * |
| I expect a lot from augmented reality services | 4.6 / 1.9 | 3.4 * | 4.3 *** | 5.3 *** |
| would delight me often | 4.6 / 1.5 | 3.6 * | 4.4 ** | 5.0 *** |
| would require too much of my attention (1) | 4.4 / 1.7 | 5.0 | 4.6 * | 4.0 ** |
| would support my activities in different communities | 4.3 / 1.6 | 3.2 ** | 4.2 * | 4.7 *** |
| would increase communication with people I don't know | 4.3 / 1.7 | 3.2 * | 4.1 ** | 4.8 *** |
| would threaten my privacy (1) | 4.3 / 1.9 | 4.7 | 4.7 *** | 3.8 * |
| would inspire me to produce and share information/arts with others | 4.1 / 1.8 | 3.1 * | 3.9 * | 4.6 *** |
| would support communication with my friends | 4.0 / 1.7 | 2.9 * | 3.8 ** | 4.5 *** |
| would help to inspire my creativity | 3.9 / 1.9 | 2.7 ** | 3.7 * | 4.3 *** |
| would help me to develop myself | 3.8 / 1.7 | 2.7 * | 3.6 ** | 4.2 *** |
| would offer me a suitable way to express myself | 3.7 / 1.7 | 2.6 * | 3.4 *** | 4.2 *** |
| would be difficult to use (1) | 3.6 / 1.6 | 4.0 | 3.6 | 3.5 |
| would interfere my understanding of what is real and what is virtual (1) | 3.4 / 2.0 | 3.8 | 3.8 ** | 3.1 |

*: $p \leq 0.05$, **: $p \leq 0.01$, ***: $p \leq 0.001$

(1) Notice that the statement is negatively phrased

Table 6 shows also the means of each technology orientation groups. The statements are in a decreasing order of overall mean – not in the original order of the survey. Again, Mann-Whitney analysis was

conducted one to one between the groups, and the results are summarized in the table. Similarly as with scenario-specific statements, we can see a clear trend in the technology orientation: the higher the orientation, the more positive ratings (higher acceptance). Based on the significance ratings, the effect in these overall statements is even more explicit – overall, surprisingly high. Statements like “...would support my activities in different communities” that could be said not to be directly benefitted from AR were also affected by the level of technology orientation. The reasons for this trend can be assumed to be based on the same aspects as with the scenario-specific statements: generally higher readiness of the technologically oriented to take new technologies in use and the less technologically oriented participants not being able to comprehend or identify with the overall idea and usefulness of AR or the scenarios overall.

3.3.1. Perceived strengths and values of AR

The general level open question “Which of the preceding situations you found most interesting? In what way?” revealed various aspects that were regarded as the benefits or drawbacks of augmented reality overall.

The utility in everyday life was seen as the most central value of AR: easier information retrieval, making tasks quicker to perform, saving time and effort. “*From such services I mostly want ease in my life, quick access to information, and saving time*” (Finnish male, 28). This often related to tasks that have to be done in any case, and are efficient to do with the help of AR. “*On the bus augmented an activity that I normally do a lot anyway*” (Finnish male, 36). On the other hand, AR was seen useful in unfamiliar environments where there often are plenty of needs for information about practical issues. “*Going to someplace I’ve never been before is always slightly nerve-racking for me, as I fear I will get lost and would not know enough to recognize my destination the first times around. Having a walking-navigation type unit leading the way would be very comforting*” (American female, 23).

AR was also seen to provide users with totally new interaction possibilities, as well as new perspectives to existing and previously familiar objects or places. “*It would be fun to walk down the street and see things that others can’t.*” (American female, 40). “*In jogging, getting information unobtrusively is difficult otherwise*” (German male, 30). “*Possibility to integrate the nice interface of a paper magazine and the*

dynamic content from the phone” (Finnish male, 36). The new perspectives and user-generated information content would create possibilities for a new type of communication but also for discovery and exploration. “*While jogging, it would bring new dimensions to even a familiar environment, and thus make jogging more exciting*” (Finnish female, 28). “*The idea of hidden messages appeals to me, as well as everyone being able to create them*” (Finnish male, 27). Also the motivating and inspiring effect of utilizing a new technology in one’s doings were brought up when considering *Jogging* and *Virtual mirror*. “*...I need to get into shape, and some kind of technological motivator like this could spark me to get started*” (Finnish male, 21).

In addition to the utility value, the way of interacting with AR services was seen comfortable and convenient. “*Seemed much more convenient and effortless than the traditional way of fitting clothes*” (Finnish male, 29). This was also emphasized in some comments about virtual glasses: “*The glasses interface is much more appealing than watching something through my mobile*” (Danish male, 31). “*The whole package would be totally sky-high. I’d go nowhere without the glasses!*” (Finnish male, 24).

A central factor in the utility value was considered timeliness and the information being real-time and up-to-date. “*At a bus stop getting real-time information would ease the pain of waiting :)*” (Finnish female, 25). This aspect of realistic content created also expectations of the integration of real and virtual: “*I’ve created models of furnishing by hand. This takes time but does not correspond to reality. An AR service for this sounds easy, fast, and truthful*” (Finnish female, 26).

3.3.2. Perceived drawbacks and risks of AR

The negative attitudes towards AR were most often manifested as the fear of virtual experiences and information replacing the real. There seems to be a thin line between merely enriching with virtual vs. the virtual taking over the existing tangible, auditory, and visual objects and activities. “*Seeing a hairdo beforehand could be handy but would take away the excitement of waiting for the result in a hairdresser. Why should all real experiences be removed and replaced with virtual?*” (Finnish female, 24). “*I want to be a human and make mistakes. I want to use measure tape and trust on intuition whether a piece of furniture is suitable to my apartment*” (Finnish female, 22).

Some comments related to a technophobic fear of losing autonomy. *“One should not rely too much on technology, but to manage with own senses and brains”* (Finnish female, 20). *“...leaving 98% of the real world visible, tangible, audible. AR should be there only when really needed or wanted by me, and even then around 70% of the real world should be there (behind the ‘transparent’ augmented reality)”* (Finnish male, 59). In addition, an anxiety of being able to be tracked or supervised by other people often came up.

The respondents were also worried about acquiring an excessive amount of information, thus being flooded with it. Several respondents mentioned that they want to receive only information that is meaningful to them in the current context and situation. Receiving information under the control of the user, as in some situations the user might not want or be able to receive any information at all. *“Too much information makes numb. Not everything visible is meant to be viewed through a camera”* (Finnish male, 32).

Additionally, the AR information was not trusted to be very realistic or of excellent quality. *“Like real graffiti is dirty and does not often blend into reality, probably much of virtual graffiti would also be very poor quality”* (Finnish male, 38). *“I’m pessimistic about the idea that touching tables would bring me reliable information about the stock or price development as those are not necessarily up-to-date even in the store’s current information systems”* (Finnish female, 25).

Users generating the AR content received also dubious thoughts regarding the long-term future. *“Would this become a new form of virtual pollution? Would I have to continuously block augmented spam from other people? Would there be an abusive virtual sticker in my back I would not even see?”* (Finnish male, 32). In addition, many expected that ads would sooner or later penetrate to become a central, unavoidable, part of the content, which would lower the contextual relevance of the content.

Finally, some criticized the camera-centered way of accessing and interacting with the AR content. Instead of using the ‘Magic lens’ and forcing the user to view the world through a camera, there should be also stationary displays or other ways independent of the position and heading of the mobile device. A comment on *On the bus*: *“Although I shoot photos a lot, it’d be odd to watch the world through a camera in order to receive information. Why is it not on an info display or transfer by other means to the mobile?”* (Finnish female, 49).

3.4. Expectations towards proactive and context-aware AR services

The previous section pointed out that most often users need be in control of the amount and type of received information. Such perception, however, was more or less expected in advance. Therefore, an additional question at the end of the survey addressed this aspect in detail: *“In what kind of situations you would like to get the information from your environment automatically? What would the information be related to?”*

Most often the mentioned situations in which such proactive and automatic features could be used were related to rather specific contexts and activities. 37% of the responses mentioned traffic, including both private passenger traffic and public transit. Participants suggested, for example, information about traffic jams, road and weather conditions, and changes in public transportation schedules. *“While driving the car it would be useful to have projected on the glasses (or directly on the windshield) info like hurdles on the road, distance from vehicles ahead of us getting too short, speed limit exceeded, etc.”* (Italian male, 32).

Using specific service modes based on which to filter the incoming information was often mentioned. In unfamiliar environments (brought up by 30%), some respondents were willing to receive e.g. navigation aids while travelling and automatic notifications about near-by sightseeings, cafés, and other points of interests for tourists. *“In terms of navigation, optimal routes, landmarks... in situations where the level of uncertainty is high”* (Pakistani female, 24).

17% mentioned shopping, including, for example, product recommendations and special prizes of browsed products, and reminders of what to buy. *“When buying furniture, I’d like to be automatically noticed if something I’m about to buy will not fit in my apartment, if it is cheaper elsewhere, or if it exceeds my budget!”* (Finnish female, 27). Some were even willing to receive ads under certain circumstances. *“Maybe while walking in the city I could select a certain shop brand to send ads to my phone when walking past a store selling those brands. Same for certain restaurants etc.”* (Finnish female, 33).

Social aspects, such as informing about near-by friends and their statuses, were also brought up by 5%. Up-to-date weather forecasts was mentioned by 3%.

Another distinct theme among the responses was exceptions and emergencies. This was explicitly mentioned by 7% but many other responses implicit-

ly included such elements as well (esp. traffic related). For example, possible dangers, crisis information, extra-ordinary events or arrangements in traffic, and other disruptions to the normal were mentioned. “*To warn me of danger. For the rest, I would like to call for the info or subscribe to it based on the situation*” (Finnish female, 34).

Regardless, 15% mentioned that they would not want to receive information automatically in any situation, and would prefer user-initiated information browsing or searching. “*I don’t want any device other than the fire alarm to notify me about surrounding information*” (Finnish female, 36). “*I’d rather define it in advance for each separate case*” (Finnish female, 22). These participants regarded AR as a tool for browsing and searching information in a user-initiated way rather than as part of a proactive and context-aware service.

4. Discussion and conclusions

When comparing the scenarios with each other, it seems that the most valuable mobile AR services were those demonstrating pragmatic usefulness for the user, e.g. by saving time and effort. This became evident both in scenarios-specific evaluations and in general AR evaluation. For example, AR content was seen to offer the user rich and contextually relevant information for ad hoc needs in the daily mobile life.

The first three of the scenarios (*On the bus*, *Jogging* and *Shopping furniture*) were considered the most utilitarian and practically useful, whereas the last two were regarded more as entertainment-related. Naturally some respondents had doubts also concerning the three first scenarios but in general the expected benefits seemed to outrun the drawbacks in the overall evaluation. *Shopping furniture* brought the most additional value to the described situation by providing practical benefits that cannot be achieved with current mobile services. Some answers reflected the fear of uncontrollable information flood, i.e. not wishing anything more than the most important information at the very moment. Most participants wanted to receive only such information that is meaningful in the current situation and context. This tendency might have caused more negative attitudes towards entertainment content and ads. The same tendency was visible in the needs for proactively provided information: pragmatically useful information about unexpected issues in the ambient environment was the best accepted.

With regard to differences between genders and differences between the three technology orientation groups, we could identify some interrelations. In qualitative questions, female respondents explicated their opinions more than male. Statistical tests showed a slightly more positive attitude within men than women.

More importantly, the expectations were highly dependent on the technology orientation: the more technologically oriented, the more positive the scenario evaluations were, and more willingly they would accept AR services in use. Highly technologically oriented people regarded most of the scenarios to be more appropriate and were more willing to try out the services than the less technologically oriented. The underlying reason might be that the highly technologically oriented put emphasis on the pure novelty value of a new technology whereas less technologically oriented assess more the suitability of the technology in their daily activities and needs. The least technology orientation dependent scenario was *Shopping furniture* in which no significant differences between the groups could be found.

The overall statements about acceptance of AR showed a similar effect of the technological orientation, and even more extensively: 17 out of 18 statements measuring the acceptance and expected user experience showed at least some differences between comparisons of the technological orientation groups. Due to its overarching effect, the technological orientation might partially explain also the differences between genders. However, no further analysis of such interactions between background variables was carried out as it was not the main focus of this research.

4.1. Reflecting the results

Mobile AR technology provides an extensive design space for new services. The extent of information that could be represented and accessed by means of AR in the mobile domain is practically unlimited. Likewise, so is the extent of potential use cases, contexts, and user groups. AR has the potential to revolutionize the way in which information is accessed and presented to people and enables people to integrate their own skills with the power of ubiquitous computing. Therefore, the grounds for evaluating the scenarios were also diverse and a complex whole.

The results show that the respondents considered the possibilities and risks of the services also outside the described contexts and use purposes. For some it

seemed to be natural to ponder also the long-term influences and how such services would affect and become affected in a larger, socio-cultural context. Furthermore, it seems that even small changes in the service can affect the overall judgment (e.g. accepting ads if they lower the price of another thing). These observations can be said to further justify the approach of studying user acceptance and expectations of mobile AR services.

Regarding *Street art* and *Virtual mirror*, it seemed that most respondents could not relate to the user groups in question; they did not share the described motivation to use the service for the described purpose. Overall, it seemed that the respondents' ability to identify with the actors in the scenario affected their answers. There were also more doubts related to how the services in the last two scenarios would actually function.

Similar results of users favoring pragmatic elements prior to the actual use have been found also by Kurniawan et al. [23]. Their focus groups pointed out that each of the physical design elements was chosen for their practicality rather than for their aesthetic and other non-instrumental values. Furthermore, for example Hsee et al. [15] and Hassenzahl [13] have shown that for people it seems easier to justify the expenses for something practical compared to something hedonic (phenomenon called "lay functionalism"). This bias in human thinking most probably occurs in situations where people have little time to ponder. This might have also affected the results of this survey, where the respondents formed and stated opinions in 15–20 minutes. Therefore, the respondents' general orientation and attitude towards novel technologies might have affected the evaluation results rather much – instead of the results describing the respondents' truly thought through views on the scenarios.

The results reflect also other technologies and domains than AR: the identified aspects are not only technology-specific but also relate to the context of use and use case. Therefore, some of the requirements and expectations can be valid also in other mobile and ubiquitous technologies. For example, the scenarios manifested also simple examples of context awareness and ambient intelligence as part of the AR service (e.g. proactive features). Such automatism strongly polarized opinions. On one hand, too proactive features caused concern because they might interrupt the user excessively, provide excessive amount of information, include advertisements, and lead to losing the user's sovereignty. On the other hand, some respondents would value automatically

receiving information about traffic, price discounts, and extra-ordinary events or arrangements.

The additional value of AR as a way of visualizing and interacting with information was not assessed separately as much as expected. Especially most negatively regarded comments related to the use case and functionality or the information content itself – not to how the information in the example was accessed and interacted with. It became evident that the respondents did not judge the services based on which technology they use but based on their functional and experiential value to the user.

One of the most AR-specific results was that some respondents were worried about how well the virtual representations correspond to the reality. As virtuality rarely excels existing interactions and conventions in reality, the reality should not be replaced by virtual correspondents. Instead, augmentation should be utilized for such purposes that are not possible in real life. Another clear expectation was the timeliness and information being real-time, up-to-date, authentic, and realistic. Some participants expected the AR information to realistically match and correspond to its real world counterparts, while some considered the very same thing as the uppermost challenge.

4.2. Methodology revisited

Studying expectations towards technology that does not exist yet is unquestionably a challenging task. The responses are inevitably influenced by the examples indicated in the scenarios, and hindered by people speculating on what they would do with technologies they do not yet have. The language-based scenarios describe and evoke a metaphorical experience, which can hardly stand in the place of actual usage. In other words, the responses here may have different emphasis after taking the technology in use, and additional issues or perceived benefits can still be discovered. Despite this challenge, using example scenarios was necessary for elucidating the hard-to-understand topics for our participants and to be able to gather potential users' perceptions of AR technology a priori.

Naturally, the results need to be understood as suggestive and contextualized in a certain type of user population. The survey respondents were mostly technically savvy and largely represented urban, well-educated people with backgrounds in, for example, science and research. For example *Street art* and *Virtual mirror* represented more artistic use cases than the others, and therefore might have received a more positive acceptance in a different user popula-

tion. In addition, acceptance of each scenario depends on to what extent the activities and tasks in the scenario are relevant to the respondent. For example, acceptance of the *Jogging* scenario can be expected to be affected by the respondent's interests and habits related to jogging. Naturally, all such background variables of the respondents' lives and behaviour could not be measured. Such underlying factors were tried to be minimized by selecting scenarios that are generally understandable.

Gaining a truly holistic picture of the acceptance of various service types, not to mention AR in general, with five scenarios is unrealistic. Nevertheless, the aim was to gain understanding as extensively as possible in a very early phase of development by utilizing an approach not before used in AR research. Any information about potential users' valuations and judgments – despite the slight questionability of future validity – is important when still exploring what kind of services to develop.

In retrospect, some decisions regarding the approach and study setup might have affected the overall validity and reliability of the results. The scenario order could have been randomized to counterbalance its possible effect on the evaluations. Especially regarding the Likert statements, the subjective scale of the participants may have changed from scenario to another as new aspects were introduced. However, here the goal was not only to put the scenarios in order of superiority but also to understand the role and acceptance of mobile AR in general. If the scenario order would have been randomized, the overall narrative of AR through the scenarios would have become different. Thus, different respondents could have received different overall pictures of what mobile AR is and what it makes possible.

Regarding *Virtual mirror*, few responses to the English survey indicated that as a service it was not as well understood as the others (e.g. wrong assumption of needing to buy the mentioned data glasses in advance). In addition, the use cases in *Virtual mirror* and *Shopping furniture* can be seen to be more directed to women. However, no clear statistical differences in the evaluations between genders were found in this regard.

The answers of the English version of the survey showed a slightly more positive attitude towards mobile AR. In the evaluative statements there were some sporadic, yet significant, differences between the responses to Finnish and English versions of the statements (not reported in detail). However, no specific themes or consistency was noticed. With such a small sample of cultures, the reason remains unex-

plained. It might be that the differences are due to higher level of technological orientation of the participants answering to the English version, or merely small differences in the nuances of the two language versions.

4.3. Concluding remarks and future research

This paper presented an online survey study aimed at gaining an understanding of potential users' expectations of AR and to evaluate specific use scenarios that demonstrate various aspects of mobile AR services. The main contribution is the variety of qualitative and quantitative user evaluations and the participants' views into the acceptance of mobile AR services. The scenarios elicited an extensive set of positive aspects and advantages, needs and expectations, as well as challenges and perceived risks.

Including both favorable and unfavorable aspects, the results show an ambivalent attitude towards future AR services. Such understanding of users' expectations and acceptance issues possesses a high potential to catalyze service success, and therefore can be utilized as design targets in development. Despite the earlier research on usability issues in AR, the community has been lacking research results with a focus on user acceptance and user experience.

An additional contribution and novelty is methodological: applying a scenario-based approach and online surveys in soliciting the perceived use of potential users of AR. Albeit the scenarios were rather simple and only textual, the online survey served well its purpose to evaluate the scenarios, and AR in general, in a multifaceted way. Despite the novel concept of AR, potential users' expectations and requirements could be gathered with a relatively simple research setup and before enforcing service implementation efforts. The novelty value of the results can be considered high as very little user research on the user acceptance and user expectations based on AR service concepts has been published.

In general, our future studies will focus on investigating how functional mobile AR applications or prototypes get accepted and are appropriated into use. It is to be seen what will be the true value of AR as an interaction paradigm and what kind of behavioral and societal implications the technology might have on people. For example, proactivity and automatism as much discussed aspects of AR services open up an interesting track of further research: in what kind of situations would such features be truly accepted when considering also the social context and possible long-term implications?

Parts of the scenarios have been lately implemented in new web or AR applications, e.g. virtually placing models of clothes on top of a webcam view of oneself and virtual furniture fitting. It would be highly interesting to evaluate the real experiences evoked from the use of those applications, and to contrast such findings with the expectations of mobile AR services presented here. Lately, a similar evaluative approach has been applied to the recently introduced mobile AR applications like Layar, Google Goggles and Junaio [30]. A cross-analysis between the expectations and the actual user experience and acceptance could further point out interesting issues with regard to how people are able to envision the actual use in a pre-use setting like with the scenarios in this paper.

Acknowledgements

The research was supported by the Finnish Funding Agency for Technology and Innovation (TEKES) under the project DIEM (grant 40223/07). We highly appreciate the participation of all the respondents of the survey. For fruitful collaboration and comments on paper drafts we thank Pirita Ihamäki, Timo Partala, Tuomas Vaittinen, and Minna Kynsilehto.

References

- [1] M. Aittala, Inverse lighting and photorealistic rendering for augmented reality, *The Visual Computer* **26**(6–8), 669–678.
- [2] L. Arhippainen, Studying user experience: Issues and problems of mobile service – Case Adamos: User experience (im)possible to catch?, Ph.D. Dissertation, University of Oulu, 2009.
- [3] R. Azuma, Tracking requirements for augmented reality, *Communications of the ACM* **36**(7) (1993), 50–51.
- [4] J. Burrell, G. Gay, K. Kubo and N. Farina, Context-aware computing: A test case, in: *Proc. of UbiComp 2002*, Lecture Notes in Computer Science, Vol. 2498, 2002, pp. 647–653.
- [5] A.D. Cheok, X. Yang, Z.Z. Ying, M. Billinghurst and H. Kato, Touch-Space: Mixed reality game space based on ubiquitous, tangible, and social computing, *Personal and Ubiquitous Computing* **6**(5–6) (2002), 430–442.
- [6] A. Dey, Understanding and using context, *Personal and Ubiquitous Computing* **5**(1) (2001), 4–7.
- [7] A. Dünser, R. Grasset and M. Billinghurst, A survey of evaluation techniques used in augmented reality studies, in: *Proc. of International Conference on Computer Graphics and Interactive Techniques, ACM SIGGRAPH (2008)*, Article No.: 5.
- [8] K. Finstad, Response interpolation and scale sensitivity: Evidence against 5-point scales, *Journal of Usability Studies* **5**(3) (2010), 104–110.
- [9] J.L. Gabbard and E. Swan, Usability engineering for augmented reality: Employing user-based studies to inform design, *IEEE Transactions on Visualization and Computer Graphics* **14**(3) (2008), 513–524.
- [10] M. Gandy, R. Catrambone, B. MacIntyre, C. Alvarez, E. Eiriksdottir, M. Hilimire, B. Davidson and A.C. McLaughlin, Experience with an AR evaluation test bed: Presence, performance, and physiological measurement, in: *Proc. of Mixed and Augmented Reality (ISMAR) 2010*, pp. 127–136.
- [11] T. Gleue and P. Dähne, Design and implementation of a mobile device for outdoor augmented reality in the ARCHEO-GUIDE project, in: *Proc. of 2001 Conference on Virtual Reality, Archeology, and Cultural Heritage*, 2001, pp. 161–168.
- [12] M. Hassenzahl and N. Tractinsky, User experience – A research agenda, *Behaviour and Information Technology* **25**(2) (2006), 91–97.
- [13] M. Hassenzahl, M. Schöbel and T. Trautmann, How motivational orientation influences the evaluation and choice of hedonic and pragmatic interactive products: The role of regulatory focus, *Interacting with Computers* **20**(4–5) (2008), 473–479.
- [14] I. Herbst, A.K. Braun, R. McCall and W. Broll, TimeWarp: Interactive time travel with a mobile mixed reality game, in: *Proc. of MobileHCI 2008*, pp. 235–244.
- [15] C.K. Hsee, J. Zhang, F. Yu and Y. Xi, Lay rationalism and inconsistency between predicted experience and decision, *Journal of Behavioral Decision Making* **16** (2003), 257–272.
- [16] J. Häkkinä, Usability with context-aware mobile applications, Ph.D. Dissertation, University of Oulu, 2006.
- [17] T. Höllerer and S. Feiner, Mobile augmented reality, in: *Telegeoinformatics: Location-Based Computing and Services* H. Karimi and A. Hammad, eds, Taylor & Francis Books Ltd., 2004.
- [18] ISO FDIS 9241-210:2009, Ergonomics of human system interaction – Part 210: Human-centred design for interactive systems (formerly known as 13407), International organization for standardization, ISO.
- [19] J. Jain, R. Ghosh and M. Dekhil, Multimodal shopping lists, in: *Proc. of Human-Computer Interaction*, 2009, 39–47.
- [20] E. Kaasinen, User needs for location-aware mobile services, *Personal and Ubiquitous Computing* **7** (2003), 70–79.
- [21] E. Kaasinen, User acceptance of mobile services – Value, ease of use, trust and ease of adoption, Ph.D. Dissertation, VTT Publications 566, 2005.
- [22] E. Kruijff, E. Swan and S. Feiner, Perceptual issues in augmented reality revisited, in: *Proc. of Mixed and Augmented Reality (ISMAR) 2010*.
- [23] S. Kurniawan, M. Mahmud and Y. Nugroho, A study of the use of mobile phones by older persons, in: *Proc. of Conference on Human Factors in Computing Systems 2006*, pp. 989–994.
- [24] P.J. Ludford, D. Frankowski, K. Reily, K. Wilms and L. Terveen, Because I carry my cell phone anyway: Functional location-based reminder applications, in: *Proc. of the SIGCHI Conference on Human Factors in Computing Systems, 2006*, pp. 889–898.
- [25] W.E. Mackay, Augmenting reality: A new paradigm for interacting with computers, *La Recherche* (1996).
- [26] P. Milgram and F. Kishino, A taxonomy of mixed reality visual displays, *IEICE Transactions on Information Systems* **E77-D**(12) (1994), 1321–1329.
- [27] A. Morrison, A. Oulasvirta, P. Peltonen, S. Lemmelä, G. Jaccucci, G. Reitmayr, J. Näsänen and A. Juustila, Like bees around the hive: A comparative study of a mobile augmented reality map, in: *Proc. of the SIGCHI Conference on Human Factors in Computing Systems, 2009*, pp. 1889–1898.
- [28] S. Nilsson, Augmentation in the wild: User centered development and evaluation of augmented reality applications, Ph.D. Dissertation, Linköping University, 2010.

- [29] T. Olsson, P. Ihamäki, E. Lagerstam, L. Ventä-Olkkonen and K. Väänänen-Vainio-Mattila, User expectations for mobile mixed reality services, in: *Proc. of 4th European Conference for Cognitive Ergonomics 2009*, pp. 177–184.
- [30] T. Olsson and M. Salo, Online user survey on current mobile augmented reality applications, in: *Proc. of Mixed and Augmented Reality (ISMAR) 2011*.
- [31] Z.-R. Peng, M.-H. Tsou, Internet GIS, *Distributed Geographic Information Services for the Internet and Wireless Networks*, John Wiley & Sons, Inc, 2003.
- [32] P. Remagnino and G.L. Foresti, Ambient intelligence: A new multidisciplinary paradigm, *IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans* **35** (1) (2005).
- [33] E.M. Rogers, *Diffusion of Innovations*, 4th ed, Free Press, 1995.
- [34] M. Rohs, Marker-based embodied interaction for handheld augmented reality games, *Journal of Virtual Reality and Broadcasting* **4**(5) (2007).
- [35] H. Sato, I. Kitahara and Y. Ohta, MR-mirror: A complex of real and virtual mirrors, in: *Proc. of Virtual and Mixed Reality 2009*, pp. 482–491.
- [36] M. Satyanarayanan, Pervasive computing: Vision and challenges, *IEEE Personal Communication* (2001), 10–17.
- [37] A. Schmidt, Ubiquitous computing – Computing in context. Ph.D. Dissertation, Lancaster University, 2002.
- [38] E. Swan and J.L. Gabbard, Survey of user-based experimentation in augmented reality, in: *Proc. of 1st International Conference on Virtual Reality*, 2005.
- [39] J. Vallino, Interactive augmented reality, Ph.D. Dissertation, University of Rochester, 1998.
- [40] V. Venkatesh, M. Morris, G. Davis and F. Davis, User acceptance of information technology: Toward a unified view, *MIS Quarterly* **27**(3), 425–478.
- [41] D. Wagner and D. Schmalstieg, Making augmented reality practical on mobile phones, part 2, *IEEE computer Graphics and Applications* **29**(4) (2009), 6–9.
- [42] M. Weiser, The computer for the 21st century, *Scientific American* **265**(3) (1991), 94–104.
- [43] P. Wellner, W. Mackay and R. Gold, Back to the real world, *Communications of the ACM* **36**(7) (1993), 24–26.