Lappeenranta University of Technology
School of Industrial Engineering and Management
Degree Program in Computer Science

Joanna Kalalahti

DEVELOPING USABILITY EVALUATION HEURISTICS FOR AUGMENTED REALITY APPLICATIONS

Examiners: Professor Jari Porras
Master of Economics Janne Paavilainen

Supervisor: Professor Jari Porras

TIIVISTELMÄ

Lappeenrannan teknillinen yliopisto
Tuotantotalouden tiedekunta
Tietotekniikan koulutusohjelma

Joanna Kalalahti

Lisätyn todellisuuden sovellusten käytettävyysheuristiikan kehittäminen

Diplomityö

2015

68 sivua, 11 kuvaa, 10 taulukkoa, 6 liitettä

Työn tarkastajat: Professori Jari Porras
KTM Janne Paavilainen

Hakusanat: lisätty todellisuus, käytettävyys, heuristiikka
Keywords: augmented reality, usability, heuristics

Koska lisätyn todellisuuden sovellusten käytettävyyden arviointiin ei ole olemassa yleistä heuristiikkaa, sen kehittäminen otettiin tämän diplomityön tavoitteeksi. Heuristiikan kehittäminen tapahtui vaiheittain. Kirjallisuuskatsauksen pohjalta muodostettiin alustava käytettävyysheuristiikka, jonka neljä asiantuntijaa evaluoi. Lopputuloksena syntyi kuusi arviointikriteeriä: 1) vuorovaikutustavat ja kontrollit, 2) virtuaalisten objektien esittäminen, 3) virtuaalisten objektien ja reaaliympäristön suhde, 4) virtuaalisiin objekteihin liittyvä informaatio, 5) soveltuvuus käyttökontekstiin ja 6) käytön fyysinen miellyttävyys. Heuristiikkaa on tarkoitus käyttää yhdessä Nielsenin (1995) yleisen käytettävyysheuristiikan kanssa. Heuristiikka ei ole vielä valmis käytettäväksi, sillä sen toimivuus tulee vielä testata käytännössä.

ABSTRACT

Lappeenranta University of Technology
School of Industrial Engineering and Management
Degree Program in Computer Science

Joanna Kalalahti

Developing usability evaluation heuristics for augmented reality applications

Master's Thesis

68 pages, 11 figures, 10 tables, 6 appendices

Examiners: Professor Jari Porras
Master of Economics Janne Paavilainen

Keywords: augmented reality, usability, heuristics

There is no generic usability heuristics for Augmented Reality (AR) applications, thus, the aim of this thesis was to develop one. The development of the heuristics was carried out in phases. Based on a literature review, a preliminary version of the heuristics was developed, which was evaluated by four experts. As a result, six evaluation criteria were formed: 1) interaction methods and controls, 2) presentation of virtual objects, 3) relationship between virtual objects and real world, 4) information related to virtual objects, 5) suitability for the usage context and 6) physical comfort of use. The heuristics should be used with Nielsen's (1995) generic usability evaluation heuristics. The heuristics are not ready to be used as such, since it must still be tested in practice.

ACKNOWLEDGEMENTS

This master's thesis is accomplished partly related my previous work as a researcher in University of Tampere, School of Information Sciences in the Spring 2014, in a project funded by Tekes (DIGILE, Digital Services program, WP3 Education, Collaborative and Rapid Prototyping of Educational Games ProGE). I thank for the 2nd examiner of this thesis, Janne Paavilainen, who has also given valuable advices for me during the course of this work, also professor Frans Mäyrä has offered his support for my work. I also thank my ex-colleagues Erika Tahnua-Piiroinen, Ismo Rakkolainen and Yrjö Lappalainen from University of Tampere who have offered me a lot of help in accomplishing this thesis. Also many people in Finland and elsewhere in the world who are working with and interested in the use of Augmented Reality have given me inspiration and shared their expertise with me deserve to be thanked for. Thanks also go to the supervisor and 1st examiner of this thesis, professor Jari Porras in Lappeenranta University of Technology.

I also want to give thanks to my current workplace Police University College for the support in finishing my MSc. studies. My friends and family, most importantly my husband Matti for his valuable help, patience and support, despite of his own hurries for example with the housebuilding project.

In Tampere 31.12.2014

Joanna Kalalahti

TABLE OF CONTENTS

LIST OF SYMBOLS AND ABBREVIATIONS

AR	Augmented Reality
AV	Augmented Virtuality
CVI	Content Validity Index
ITSM	IT Security Management
MR	Mixed Reality
RAD	Rapid Application Development
2D	Two-dimensional
3D	Three-dimensional
VE	Virtual Environment
VR	Virtual Reality

1 INTRODUCTION

Even though Augmented Reality (AR) is not a new technology, the development of it has been technologically oriented, and very little attention has been paid to HCI (Human-Computer Interaction) issues. Still, these issues would be very important for the development of applications which are experienced as usable and at which satisfy the needs of the users. No generic usability evaluation heuristics exists, creation of which has been taken as the aim of this master's thesis. The background, further refined goals, limitations and structure of this work are presented in this chapter.

1.1 Background

AR originates from 1960's or even earlier depending on the viewpoint. Still, it has made a breakthrough and gained more attention within the consumers only during the last five years because of the increase in the computing power and the use of mobile devices, and the expectations are high towards the soon-to-be available data glasses like Google Glass (Google Glass).

Despite of the long history, AR as technology is technically immature in some respects. The applications are not necessarily functioning in and optimal way — for example, registration and tracking problems exist. In order to develop well-working and high-quality AR applications, technical development is still required. On the other hand, strong technology-orientation has caused development of applications disconnected from the users and usage contexts. Thus, user requirements, usability of the applications and user experiences have not been considered enough, and there is a lack of AR applications which would be useful and user-centered (Olsson 2012, 45−46; Dünser et al. 2007).

There have been some attempts to evaluate usability of AR applications by using the existing, generic usability heuristics which, unfortunately, are not quite capable of reaching all the essential and specific features of AR. Some application specific usability heuristics for AR applications have also been developed. An evaluation heuristics is needed, which could take into account the special requirements of AR. Even though it is a challenging task because of the several platforms and types of AR applications, it has been seen something worth persuing. (Dünser et al. 2007, 37–38.)

1.2 Research task, goals and limitations

At the moment there is no generic usability evaluation heuristics for AR applications, and for this reason, it has been taken as the goal of this master's thesis. Since the heuristics should be generic enough to allow evaluation of the wide variety of AR applications, the challenge is to operate on level which is concrete enough to allow useful evaluation results.

Main features of AR as a technology, already known problems causing usability issues, heuristics developed for the near-fields of AR (Virtual Environments, VEs and 3D user interfaces) must be taken into account. These issues are studied in a literature review, which is mainly based on writer's familiarity with the research in the domain of AR. Based on the literature review, a preliminary version of the heuristics is formed. Because AR heuristics should be used with generic usability heuristics of Nielsen (1995), the developed heuristics will concentrate of the issues specific to AR.

The heuristics developed need to be validated to see if it is actually useful for the purpose it was developed. The validation carried out in this work is based on expert evaluation and the insight of the researcher. Only a small amount of experts is used, which restricts the use of statistical methodology in validation of the heuristics. The development of the heuristics will follow the steps of the heuristics development framework of Rusu et al. (2011), including phases of validation and refinement of the heuristics based on the feedback of the evaluators. The idea is to evaluate the relevance and validity of the heuristics, and as a result of this work, a priori validated version of the heuristics is developed. Still, it must be emphasised that the heuristics are not ready for use as such. The effectiveness of the heuristics must be validated afterwards by testing their applicability for the evaluation of an AR application, which was left out of the scope of this work.

This work is not the first effort of developing AR heuristics. The value of this work can be seen to be in its aim of developing a first version of a compact set of generic heuristics, which might be easy to use in evaluation of AR applications. Also an effort to validate the heuristics a priori to the experimental, a posteriori validation was made. The generated AR heuristics should not be seen as a final version, rather, a beginning which might be improved and used with different kinds of AR applications. Finalising the heuristics will require combined effort from several AR and usability experts.

1.3 The structure of the research

Introduction for the work is given in chapter 1. Theoretical framework consisting of usability, heuristic usability evaluation, AR as a technology and its usability considerations is presented in the chapters 2–3. The methodology used in development of the heuristics is presented in chapter 4. Chapter 5 presents the development process of the heuristics. Discussion about the work is in chapter 6 and conclusions are drawn in chapter 7.

2 USABILITY AND HEURISTIC EVALUATION

Usability is a part of user-centered design. Several methods for evaluating the usability are developed, one of them is heuristic evaluation. It is accomplished by using certain principles called heuristics as a help. Heuristics can also be used as guidance when applications are developed. This chapter describes the basics of usability and heuristic evaluation, which will form the baseline for the development of AR usability evaluation heuristics.

2.1 Usability

An application that is easy and quick to use can be said to be usable. According to ISO standard (ISO 9241–11 1998, 2) usability is the *"Extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use."* Another and quite similar definition of usability of Nielsen is about the question of how well users can use the functionality of the system. (Nielsen 1993, 24). Usability is not a single, one-dimensional property of the user interface, instead, it is formed of several components. Traditionally the attributes of learnability, efficiency, memorability, errors and satisfaction are associated with usability. (Nielsen 1993, 26.)

It is important to make a distinction between utility and usability. According to Nielsen (1993, 25), the utility can be defined as whether the functionality of the system in principle can do what is needed (e.g., learn with the system), and usability concerns how well users can use that functionality (e.g., use the learning environment to learn) (Fig. 1).

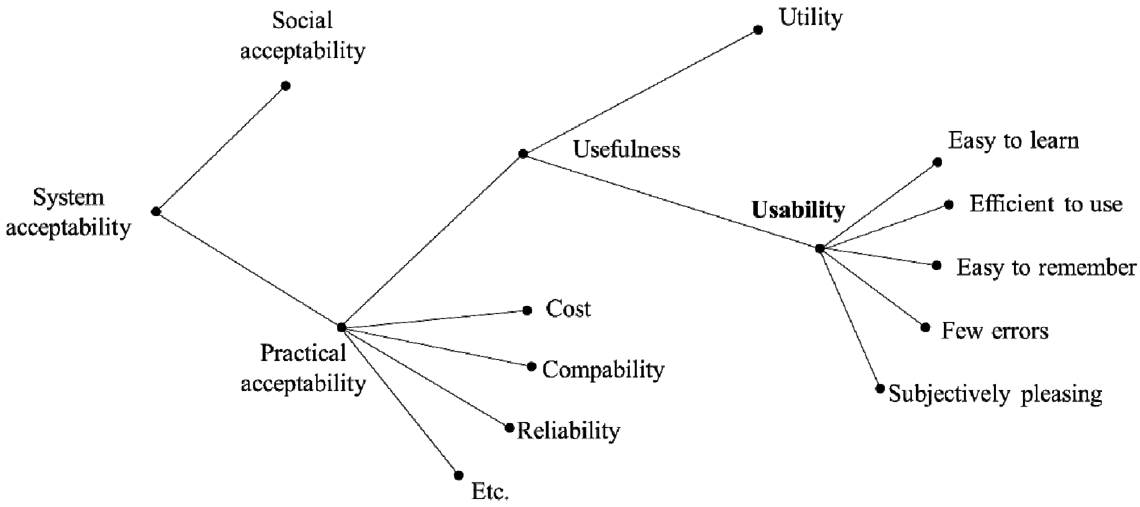


Fig. 1. A model of attributes of system acceptability (Nielsen 1993, 25).

Utility of applications is an important aspect, but the evaluation of utility of AR has been left out of the scope of this thesis and the starting point is a situation where the user of AR application has already been convinced about its utility.

Principles according to which usability is evaluated are called heuristics. One of the most known generic heuristics is developed by Nielsen and Molich (1990, 339) and further refined by Nielsen (1995). It consists of ten different criteria. According to the heuristics, a usable application should:

- Show the system status to the user and give feedback within reasonable time
- Match with the real world by speaking the user's language with familiar words, phrases and concepts and present information in a natural way
- Allow user control and freedom for example after mistakes when choosing system functions
- Be consistent and follow standards and platform conventions
- Prevent errors
- Minimize user's memory load by allowing possibilities for recognition rather than recall and offer instructions which are easily retrievable whenever appropriate
- Be flexible and efficient to use for novices as well as experts
- Be aesthetic and minimalist by design
- Support error recognition of user made mistakes and recovering from them

- Contain help and documentation (even though the system should be ideally be easy enough to use without them).

Since Nielsen’s heuristics is regarded suitable for generic usability evaluation, it will be used in this thesis with the AR heuristics. It has been already used as such in evaluation of AR application, and it was possible to detect usability problems with it (Dünser 2007, 37). Since Nielsen's heuristics has been developed originally for evaluation of web-pages, new devices and technologies may require more tailored and fine-grained heuristics (Rusu et al. 2011, 59).

Usability guidelines contain well-known principles for user interface design, and they are also used for usability evaluation heuristics. The guidelines can vary according to their abstraction level. General, high-level guidelines are applicable to all user interfaces, while more detailed, low-level guidelines are applicable to certain individual products. The low level would contain components dealing with the perceptional issues, and the high level be more focused on the interaction techniques and input devices. (Nielsen 1993, 91–92; Bowman et al. 2002, 409; Dünser & Billinghurst 2011, 295–296.)

The amount of usability guidelines can vary from a few to thousands, but a large amount of the guidelines is experienced as intimidating by the evaluators. Since the idea of heuristic evaluation is to find the most obvious and typical usability violations, there need not be so many criteria. It is typical for experienced evaluators to know the heuristics by heart and keep them all in mind while evaluating the application. If the list seems to be too short and there is a fear of the evaluators missing important details and accomplish the evaluations in a too rough and abstract level, descriptions of the different criteria can be used to help the evaluator to focus on typical issues and usage situations. For these reasons, as short list of heuristics as possible with more detailed descriptions will be the aim of this work. (Nielsen 1993, 155; Nielsen & Molich 1990, 249; Dünser & Billinghurst 2011, 296.)

2.2 Heuristic evaluation

Usability can be evaluated using different methods, of which one is heuristic expert evaluation. Heuristic evaluation is a systematic inspection for a user interface design for usability. The fundamental idea of heuristic evaluation is that experts go through the user interface according to usability heuristics, detecting violations against the used heuristics and the severity of them. The method is efficient, easy to learn and carry out, and it is also quite cost-effective, since only a few evaluators are needed to carry out the evaluation usually completely on their own. Heuristic evaluation is typically carried out in the development phase of the application and focused on the prototype version. It is used to detect the most coarse usability problems before proceeding in to the more detailed level in application development. (Nielsen & Molich 1990, 249; Nielsen 1993, 155–160.)

The cost-benefit ratio of heuristic evaluation has been found to be very high. Usually at least three evaluators are suggested to be used to in heuristic evaluation. As Fig. 2 illustrates, if three evaluators are used to carry out the evaluation, 60% of the usability violations can be found. (Nielsen 1993, 155–156.)

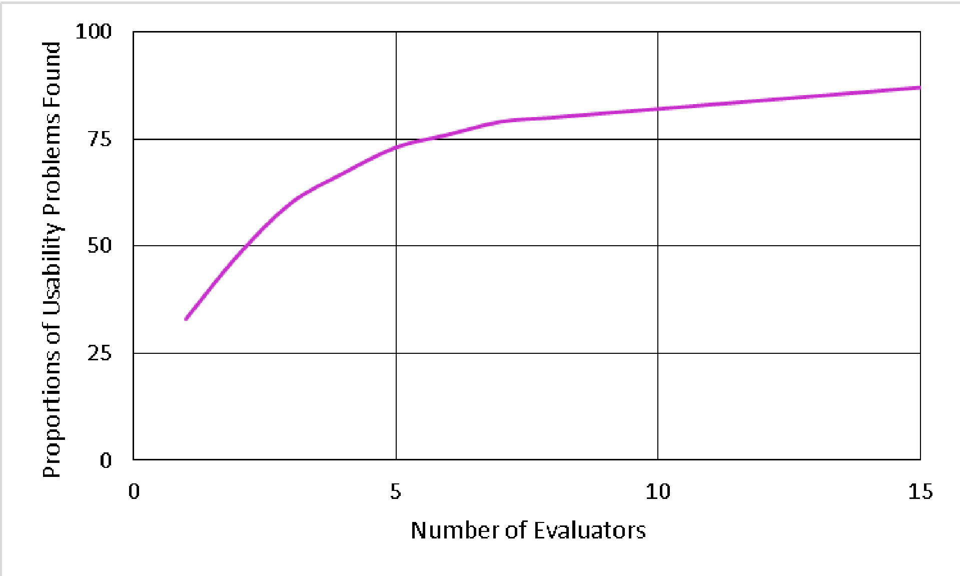


Fig. 2. Usability problems found by heuristic evaluation as a function of the number of evaluators (Nielsen 1993, 156).

The heuristics developed in this thesis are meant to be used in heuristic evaluation of AR applications. Also, the suggested next phase of the development of the AR heuristics (a posteriori validation), which is left out of the scope of this work, can be carried out using heuristic evaluation based on the use and comparison of two separate heuristics.

3 AUGMENTED REALITY

Augmented Reality (AR) is a technology which allows the environment to be explored in real time through different displays with integrated computer-generated content. This chapter presents the main features of AR and the most common problems of the technology which should be considered when AR applications are developed, especially from the viewpoints of usability. Also, the usability evaluation heuristics developed in the near fields of AR and application-specific usability heuristics developed for AR are presented.

Literature review is mainly based on the literature writer has studied as a researcher in the domain of AR during 2010–2014. Searchers have been made in the reference databases such as Web of Science and Scopus using the keywords related to the topic of the thesis, e.g. "augmented reality" AND (usability OR "user-centered design").

3.1 Features of AR as a technology

According to the definition of Azuma (1997, 356) AR application should qualify three criteria:

1. Combine real and virtual views
2. Allow real-time interaction
3. The objects augmented should be aligned accurately and registered in 3D (three dimensions).

A more loose definition may also allow 2D (two-dimensional) objects, if they are registered in 3D — for example, text tags placed within a building about which they are giving additional information (Bowman et al. 2005, 389). According to a definition of Specht et al. (2011, 117), AR is a system which amplifies visual, auditory or tactile senses digitally, making things which are not naturally observable visible. Augmented objects may also be audio files according to some of the definitions (Bowman et al. 2005, 389; Mariette 2013, 11–12). This is an important addition, since it has been observed that especially in place-based AR applications additional information which is presented with audio improves the usability of the applications (McCall et al. 2011, 34). Also, for the viewpoint of accessibility, audio files are important for several user groups[1].

One illustration often used with the definitions of AR is virtuality continuum of Milgram and Kishino (1994) (Fig. 3). Real environment and virtual environment form the two dimensions of the continuum. AR is situated nearer the real environment, since the basis for activity lies in real environment when AR is concerned.

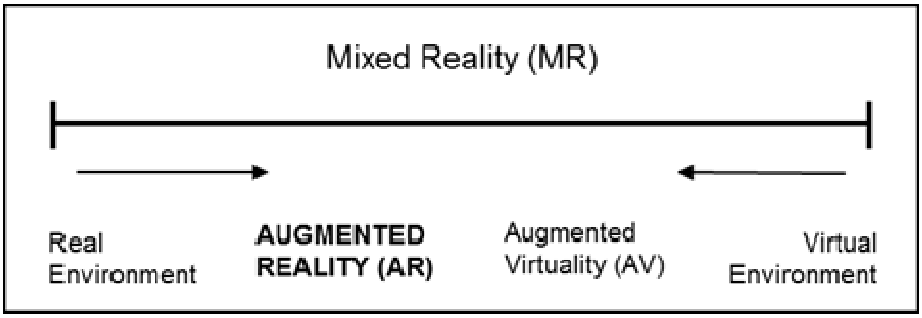


Fig. 3. Virtuality continuum (Milgram & Kishino 1994).

The continuum has also been presented in a way that the area of Mixed Reality (MR) consists only of AR and Augmented Virtuality (AV)[2] (Wang & Dunston 2009, 5). The continuum illustrates very well the discreteness of the limits between real and virtual environments and the applications remaining in between them, referring to the stricter and looser definitions of AR presented earlier. A loose definition for AR is applied in this work, e.g. also 2D objects registered in 3D and also other than visual augmentations are regarded as AR.

AR as a technology dates back to 1950–60s, and to be more accurate, the ideas behind AR can be seen to be from the beginning of early 1900 (Carmigniani & Furht 2011, 4; Wagner 2013). When the processing power of computers has increased and the use of mobile devices has become popular, AR applications have become more popular amongst consumers around the world.

[1] An application called BlindSquare (BlindSquare) is developed for visually impaired, which provides audio information about the targets near the user.

[2] Augmented Virtuality can be defined as a virtual environment which is connected to the real environment for example through movements of user's body when steering the avatar.

To be able to understand different requirements and challenges for AR application design and to develop well-working AR applications, understanding is needed about the background technologies and devices of it, which are multiple. Carmigniani & Furht (2011, 9–14) and Wang and Dunston (2009, 12–23) have classified technologies and devices behind AR[3], and a combination of their classifications gives a thorough view of them (Table 1):

Table 1. AR technology overview (based on Carmigniani & Furht 2011, 9–14 and Wang & Dunston 2009, 12–23).

<table>
<tr><td colspan="2" rowspan="2">Content and media types presented with AR can be classified in a continuum...</td><td colspan="2">...from abstract (text)...</td></tr>
<tr><td colspan="2">...to more realistic (picture, video, three-dimensional contents)</td></tr>
<tr><td rowspan="7">Controls:</td><td>Input mechanisms can be classified in a continuum...</td><td>...from two-dimensional control devices (e.g. traditional graphical user interfaces and typical control devices like keyboard, mouse and possible game-controls)...</td><td>...to more intuitive three-dimensional and tangible user interfaces (touch-screen, data gloves, wristband, phone as a pointing device, gaze control, gesture control).</td></tr>
<tr><td rowspan="6">Output mechanisms can be classified as:</td><td colspan="2">Monitor displays like traditional computer display or bigger spatial screen</td></tr>
<tr><td colspan="2">Handheld displays like smartphone and tablet computer displays</td></tr>
<tr><td colspan="2">Head mounted displays (HMD) from data helmets to eye-glass and contact lenses based lighter displays which are becoming common at the moment</td></tr>
<tr><td colspan="2">Spatial displays / Spatial Augmented Reality (SAR)</td></tr>
<tr><td colspan="2">Audio output (device loudspeaker or headphones)</td></tr>
<tr><td colspan="2">Haptic output</td></tr>
<tr><td colspan="2" rowspan="3">Technology behind the displays can be classified in a following way:</td><td colspan="2">Video see-through: display device also contains a video camera filming the environment of the user and integrates the augmented objects beforehand to the video displayed with a very short delay. This kind of displays are typical in monitor displays and handheld displays, also spatial displays can use the technology.</td></tr>
<tr><td colspan="2">Optical see-through: objects augmented are integrated to the display with a see-through mirror in real-time. This kind of displays are typical in head-mounted displays.</td></tr>
<tr><td colspan="2">Projector displays: special cases of monitor displays with larger device area like whiteboard or pictures projected on a surface. Also different see-through technologies can be used like holography and fog screens. This kind of displays are typical in spatial AR and they make multi-user applications possible. Also head mounted displays have used projector displays with smaller projectors.</td></tr>
<tr><td colspan="2" rowspan="4">Tracking technologies used with AR applications (i.e. technologies used to align the augmented information with the environment):</td><td colspan="2">Image recognition (digital cameras)</td></tr>
<tr><td colspan="2">Place-based recognition (GPS, compass)</td></tr>
<tr><td colspan="2">Other, rarely used sensors like optical sensors, inertial sensors like accelerometers and gyroscopes, magnetic sensors, acoustic sensors, other wireless sensors</td></tr>
<tr><td colspan="2">Hybrid sensors which are combinations of different sensors.</td></tr>
</table>

[3] Even though Wang & Dunston (2009) use the term Mixed Reality, the issues they discuss about apply to AR as well, and the term AR is used in this work when referring to them.

Computers i.e. data processing units:	Traditional computer with an application running on the computer
	Distant devices over the internet
	Mobile devices (like smartphones and tablets)

From the viewpoint of user-centered design and usability, the classification has a central meaning. The area of application and the goal of the activity, usage environment and its requirements should be analysed carefully in order to select the technologies which best fit to the requirements. In this way, it is easier to make sure the pre-requisites for the use of the applications are met. Wang and Dunston (2009, 24–42) recommend task analysis and linking it with the technology selection already when considering the use of AR and the development of the applications. The best possible format for the presentation of contents should be selected to minimize the cognitive load. The selection of input mechanism is connected with the usage context and task — for example, it must be considered if user's hands are free or reserved for the task itself. Display technologies are also connected with the usage environment and its requirements, for example, lighting conditions and the need for co-operation with other users need to be concerned. Different tracking technologies work in a different way in different environments, e.g. inside and outside. Also their accuracies differ and must be taken into account in each situation. Wang & Dunston (2009, 20) illustrate different requirements in a following combination (Fig. 4):

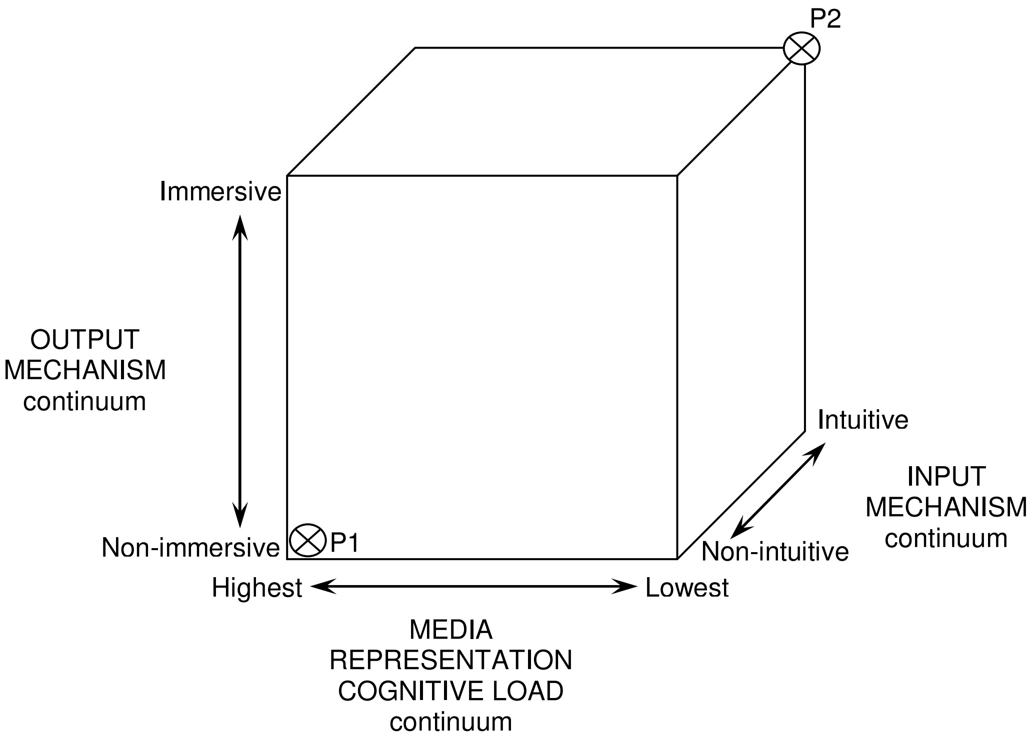


Fig. 4. Mixed Reality global continuum (modified from Wang & Dunston 2009, 20).

The closer the point P1 is, the more the user needs to accomplish mental transitions in using the application, and the closer the point P2 is, the fewer mental transitions are needed (Wang & Dunston 2009, 20).

Because the basis of AR is strongly in the real environment, it gives a good hint of what kind of contexts and usage situations AR is most appropriate. Carmichael et al. (2012, 1768) have distinguished a few clear criteria for assessing the utility of AR:

- The relationship of virtual objects and real environment must be clear and meaningful: *"When reality doesn't play a prominent role in the application, it is difficult to make a meaningful connection between virtual and real objects."*
- When context-relevant meaning must be offered to virtual information, AR will also prove to be useful.
- AR is useful when it is critical to remain the attention of the user in the task without splitting it elsewhere.
- AR is useful also when natural user interfaces and direct manipulation of the object are strived for.

Wang and Dunston (2009, 26–28, 35) have presented a quite similar classification about the benefits of AR as a support for construction, manufacturing and engineering work. According to them, the benefits of AR are mostly connected with situations of information processing, which are a central part of all manual tasks (cf. Neumann & Majoros 2008, 4–5). When cognitive components are integrated as a part of manual work, the accomplishment of tasks is enhanced and speed up, because:

- AR minimizes the costs of accessing theoretical information (e.g. information search and internalisation).
- The problem of split attention between cognitive and manual component of the task can be avoided with AR, when the theoretical information needed is integrated as a part of manual task.
- Cognitive information connected with physical contexts can be integrated with AR and ease the memorization of things.

3.2 Typical problems of AR and application design recommendations

Some of the commonly appearing problems in AR applications seem to be registration and tracking errors. Real objects and virtual objects must be properly aligned with respect to each other to create an illusion of the coexistence of the two worlds, which is called registration. Errors in registration can be divided into two types (Azuma 1997, 372–379):

- Static registration errors, which appear even though the user's viewpoint and the objects remain completely still. Static errors are caused by optical distortion, errors in the tracking system, mechanical misalignments and incorrect viewing parameters.
- Dynamic registration errors, which appear when the viewpoint of the user or the objects move. Dynamic errors are caused by system delays or lags. The end-to-end system delay is defined by Azuma as *"the time difference between the moment that the tracking system measures the position and orientation of the viewpoint to the moment when the generated images corresponding to that position and orientation appear in the displays"*.

Registration requires accurate tracking of the user's and surrounding object's position in relation to it. Accurate tracking systems, greater input variety and bandwidth and longer range are needed. (Azuma 1997, 383–386.) Hybrid tracking systems have been used to compensate the weaknesses of separate tracking technologies, and it is expected that future AR systems will be common (Wang & Dunston 2009, 22; Dünser et al. 2007, 40).

Problem with occlusion is brought up in many studies and articles. Occlusion deals with the depth perception and it occurs when real objects appear in front of the virtual objects even though they should appear behind them. Occlusion handling is used to enhance the illusion of virtual objects appearing as a part of the real environment, and it is important for a correct spatial perception about the relationships of the objects and possibly to prevent physical issues like eyestrain and motion sickness. (Tian et al. 2010, 2886; Kruijff et al. 2010, 6.)

An example of proper registration, tracking and occlusion handling is presented in Fig. 5, where virtual eyeglasses appear to be real, since the application recognises the position of eyes very well when the user looks to the web camera. Eyes seem to be behind the lenses and when watched from different viewpoints, the glasses adjust to them very well, as well as the movement of the head of the user without any lag.

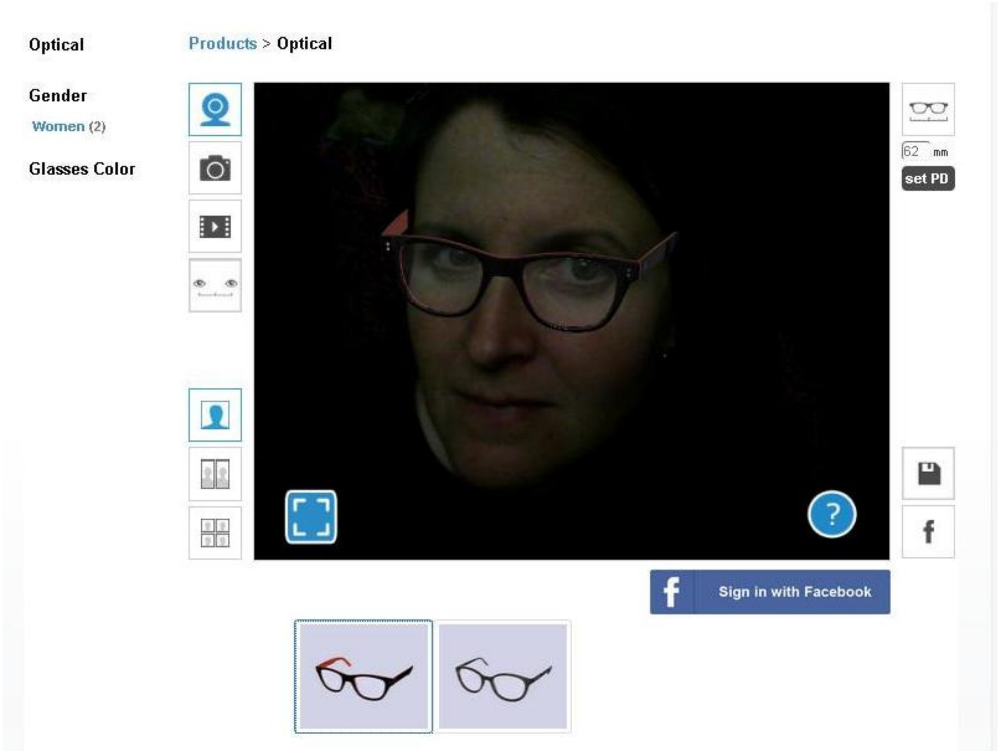
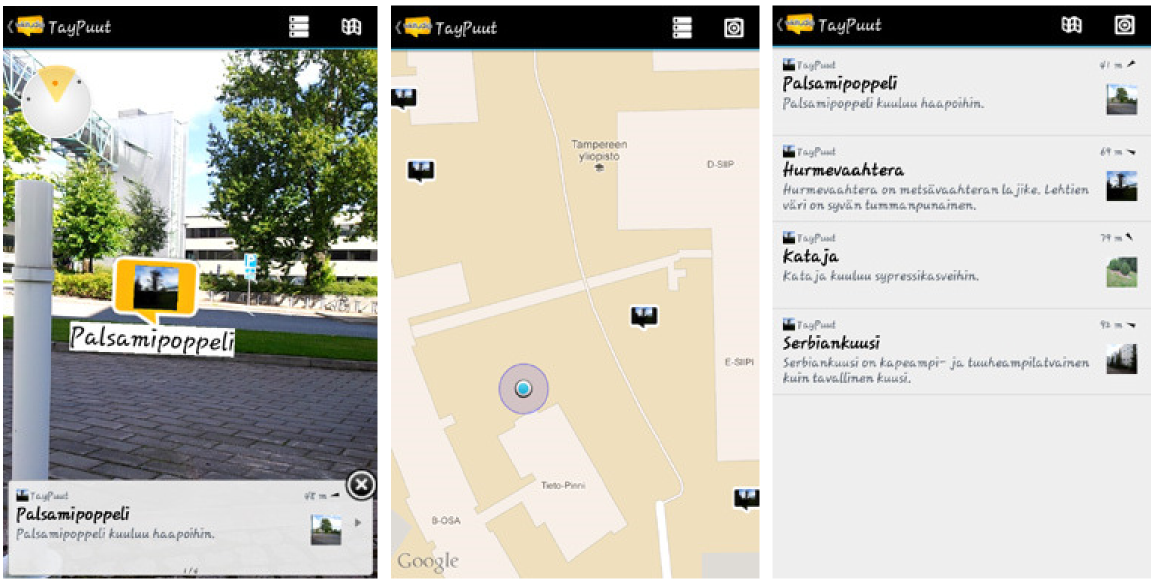


Fig. 5. Accurate registration, tracking and occlusion handling (TryLive).

Li & Duh (2013, 110) present cognitive issues which are important from the viewpoint of human-centered design of mobile AR applications. Three central categories concerning the cognitive issues in mobile AR interaction are information presentation, physical interaction and shared experience. Ganapathy (2013, 177–179) has presented design principles for mobile AR, which are in many respects similar to the presentation issues presented by Li & Duh. Bowman et al. (2001, 98–103) discuss about specific issues concerning the 3D interaction methods of VEs, but some of them are also typical interaction tasks for AR applications: wayfinding, selection, manipulation and system control. According to Li & Duh (2013, 116), typical interaction methods in AR applications are navigation, direct manipulation and content creation. These issues are discussed in more detail.

According to Bowman et al. (2001, 100), wayfinding is a cognitive part of the navigational task, and the other component, moving, does not apply to AR applications since in AR the user is not moving in a virtual environment. In wayfinding, the user must be aware of her own position, objects around her, spatial relations between them and expectations about the future status of the environment. User must be able to change the perspective from egocentric camera view to exocentric map view (Fig. 6). In addition to the issue of different perspective taking, the issue of how smoothly the user can change the attention between the AR environment and real environment is important. The ability to deal with the transitions from real to virtual which are encountered in different levels. The user should be able to transfer the knowledge from AR application to the real world. Different parts of the environments should be identifiable and the user should be able to relate them to other parts, for example, when looking outside the camera view to the real environment and then back to the device. Also real-world wayfinding principles should be transferrable to the usage of AR applications. (Bowman et al. 2001, 98–100; Li & Duh 2013, 116–117.)



(a) (b) (c)

Fig. 6. Different perspectives in AR applications: a) camera view b) map view and c) list view of a location-based AR application (Wikitude).

Manipulation, especially object selection, resembles system control techniques in some respects (Bowman et al. 2001, 102), and it is considered with all interaction methods in this work. Since the direct hand manipulation is a major interaction modality in natural physical environment, it should also be applicable in AR environment. Also the selection methods which are familiar from traditional graphical user interfaces could be used if possible. Direct manipulation is at its best as natural as possible. Tangible user interfaces allow direct interaction with the target. If manipulation by hands is supported, it may be challenging if also 2-dimensional interaction methods are used. If a normally two-dimensional task becomes three-dimensional in an AR application, it reduces the effectiveness of traditional interaction technique. Combining different input and output modalities gives new possibilities for different situations, but also requires skills to combine them so that the whole interface is well-functioning. Different modalities and their strengths and weaknesses must be utilised according to the requirements and opportunities of the situations. Also used interaction methods should be replaceable and switched according to the context (Ganapathy 2013, 179). There should be a balance between the different interaction methods — when and how they are used regarding to other interaction methods available. The transition from one to another should be as smooth as possible. When the system control is considered, the user should be able to change the state of the system, which may include the selection of an element, and changing the mode of interaction. According to Ganapathy (2013, 177), critical for the user is to receive feedback on actions user has committed and identifying that the application is in proper state to accomplish the action. (Li & Duh 2013, 118–121; Bowman et al. 2001, 100–104.)

User-generated content is getting more and more common in AR applications and it enriches the user experience. Typically user-generated content is added in physical environment and it contains text, image and audio. It is challenging and important to position the information in the required place. Different viewing perspectives may help the process. Adding content is difficult when the user is on the move, and different techniques like freezing the views have been applied. (Li & Duh 2013, 122–123.)

When information presentation is considered, several issues need to be considered (Li & Duh 2013, 112–116; Ganapathy 2013, 177–179):

- The amount of information (too much or too little). In Fig. 7, there is too much information visible, which makes it difficult to study the environment through the display.

– The relation of information and its background (e.g. contrast between the augmented text and the background, also when different backgrounds are used). In Fig. 7, the background of the text is grey, which allows the texts to be separated from the background, but it is more difficult to read white text in the bright sunshine, and the text labels conceal the background.

Fig. 7. A crowded view with inaccurate registration of objects and labels (Wikitude).

– The form of information presented (text, image, 2- or 3-dimensional) affects how strongly the virtual information is experienced as part of the physical reality. In Fig. 8, the dinosaur seems to integrate to the physical reality quite well, because the presentation of it is three-dimensional and for example shadows have been used to create an illusion of an even more realistic appearance. Also the example of the virtual eyeglasses in Fig. 5 is an example of well accomplished integration.

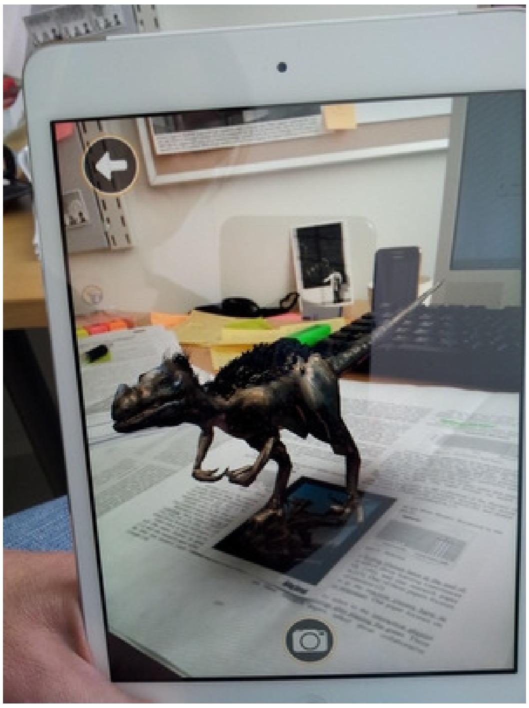


Fig. 8. A 3D model of a dinosaur which appears to be part of its surroundings because of shadows and accurate registration (Dinosaurs — Live!).

- How clear the textual information is (e.g. font that is easy to read).
- Positioning and placement of the virtual information: taking into account the overlaps (items of interest should not be obscured), occlusion of objects, proximity of virtual information and the physical object connected with it (should not be too big), depth cues. Cf. Fig. 7 in which the text labels conceal quite a big part of the physical reality, and the virtual objects are not properly aligned near the physical objects (different trees and plants) they are connected with. On the other hand, in Fig. 8 and Fig. 5, virtual information is accurately aligned with the marker and the background is visible properly from the back of the legs of the dinosaur, exactly in a similar way it should be if the model of the dinosaur was physical.
- Organization and grouping of the information: there should be a possibility to filter the received information, distinguishing different icons and information based on them without reading the text label (e.g. the category of presented objects, visibility of objects and distance between objects). In Fig. 9, the user may filter the information based on its proximity.

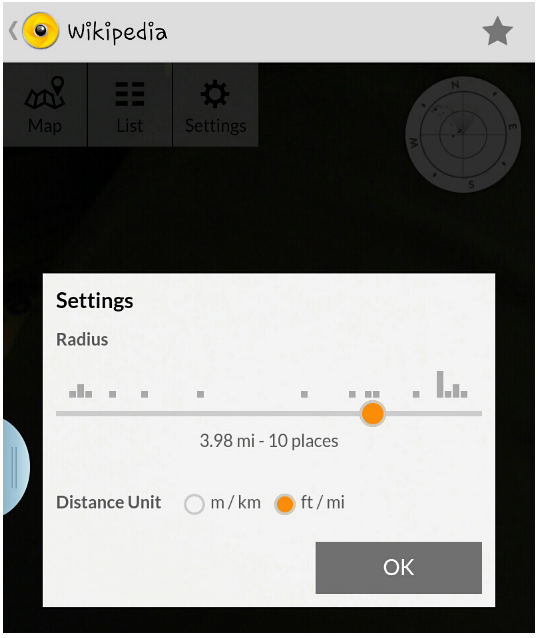


Fig. 9. Information filtering according to distance (Wikitude).

- Identifying how relevant the information is for the user: important information requiring action needs to be identified easily, especially when medical or learning applications are concerned.
- Different views to the information should be offered (general, detailed, zooming in and out, ego- and exocentric) since the user should be able to study the object from different perspectives. Different views in wayfinding tasks were presented in Fig. 6, different perspective-taking possibilities in a 3D modeling application are presented in Fig. 10:

(a) (b)

Fig. 10. Different perspectives in AR applications: a) front view and b) upper view of an image-based AR application (Viking Shoe).

3.3 Usability of an AR application

As it is usually the case with emerging technologies, user requirements, usability of the applications and user experiences of AR applications have not been concerned enough. There is a lack of AR applications which are useful and user-centered. (Olsson 2012, 45−46; Olsson 2013, 203-204; Dünser et al. 2007, 37). Information technology research and advisory company Gartner's hype cycle for emerging technologies (Gartner 2014) illustrates very well the situation of AR in this respect (Fig. 11). AR will soon reach the bottom of the curve and next couple of years from this will tell if it will meet the expectations claimed for it and reach the plateau of productivity. This depends strongly on the usefulness and user-centeredness of AR applications, of which, usability is one aspect.

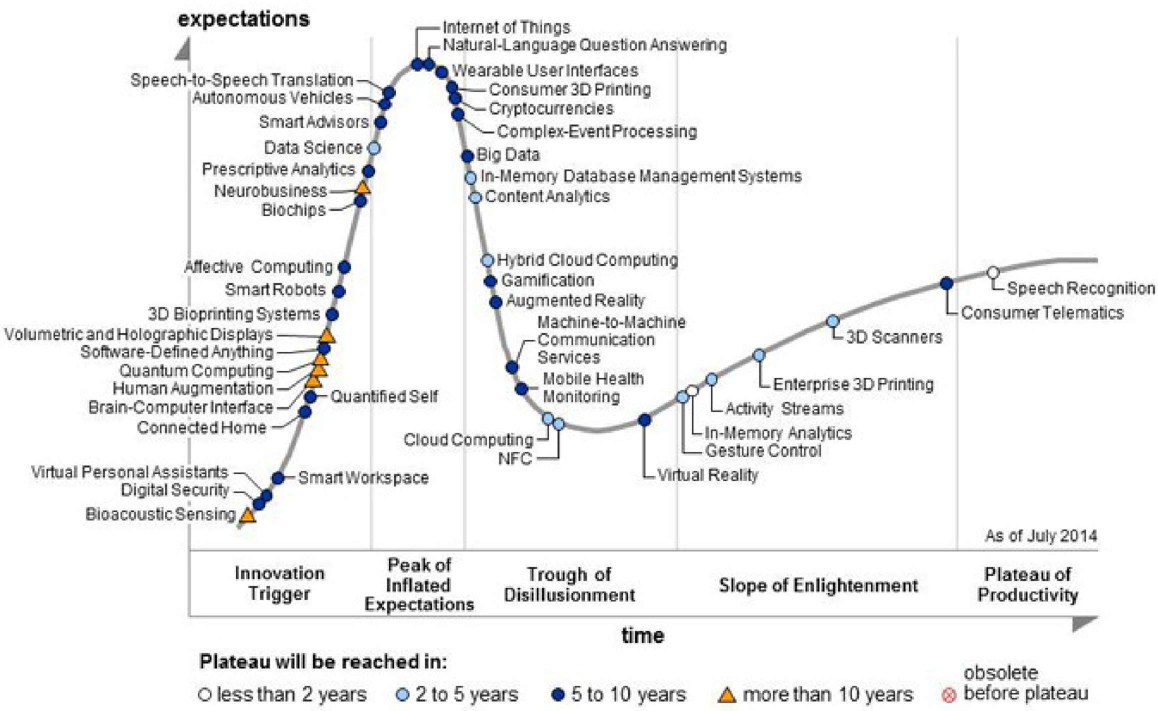


Fig. 11. Gartner's 2014 Hype Cycle for Emerging Technologies (Gartner 2014).

Dünser et al. (2008, 2) and Dünser et al. (2007, 37) refer to a literature survey of Swan & Gabbard (2005) indicating that in 2004, there were only 14% of AR-related articles in the leading journals and conferences which addressed an aspect of human computer interaction. There is no standardized or generally accepted usability heuristics for AR applications. As mentioned earlier, applying generic usability heuristics (e.g. Molich & Nielsen 1990, 339; Nielsen 1995) for evaluating AR applications has already been tried out, and some usability problems have been detected with it (Dünser et al. 2007, 37). Still, special requirements of AR must be taken into account and heuristics developed for traditional user interfaces are not enough for evaluating all the interaction techniques used in AR applications. Especially locating, selecting and manipulating objects in 3D space are missing from Nielsen's heuristics, also input and output modalities can be different for AR interfaces. Increasing user's effectiveness and efficiency may not always be the primary goal of some AR applications, instead of providing a novel user experience. (Dünser & Billinghurst 2011, 292, 297.)

Some of the most promising attempts to develop heuristics for AR applications and the generic criteria derived from the existing AR application-specific heuristics are presented in this chapter and used as a skeleton of the generic AR heuristics. Also the literature review in chapter 3.1 is used to identify what is specific for AR as a technology, and known problems and suggested design principles for AR applications (Chapter 3.2) can be used to inform the development of the heuristics. Identifying some common denominator from VE and MR tasks could be applied in more application-specific or high-level task analysis level, and for this reason, usability heuristics developed for VE applications are also studied (cf. Träskbäck 2004, 39).

When developing usability evaluation heuristics for AR applications, it is important to understand that AR applications are very diverse, used with multiple devices, displays, interaction techniques and user interfaces, as was pointed out in the chapter 3.1. Multitude must be accepted as a starting point, since it is likely that there will never be just one standard user interface for AR applications, as in the case of traditional computers. On the other hand, fast technological development of devices may change the situation very quickly. (Dünser et al. 2007, 37–38; Bowman et al. 2002, 409.)

The situation for AR applications is a lot like with virtual environments (VEs) twelve years ago, described by Bowman et al. (2002, 409): there are no interface standards or good understanding of the usability of various interface types. For this reason, applying design principles developed for specific usage contexts is not necessarily a very good approach for usability evaluation of different AR applications. One possibility is to develop generic usability criteria for AR applications, which allows different kinds of AR applications. Still, it is not an easy task, and if generic heuristics are developed, they will only serve as a starting point and apply to high level design issues. (Azuma 2001, 43; Dünser et al. 2007; Dünser & Billinghurst 2011, 4, 15; Träskbäck 2004, 11.) Bowman et al. (2002, 409) warn, considering the situation with VEs before, that it is tempting to over-generalize the results of VE usability evaluations in a generic context. Even though generic heuristics were used, one should describe the environment of the evaluation. Evaluations should also be carried out in a range of different environments and devices.

As Träskbäck (2004, 18) points out, not all AR applications need to fulfill all of the requirements of an ideal system. It would be wise to adjust the heuristics considering the application in focus, for example, to include a possibility to tick-mark if some of the usability requirements of an AR application are not applicable to it.

3.3.1 Usability evaluation heuristics developed for Virtual Environments

Heuristics developed for VEs share some common features with AR applications, and they have been applied partly for AR application evaluation heuristics. For example Gabbard & Hix (2001) have developed a set of guidelines for AR, but as Dünser & Billinghurst (2011, 296) point out, Gabbard and Hix's guidelines are so extensive that they are not easy to apply for practitioners and researchers. Guidelines are also taken from papers of Virtual Reality (VR) systems, and according to Dünser and Billinghurst, they are not very well applicable for AR. The heuristics developed for VEs concerning interaction (like selection and manipulation of the 3D objects), multimodal output and side-effects are applicable for AR environments (Dünser & Billinghurst 2011, 292–295). Also some of the wayfinding guidelines can be applied, as pointed out in chapter 3.2. Some of the most promising and referred VE heuristics were studied, and the following criteria based on VEs are applied in this thesis (Table 2):

Table 2. Usability evaluation criteria from VE heuristics which are applicable for AR applications.

Sutcliffe & Gault (2004, 833)	**Sutcliffe & Kaur (2000, 419)**	**Stanney et al. (2003, 449–467)**
Natural engagement	Recognizable objects	Interaction usability concerns: – Wayfinding – Object selection and manipulation Input devices should be easy to use. Object selection points should be obvious. It should be easy to select multiple objects.
Compatibility with the user's task and domain		
Natural expression of action	Approachable object	
Close coordination of action and representation		
Realistic feedback	Affordance for action	Multimodal system output usability concerns: – Visual output – Auditory output – Haptic output Visual, auditory, and / or haptic should have high frame rate and low latency and be seamlessly integrated into user activity.
Faithful viewpoints		
Clear entry and exit points	Clear object components	
Support for learning	Locatable areas for manipulation	Side effects usability concerns: – Comfort – Sickness – Aftereffects System should be comfortable for long term use
Clear turn-taking		

3.3.2 Already existing attempts to develop usability evaluation heuristics for AR applications

Some attempts have been made to develop specific heuristics for AR applications, but they are in many cases developed for application specific use (Dünser & Billinghurst 2011, 291). With these kinds of heuristics, one must be careful if adapting the guidelines to other applications, like Kaufmann & Dünser (2007, 663) emphasis. Examples of that kind of heuristics are Wang & Dunston (2009), Pribeanu et al. (2009), Martín-Gutiérrez et al. (2010) and Ko et al. (2013). The criteria adopted from them which can be seen to be common and important to all AR applications are presented in Table 3.

Table 3. AR application criteria adopted for generic heuristics from application-specific AR usability evaluation heuristics.

<table>
<tr><th>Wang & Dunston (2009, 94–99)</th><th>Pribeanu et al. (2009, 180)</th><th>Martín-Gutiérrez et al. (2010, 303–304)</th><th>Ko et al. (2013, 507)</th></tr>
<tr><td rowspan="2">Did you feel disoriented?
With the AR system, are you isolated from and not distracted by outside activities?
Were you able to actively survey the environment and easily locate objects?
Did the surrounding real background help your spatial comprehension of the model?
Is the AR display effective in convincing senses of models appearing as if in the real world?
Did you have a natural perspective [...] while manipulating the tracking marker?</td><td>Adjusting the "see-through" screen / stereo glasses / headphones is easy.</td><td rowspan="2">The Augmented Reality application has been stable (doesn't block).</td><td rowspan="2">User information:
– Defaults
– Multi-modality
– Visibility</td></tr>
<tr><td>The work place is comfortable.</td></tr>
<tr><td rowspan="2">Did the visual display create difficulties for performing?
Was the FOV (field of view) appropriate for supporting this activity?
Did the visual display maintain adequate stability (no distortion) of the image as you moved?
Does visual output / display have / exhibit an acceptable degree of response delay with no perceivable distortions in visual images / lag in image updating?</td><td>Observing through the screen is clear.</td><td rowspan="2">The familiarization with gestures and manipulating virtual objects has been easy.</td><td rowspan="2">User-interaction:
– Direct manipulation
– Low physical effort</td></tr>
<tr><td>Understanding how to operate the [AR application] is easy.</td></tr>
</table>

Can you predict responses to your actions? Did you have satisfactory control over the system?	The superposition between projection and the real object is clear.	Upon manipulating the virtual figures there is no delay in the screen, the virtual image does not produce "image leaps".	User-usage: – Context-based
	Understanding the vocal explanation is easy.		
Is tracking marker lightweight, portable, non-encumbering, and comfortable, thereby avoiding issues of limited your mobility and fatigue? Is display lightweight, portable, non-encumbering, and comfortable thereby avoiding issues of limited your mobility and fatigue? Did the real-world props (tracking devices) introduce body fatigue while you interacted with the AR system? Did the real world props (tracking devices) introduce hand / arm fatigue while you interacted with the AR system? Did you experience high levels of general discomfort during interaction with the AR system? Did you experience nausea during your interaction with AR system? Did you experience excessive eye fatigue? Is the AR system comfortable for long-term use?	Reading the information on the screen is easy.	The three-dimensional virtual figures are clear and do not present definition difficulties.	
	Selecting a menu item is easy.		
	Collaborating with colleagues is easy.	Utilizing materials (design notebook) and Augmented Reality technology has been easy and intuitive.	
	I like interacting with real objects.		

Also Vallino (1998, 19–20) has presented ideal requirements for an AR system. It combines many important issues which can be derived from the generally known AR features and problems. The requirements constitute of the following issues:

- Constrained cost to allow for broader usage
- Perfect static registration of virtual objects in the augmented view
- Perfect dynamic registration of virtual objects in the augmented view
- Perfect registration of visual and haptic scenes
- Virtual and real objects are visually indistinguishable
- Virtual objects exhibit standard dynamic behaviour
- The user has unconstrained motion within the workspace
- Minimal a priori calibration or run-time setup is required

Dubois et al. (2013, 181–199) have attempted to develop an evaluation heuristics for AR application based on AR research. The central idea is to accept the multitude of the applications developed for different usage areas and contexts, and list components already found. When a database contains enough content about different applications, usage areas and contexts, it is possible to retrieve best references for each design and evaluation situation and apply them. According to the researchers, the heuristics has been already applied successfully[4]. The aim is ambitious and one alternative in approaching the multitude, but when the component list was studied further, it seemed that unambiguous definition of the components and understanding the definitions universally is difficult. Also, for the concrete need to find quick help in evaluating AR applications under development, this tool will not be much of help. For this reason another, more generic heuristics are seen to be a more productive approach in this thesis.

Dünser et al. (2007) have made a good start in describing what kind of usability evaluation issues need to be considered in the case of AR applications, without considering the devices the application is developed for. They point out that the criteria is not complete, and there are no specific rules used in developing it. The aim has been in recognizing some important design principles for AR applications and discuss their relationship with AR system design and offer examples of how to apply usability principles for AR. The criteria have been presented in Table 4.

[4] The developed heuristics has been available for testing in the internet, but at the time of writing this work, it was not found anymore.

Table 4. Examples of design principles and usability heuristic for AR systems (Dünser et al. 2007, 38–40).

Design principle	**Description**	**What it means for AR applications?**
Affordance	Affordance communicates the connection between a user interface and its functional and physical properties to the user – by appropriate interaction metaphors it is easy to communicate what the device is used for.	An affordance of AR applications is direct object manipulation in a three-dimensional space, thus, interaction devices which are registered in 3D should be preferred.
Reducing cognitive overhead caused by interaction with the application	Cognitive overhead caused when interacting with the application may hinder focusing on the actual task and reduce learning effects. It may be caused by novelty of interaction techniques and can be high especially for novice users.	Especially registration errors in AR systems requires cognitive effort of the user when virtual objects are aligned inaccurately to the real objects.
Low physical effort as a goal while using the AR application	The user should be able to accomplish tasks without unnecessary interaction steps without fatigue.	Fatigue may be caused by the heavy or unpleasant user worn parts of the system (e.g., data helmet). Simulator sickness may occur also with AR. When user's viewpoint move from an AR presentation to a VR presentation, motion sickness and disorientation may be caused without a transitional AR interface. Usage times of AR applications should be short enough to reduce the negative physical effects.
Learnability	Learning to use the system should be easy for the user.	AR applications allow realization of novel interaction techniques which need to be learned before the user can use the system efficiently. On the other hand, natural and intuitive interaction techniques and methods are available within AR applications which reduce the need of learning to use the application. Traditional user interface elements may be combined with AR user interfaces because they are already familiar to users. The user interface should be as consistent as possible by its appearance and behaviour, and it should be designed to be as similar as the ones used in the target application domain.
User satisfaction — objective and subjective measurable experiences	User experience is important especially when the application is not used to accomplish tasks but engage user. Subjective user perceptions when interacting with the application are also important for usability, not just the objective measurements.	Physical and virtual elements should be matched in a way that the real context is integrated with the AR experience. For example in an AR game, enjoyment depends on the suspension of disbelieve and registration errors should be avoided because they may break point for natural interaction.

Flexibility in use	User interfaces of AR applications should be designed for different user preferences and abilities.	AR offers different kinds of input and output devices and allows their integration to accommodate different user preferences. On the other hand, certain input modalities are more useful for certain kinds of tasks. The balance must be found between offering different possibilities and selecting the most suitable modalities.
Responsiveness and feedback towards user actions	The lag between commands and feedback cannot be too long, or user is unable to build a persistent model of cause and effect. User should be aware of the status of the system, for example, when a control is used.	Slow tracking performance can cause lag and problems with current AR systems, which should be diminished with the evolution of the technology. Meanwhile it should be taken into account in a way that poor tracking does not interfere too much with task performance.
Error tolerance	Systems should be robust and error tolerant.	Many AR systems are still prone to instability because of the early development stages, and tracking stability is a major problem. Inaccuracies can be caused by numerical error estimations, environmental conditions or human errors, and cause virtual information jumping, jittering or disappearing. Hybrid tracking technologies may help in resolving this problem as well as identifying and resolving error scenarios.

4 METHODOLOGY

Several methods and combinations of them can be used to develop a new heuristics and validating it. Still, as Jiménez et al. (2012, 51) state, there is no evidence of a formal process or methodology which would have been used in establishing heuristics. Overall, it seems that literature study, practical experience of the domain of new heuristics or existing heuristics (such as Nielsen's or something else) have been used as a starting point, and new heuristics have been developed based on them (cf. Jiménez et al. 2012, 51; Ko et al. 2013, 504–505; Muñoz et al. 2011, 172; Stanney et al. 2003, 448–449; Sutcliffe & Gault 2004, 832; Pribeanu et al. 2009, 177–179; Martín-Gutiérrez et al. 2010, 302–303). Jaferian et al. (2014, 316–318) provide a thorough review of the most prominent literature of systematically developing usability evaluation heuristics. They distinguish between the bottom-up and top-down approaches, of which the first is based on the use of real-world data when developing the heuristics, and the latter is based on high-level expert knowledge and / or theory. Even though using both of the methods is suggested to be used, the approach in this study was mostly based on the top-down approach.

The developed heuristics need to be evaluated for their effectiveness. According to Jaferian et al. (2014, 326–327), four ways to tackle the problem are been used in heuristic creation literature: 1) no evaluation / informal evaluation, 2) long-term evaluation (using and refining the heuristics in practice), 3) controlled study of the effectiveness without a control group and 4) controlled comparative evaluation (comparison against existing heuristics). For example, Korhonen & Koivisto (2006, 14) have used the long-term evaluation approach while developing game playability heuristics, while experts evaluated several applications with the developed heuristics and modifications on it were made based on the feedback. Expert evaluations of the relevance of evaluation criteria are used (Jiménez et al. 2012, 52) which might fall into the category of informal evaluation or controlled study of the effectiveness without a control group, depending on the case.

Other methods may be used to validate the heuristics before the effectiveness evaluation is carried out. For example in the field of healthcare and education, there are examples of measurement instrument validation. According to Beck and Gable (2001, 202) also a priori validation should be carried out before testing the measurement instrument. Engels (2013,

2) points out that standardized procedures about how appropriate validation should be accomplished are not available. In some cases, heuristics are validated by using different methods of correlational analyses between the heuristics (Ko et al. 2013, 505–506).

Two examples of heuristics development and validation processes are given to illustrate some of the differences and also similarities of them, and third, the model of Rusu et al. (2011) is described. The basis in this thesis lies in the latter model, but ideas from other processes are also applied.

Jaferian et al. (2014, 318–330) used a very systematic and thorough method of developing heuristics for evaluation of an IT security management (ITSM) application. They started with the previously mentioned bottom-up approach by getting and understanding of the characteristics of ITSM tools from the publications in the field. They aimed at a saturation of the themes which came up from the papers, and ended up with 19 guidelines. After that, they moved to the top-down approach by finding an appropriate theory which they used to further study the guidelines describing the characteristics of ITSM domain. They used the theory to build 10 principles which they used to help in explaining each of the guidelines. The guidelines were concentrated to 13 main explaining principles and placed under 6 categories with supporting principles. Each category was then converted into a heuristics, which were discussed with peers iteratively and some of them were reworded if necessary to be more understandable. The heuristics were then titled, described and presented with the empirical support for them from the literature. After the set of heuristics was finalised, an effectiveness evaluation was carried out by a controlled comparative evaluation, where the heuristics were compared to existing, in this case, Nielsen's heuristics.

Ko et al. (2013, 503–507) analysed existing Augmented Reality research regarding their own study area (location based mobile AR applications). The problems observed and reported were categorized in four different categories. Based on the literature study, the usability principles were collected together and an expert meeting was arranged, where the 61 usability principles were discussed through. Part of the criteria were integrated and part of them were discarded, since they were not seen as relevant for the application which was evaluated in the research. Next, a classification system was developed with the relationships matrix used to illustrate the relations between different criteria. If there was a relationship between two criteria, it was marked with number 2. If the relationship was ambiguous, it was marked with number 1. If there was no observable relationship, it was marked with 0. Ten experts participated in the classification. Based on the classification, the principles were divided in five different categories and definitions for them were written.

Rusu et al. (2011) developed a six-step method to develop a usability heuristics (Table 5):

Table 5. Methodology for developing usability heuristics
(Jiménez et al. (2012, 52) according to Rusu et al. (2011)).

Stage	Description
Step 1: Exploratory	For collecting bibliography regarding specific topic of study, including general or related features (if there are some).
Step 2: Descriptive	For highlighting the most relevant characteristics of the previously collected information, in order to formalize the main concepts associated with the topic of study.
Step 3: Correlational	For identifying the characteristics that heuristics for specific applications should have, taking into account traditional heuristics and the analysis of cases of study.
Step 4: Explicative	For formally specifying the set of heuristics, using a standard template.
Step 5: Experimental validation	For checking new heuristics against traditional (Nielsen's) heuristics by experiments.
Step 6: Refinement	For refining the heuristics in base of the feedback obtained through the validation stage.

The steps 1–4 and 6 of this method are applied in this thesis, but the order of some of the steps is changed and some minor modifications are made (Table 6). One reason for changing the order is that the experimental validation step (5) as described above was replaced with a priori validation step.

Table 6. Modified methodology for developing usability heuristics (Jiménez et al. (2012, 52) according to Rusu et al. (2011)), modifications marked with grey background colour.

Stage	Description
Step 1: Exploratory	For collecting bibliography regarding specific topic of study, including general or related features (if there are some).
Step 2: Descriptive	For highlighting the most relevant characteristics of the previously collected information, in order to formalize the main concepts associated with the topic of study.
Step 3: Correlational	For identifying the characteristics that heuristics for specific applications should have, taking into account traditional heuristics and the analysis of cases of study.
Step 4: A priori validation	For validating the heuristics with the help of experts to test the possible overlaps and relevancy of the proposed criteria.
Step 5: Refinement	For refining the heuristics in base of the feedback obtained through the a priori validation stage.
Step 6: Explicative	For formally specifying the set of heuristics.

In this thesis, literature review was carried out in step 1, and the specific applications are explored which require new usability heuristics. In step 2, the very meaning of usability and its characteristics are re-examined in the context of AR applications. In correlational stage (step 3), a list of preliminary heuristics is suggested. In step 4, two kinds of procedures are carried out. First, a list of preliminary heuristics is presented to the evaluators and they are interviewed. Interviews are chosen as an alternative for expert meeting (cf. Ko et al. 2013), because of resource issues. After the interviews, modifications will be made for the preliminary heuristics based on the comments of the experts. Second, a further developed version of the heuristics will be given to the experts in an evaluation table, and the experts are asked to evaluate the relevance of each item and the cohesion of the item against formed heuristics categories. Step 5, the refinement of the heuristics, a set of AR heuristics is finalised to be used. Step 6 is the final stage of the method applied in this work. The heuristics are formalised, but the template in this study will not be as thorough as in the original model of Rusu et al. (Table 7). Only ID, name and definition will be used, even though it would be advisable to add also explanations, especially in the case of the technology which may not be familiar for all. Experimental validation of the AR heuristics (step 4 in the original model) will be carried out after the refinement step, but it was left out of the scope of this work.

Table 7. Standard template for formalization and specification of the set of proposed heuristics (Muñoz et al. (2011, 172).

Issue	Description
ID, name and definition	Heuristic's identifier, name and definition
Explanation	Heuristic's detailed explanation, including references to usability principles, typical usability problems, and related usability heuristics proposed by other authors.
Examples	Examples of heuristic's violation and compliance.
Benefits	Expected usability benefits, when the heuristic is accomplished.
Problems	Anticipated problems of heuristic misunderstanding, when performing heuristic evaluations.

Analysis methods used in the validation step 4 were considered. The small number of evaluators would not allow any statistical methods to be used. Still, some basic calculations were used, but mainly to help in the considerations of the researcher, who was a content area expert in AR. If more AR content area experts would have been available, much more emphasis could have been put on the evaluations.

Averages were used when evaluating the cohesion of each of the items with the categories, since statistical methods such as confirmatory factor analysis would not be suitable. When evaluating the relevance of the items for AR applications, Content Validity Index (CVI) was used as indicative. CVI has been used in validations of measurement instruments like question forms for health care. CVI is a value calculated from expert ratings of the content relevance against the items on the instrument. (Beck & Gable 2001, 209.) There are some slight variations in the use of CVI — the scale from 0 to 2 has been used where the number value of 1 means neutral. The four-point Likert scale has been used in a similar manner — the number values 1 and 2 are treated as a group and the values 3 and 4 likewise, and there is no neutral value. In a four-point Likert scale, CVI is calculated by dividing the number of values 3 or 4 with the number of evaluators. The number of evaluators has been varied from 3 to more, and the suggested amount of evaluators is six or more if a CVI could be used. If there are less evaluators, it has been suggested that all of the evaluators should agree with their evaluation (e.g. give the number value of 4 to the items) in order to conclude that the content is valid. Also some calculation checks can be used to be sure that the variations between the evaluators are not based on chance. (Beck & Gable 2001, 209; DeVon et al. 2007, 158; Wynd 2003, 509–513.) Since the backgrounds of the evaluators were different and there were only few items of which all of the evaluators agreed with the value 4, the use of CVI as validation method would not have worked.

The heuristics lists were originally written in Finnish. Also the interviews were carried out in Finnish, as well as the evaluation instructions etc. They were later translated to English.

5 DEVELOPMENT OF THE AR USABILITY HEURISTICS

The development process of usability heuristics is presented in this chapter. The process consisted of six stages: exploratory, descriptive, correlational, a priori validation, refinement and explicative stages. The preliminary version of the heuristics was developed based on literature review in the subject area and its near fields. Expert evaluators gave feedback about the version in interviews, and some modifications were made. The relevance and cohesion of the next version of the heuristics was validated by the evaluators based on two Excel sheets specifically developed for the purpose. After that, a refined version of the heuristics was created.

The final version of the heuristics created in this work consists of six criteria: 1) interaction methods and controls, 2) presentation of virtual objects, 3) relationship between virtual objects and real world, 4) information related to virtual objects, 5) suitability for the usage context and 6) physical comfort of use. Before actually using the heuristics for evaluation of AR applications, it should be further validated in practice, by testing it with different AR applications.

5.1 Exploratory stage

Literature review consisting of the typical AR features, different application types, suggested design principles and the most commonly found problems of AR applications was accomplished and reported in chapters 3.1 and 3.2.

5.2 Descriptive stage

The issues which would need to be considered in AR usability evaluation heuristics are presented in chapter 3.3. They consist of issues from VE applications which are also relevant for AR, and the attempts to develop usability evaluation heuristics for specific AR applications.

5.3 Correlational stage

The issues found out in stages 1 and 2 were studied thoroughly, and some common aspects for AR heuristics were distinguished. It was decided that Nielsen's heuristics would be used together with AR heuristics, since many of the criteria seemed to be suitable for AR applications as well (cf. Dünser et al. 2007, 37). For this reason, some issues found to be critical for the usability of an AR application as well as any other applications do not have an own criteria in AR heuristics. Nielsen's heuristics would be otherwise used as such, but two of the criteria concerning errors (*Error prevention* and *Help users recognize, diagnose, and recover from errors*) were suggested to be combined, since the total amount of the criteria should stay as small as possible. A modular structure was considered, according to the ideas of Korhonen & Koivisto (2006, 10). The amount of separate AR criteria was high at this point (total of 34), and at this step, the idea was to receive comments about them from the experts through interviews:

1. Is it easy to identify the controls which can be used to interact with the application and know how the controls can be used?
2. If different kinds of user interface features or controls (e.g. traditional computer mouse and keyboard and newer controls such as touchscreen or gesture-based controls) are used in the same application, is it confusing to use them together?
3. Is it possible for the user to replace the interaction methods used with other which are better suitable to the context?
4. Is the manipulation of augmented objects natural (e.g. using touch-based controls?
5. Are the augmentations presented aligned accurately with the physical objects connected to them?
6. If moving while using the application, do the augmentations stay still regarding to the place they should appear in relation to the user's movement?
7. Do the augmented objects in the application correspond to the user's expectations of the real world objects and their behaviour (i.e. what can be done with the object, exploring in a natural manner, feedback of the actions on virtual objects)?
8. Is it in some ways confusing that real and virtual are combined in the application (i.e. did you try to manipulate real world object when you should have manipulated the augmented object, or did you immediately understand what kind of connection there

was between the real and augmented object)?

9. Is it possible to explore the virtual objects from different viewpoints and perspectives (e.g. using predefined bird's eye / map / camera views in location-based applications or by manipulating the objects in different ways like zooming and rotating them in modelling applications)?
10. Is the user aware of her own position, objects around her, spatial relations between the objects and expectations about the future status of the environment?
11. Is changing the attention between the application and physical environment smooth and easy?
12. Are the augmented objects concealed with each other or with real objects in a way which interferes the use of the application?
13. Is the distance of the augmented objects related to physical environment and to other augmented objects (if present) convincing?
14. Are there too many objects visible?
15. Is the size of the objects appropriate?
16. Can the objects and the background be easily differentiated from each other (i.e. is the brightness and contrasts within the objects appropriate)?
17. If the augmentations contain text, is it legible (font, size, relation to its background, position etc.)?
18. Is it possible to identify the function, type and the category of different icons in relation to other icons and by itself (without reading the text label)?
19. Is the information offered by the application organised and grouped clearly?
20. Is the user able to filter the offered information based on her interests?
21. Can the important information which requires action be identified easily (is it highlighted or differentiated in any other ways)?
22. If objects are highlighted, do they still allow noticing issues concerning with other objects or the background environment?
23. If user is able to generate content to the application, is it easy?
24. Was the basis for the activity in real world — i.e. does the application naturally integrate to authentic real world environment or context and present some additional information about the real world which would be otherwise invisible? Or is the connection unnatural and artificial? Does the application need or benefit of the real

world environment?

25. When beginning to use the device and application, was it straightforward or did the device need any procedures which had to be carried out before it was ready to use (e.g. calibration, adjusting usage settings)?
26. Does the application make it faster and / or easier to get information of the physical environment?
27. Was the device and application used appropriate for the usage environment — e.g. was it easy to see what was on the display or hear if audio was used in the application?
28. Was the device used appropriate for carrying out the task it was designed for, or would some other kind of device been better?
29. If the application is used together with others, is it easy and does it give added value (e.g. make the use of application more fun and engaging or help in accomplishing different tasks)?
30. Was the device too heavy, difficult to hold or did it cause pressure on body?
31. Did the use of the application cause nausea or headache or any other physical symptoms?
32. Was the time of the usage of the application appropriate?
33. Did you have to be on any uncomfortable positions when using the application?
34. Was the application unstable or did it even crash while using it?

Also categories were considered (*Interaction*, *Device and the application*, *Ergonomics* and *Presentation*).

5.4 A priori validation stage

The heuristics was validated in two phases. First, three experts were shortly interviewed and the feedback was gathered of the preliminary version of the heuristics. After the heuristics was further modified based on the interviews, four experts evaluated the relevance and cohesion of the heuristics by using an Excel sheet designed for the purpose. Both phases are described below.

5.4.1 Interviews based on the preliminary list of heuristics

The list of preliminary heuristics was presented to the evaluators with 9 heuristics of Nielsen. Evaluators were instructed to read the heuristics through independently before an interview, which would follow approximately one day after receiving the heuristics lists. Short introduction for each heuristics list was also included in the document, which described roughly the principles of how the items were generated:

The usability heuristics for AR has been created based on the typical features of AR, already existing heuristics dealing with virtual environments (some of the features are also common with AR), existing literature of the usability of AR applications (even though any commonly shared, generic heuristics does not exist, only more specific heuristics developed to evaluate separate applications have been tried out), also some of my own experiences have probably affected the formation of them. I tried to make the heuristics suitable for evaluation of all kinds of AR applications, but achieving this goal is not guaranteed. The heuristics is meant to be a tool used in the early phase of application development or to be used to identify the most important usability issues.

The expertise area of one of the experts consisted of learning technology (especially multimodal learning applications) with a general level knowledge of AR. She had also expertise in usability evaluation. Second expert had experience in game development, especially in the area of usability, as well as generic usability expertise. Third expert had expertise in virtual environments and AR, but less expertise in usability issues. Because each of the experts had a bit different expertise area, they were guided to give feedback on areas of the heuristics they felt comfortable with, based on their expertise. Still, comments were specifically asked about the following issues already in the heuristics list:

- *Your opinion about the modularity of the heuristics (general i.e. Nielsen's heuristics, AR heuristics) instead of a single list containing all of the aspects?*
- *If you find any overlaps, please mark them and suggest how they should be combined.*
- *Used language is not finalized, better expressions may be suggested!*
- *Comments considering used terms are welcome — for example, should the term*

augmented object, virtual object or object be used?

– *Since the lists contain many separate items, they should be concentrated and the items organised according to more general categories. Please suggest categories!*

A short (30–60 minutes), informal and loosely structured interview was carried out with each of the experts. Experts' general impression of the heuristics, comments and suggestions for improving them were discussed. Some experts gave general level advice, some commented on the used language and terminology. Based on their expertise area, they emphasized different issues.

The idea of modularity was supported. Nielsen's and AR heuristics would form two modules to be used at the same time. Still, some overlaps between them should be inspected beforehand. There would also be an additional benefit because of the modularity. If the evaluated AR application would be, for example, a learning application, a separate heuristics for evaluating learning applications could be used with Nielsen's and AR heuristics. In this way, the modularity would easily allow the evaluation of different kinds of AR applications.

One of the experts had gone through the heuristics very thoroughly, and suggested categories which would form the final heuristics. The 34 items were suggested to be used as descriptions for appropriate categories. The descriptions are also important for the evaluators, especially if they are not used to evaluate for example AR applications. The items should be changed to a statement format from the question format. According to the expert, it would be important to limit the lists as short as possible so it would be easier for the evaluator to keep all the items in mind at the same time.

The heuristics were modified according to the comments, and the main categories (to be used as the criteria of the heuristics) were formed (Appendix 1). Two items were added: *The application should be tailored for different device platforms* and *If a task in real world needs to be accomplished simultaneously while using the application (e.g. going to a place or an assembly task), the device used must be appropriate for the task.* Two items in the original list were combined (items 26 and 27) and two items were divided into two different items (29 into *Using the application with other users (in physically same place or from distance) should offer added value* and *Using the application with other users (in physically same place or from distance) should be easy* and item 6 into *If the user moves while using the application, virtual objects should stay where they are supposed to be situated, not move around* and *Virtual objects should adjust to user's movements and changed viewpoints*). An Excel sheet to be used in the next phase (first page of the Excel sheet presenting the idea is included in Appendix 2). Even though the categories were formed, the descriptive items would still be treated separately in the evaluation of the cohesion between different items against the proposed categories and the relevance of the items in the heuristics.

5.4.2 Cohesion and relevance evaluations

The relevance evaluations would be accomplished to identify the items not relevant for the usability of AR applications, and the cohesion evaluation would be accomplished in order to gain insight of the possible overlaps of items and categories. The Excel sheet was e-mailed to the same evaluators as in the previous phase, but one additional evaluator was also used since it became possible. The additional evaluator had a background in VEs, information architecture and usability, and he was also familiar with AR.

For the relevance evaluation, the evaluators were asked to mark a value between 1 to 4 indicating the relevance with each of the items and categories in the first column, where *1 = not relevant at all* and *4 = very relevant*. The categories were bolded to help separating them from the items.

For the cohesion evaluation, the same matrix was used. Nine categories in the first row formed a matrix with the items and categories in the second column. The evaluators were instructed to indicate their opinion about the strength of the cohesion between the items and categories in the cross-section cell with the numeric value from *1 = not related at all* and *4 = strongly related*. The order of the items in the second column was shuffled, so that the items most probably falling into same category would not be listed close to each other, and in this way, the evaluators would be forced to think about each of the items thoroughly.

It would have been more interesting if each of the items would have been compared against each of the others items instead of only the categories, but this would have been too time consuming for the evaluators. Since the evaluation of one of the evaluators could make an enormous effect on the results, lots of emphasis was also put in the considerations of the researcher. It would have also been possible to calculate medians, but for a small amount of evaluations, it was possible to evaluate the results with visual estimate.

The categories were evaluated against each other as well, to see if there was a strong cohesion between some of the categories and as an indicator of possible overlaps of them. Again, the averages of the numerical values evaluators gave were calculated and compared with visual estimate.

5.5 Refinement stage

For the relevance evaluations, the limitations of Content Validity Index (CVI) were obvious in this case because of the limited amount of the evaluators (see chapter 4), and the method was used only as indicative. The calculation of CVI would not be the most important method, instead, it was possible to check if there were some items which had gained very low values from each of the evaluators. No items got alarmingly low evaluations compared to the others, but one item concerning the device platform and possibilities for content production (*The application should be tailored for different device platforms*) got a bit lower CVI than the others (0,5 while others got at least 0,75) (Table 8).

Table 8. Deviation of CVI for the items.

CVI	Amount
1	20
0,75	16
0.5	1

Also one of the categories (*Possibility for content production*) got lower CVI than the others (0,5 while others got at least 0,75) (Table 9):

Table 9. Deviation of CVI for the categories.

CVI	Amount
1	5
0,75	3
0.5	1

These items would be important to consider when developing AR applications, but because they would be more utility than usability issues, it was decided that they would be left out from this heuristics. Also categories *Social usage of the application* and *Usage of the application* were left out, based on researcher's own decision, since they were more connected with utility than usability. It would probably be a good decision to evaluate an AR application in respect of these items, since in the literature they are mentioned as important (Azuma et al. 2001, 42; Li & Duh 2013, 123-125). A category and items of which the inclusion was considered for the same reason was *Relationship between virtual objects and the real world*, especially the item *The basis for using the application should be physical real world*. It was also discussed with one of the evaluators and he also agreed that the item would be more related to utility. Because it is so essential aspect of AR and might also affect the usability of the application, it was still left intact. The complete evaluation results are presented in Appendix 3.

No big surprises appeared when evaluating the cohesion between the items and categories, but some changes and modifications were made. The averages of the values each evaluator gave were calculated, and by ordering the items in regard to the categories they got the highest values, it was easy to see which categories were the strongest candidates for the items (Appendix 4). Most of the items were connected to the same categories as in the previous phase after the interviews, but some of the items in the three categories seemed to be connected to different categories as after the interviews:

- The item *Virtual objects should be accurately aligned with the real world objects linked with them* which was associated more strongly to the category *Relationship between virtual objects and real world* (average of 3,75) than to the category *Virtual* *objects* (average of 3,25) in which it was originally matched.

- The same concerned the item *Virtual objects should adjust to the physical environment and other visible virtual objects in a way that they seem natural and believable in respect to the distance and location* which also got same averages related to the same categories.
- The item *It should be possible to identify the purpose of virtual action or symbol icons based on their appearance* in the category *Virtual objects* was more associated to the category *Information related to the virtual objects*, as it got the average of 3,5 connected to the latter of the categories and only an average of 2,75 connected to the category *Virtual objects*.
- The item *The used device should not be too heavy, difficult to handle or cause depressions on the body disturbing the user* was originally associated with the category *Usability of the device*, but it was moved to the category *Physical comfort of the use*.

Two items were left out from the categories:

- The item *The interaction methods and controls and their functionalities should be easily recognisable by the user* in category *Interaction methods and controls* seemed to be connected in usability in general and it should probably be added to the generic usability heuristics as an additional item.
- The item *The information offered by the application should be organised and grouped clearly* in category *Information related to the virtual objects* was somehow connected to another item (*It should be possible to identify the purpose of virtual action or symbol icons based on their appearance*).

The evaluators related the item *If virtual objects contain text, it should be legible in respect of its size, font, location, colour and how well it can be separated from its background* into the category *Information related to the virtual objects* (average of 3,75) but only with an average of 3 into the category *Virtual objects*. A decision was made to put the item still on the latter category, since text itself does not relate to information, and the item is concerned with the presentational issues making it more related to the category of *Virtual objects*.

What comes to the categories, there was not a very strong cohesion between most of them. It was expected that three of the categories (*Virtual objects*, *Information related to the virtual objects* and *Relationship between virtual objects and real world*) would be strongly associated to each other, as they did — each of them got an average of 3,5 in relation to others. It would not still be advisable to combine those categories, since one category would have been too large and general for all of the issues regarding the virtual objects. Cohesion evaluations between categories are presented in Appendix 5.

Also the category names and category descriptions were modified after the evaluations in order to make them more compact and easier to understand in a way they were meant to. Even though the original texts were in Finnish, the texts were also translated into English at the same time as the heuristics were developed, and also the English texts were modified. The category *Usability of the device* was renamed to *Suitability for the usage context*, since it was more descriptive and highlighted the viewpoint important in selecting the device for AR applications. Some examples were also given and added to the description texts, for example, in the category *Suitability for the usage context* examples were given of bright sunlight and the usage of the application while both hands are occupied.

5.6 Explicative stage

The heuristics were formalised based on the comments of the evaluators. The references to literature, (research and theoretical considerations) which have supported including each of the criteria in the heuristics from the very beginning, are attached to each criteria.

The following list of heuristics and descriptions were formed (see Appendix 6 for more comprehensive summary of references mentioned below the descriptions):

AR1 Interaction methods and controls

It should be possible for the user to choose between different interaction methods. If different interaction methods and controls are used in the same application, (for example, mouse and keyboard controls and touchscreens or gesture controls), the co-usage of them should be clear. Manipulation of the virtual objects should be carried out in a way that supports the natural interaction methods and characteristics of the object — for example, the user should be able to manipulate the three-dimensional object by touching it somehow instead of using indirect menu commands.

References:

- Dünser et al. 2007, 38–40
- Billinghurst et al. 2005, 17–18
- Bowman et al. 2001, 97–98, 102–103
- Sutcliffe & Gault 2004, 833
- Sutcliffe & Kaur 2000, 419–420
- Stanney et al. 2003, 463–466
- Pribeanu et al. 2009, 180
- Martín-Gutiérrez et al. 2010, 303
- Ko et al. 2013, 507
- Wang & Dunston 2009, 15
- Kaufmann & Dünser 2007, 663
- Ganapathy 2013, 179
- Li & Duh 2013, 118–121
- Azuma et al. 2001, 38–39

AR2 Presentation of virtual objects

Virtual objects should not occlude each other or real world objects in a way which disturbs the usage of the application (for example, by hiding important information). Too many virtual objects should not be visible at the same time, so that the view will not be crowded and confusing. The size of the virtual objects should be appropriate, since it is difficult to observe too small or large object, and the latter may also conceal the view behind it. Virtual objects should be clearly separated from the background in different usage situations, e.g., be adequate in terms of brightness and contrast. If virtual objects contain text, it should be legible by font, location, colour and separation from the background.

The user should be able to view the virtual objects from different viewpoints: when three-dimensional objects are presented to the user, (s)he should be able to view them from different distances, sides and angles. In the case of location-based AR application, the user should be able to switch between the map view, camera view and list view. Virtual objects should adjust to the movement and viewpoint of the user as well as in the case of an equivalent physical object. Despite of the movement, the object should remain related to the physical object or point in which it is registered.

References:
- Livingston 2005, 8
- Dünser et al. 2007, 40
- Stanney et al. 2003, 464–466
- Sutcliffe & Gault 2004, 833
- Sutcliffe & Kaur 2000, 419
- Martín-Gutiérrez et al. 2010, 303
- Azuma 1997, 372–373
- Gabbard & Swan 2008, 523
- Ko et al. 2013, 507
- Vallino 2008, 19
- Wang & Dunston 2009, 97
- Ganapathy 2013, 177–178
- Li & Duh 2013, 112–117
- Azuma et al. 2001, 36, 39–40

AR3 Relationship between virtual objects and real world

The basis for the application should be in the real world, in which the application should naturally integrate. It should, for example, present virtual information which is not visible in the real world as such.

Virtual objects should respond to the expectations of the user about how corresponding real world objects would behave (for example, what kinds of activities are possible with the object, is it possible to study the object in a way that is possible with real world objects, do the virtual objects react to the actions of the user in a similar manner as the real world object would react). Virtual objects should relate to the real world and other virtual objects possibly visible in a way that they look natural and convincing what comes to the distances and locations of them.

Virtual objects should be accurately aligned with the real world objects linked with them. For example, if the application offers virtual information about the nearby building, it should be clear, for which building the information is related. If the application illustrates the invisible parts inside of a device augmenting them on top of the device, the augmentation should be accurately aligned with the corresponding physical location of the invisible parts.

The user should be aware of her/his own location, the virtual objects around her/him and the spatial relations between self and the virtual objects. Switching the attention between the application and real world should be smooth and easy. The relationship between the physical and virtual objects should be understandable in a way that the user will not get confused and try to, for example, manipulate the physical object instead of the virtual one.

References:

- Carmichael et al. 2012, 1768
- Livingston 2005, 8
- Billinghurst et al. 2005, 17–18
- Dünser et al. 2007, 38–39
- Sutcliffe & Gault 2004, 833
- Stanney et al. 2003, 463
- Pribeanu et al. 2009, 180
- Wang and Dunston 2009, 26–28, 97
- Neumann & Majoros 2008, 4–6
- Vallino 2008, 19–20
- Azuma 1997, 372–373
- Li & Duh 2013, 113–117

– Azuma et al. 2001, 40

AR4 Information related to virtual objects

Important information and information requiring actions from the user should be identifiable easily (for example, by highlighting it or separating it from the other ways from other information). If part of the information offered by the application is highlighted for quick identification, it should be possible to identify issues from the background information which might affect the actions and decisions. The user should be able to filter the information offered by the application based on her/his own interests and needs.

Virtual icons should allow identification of their affordances and categories without the need to read separate textual explanations. Different icons should also be separable from each other.

References:

– Ganapathy 2013, 178–179
– Li & Duh 2013, 114–117

AR5 Suitability for the usage context

The device and application should adjust to the usage context — the user should be able, for example, to see clearly what is presented in the display or hear the audio related to the application clearly. For example, the usage environments are different in the clear sunlight outside or when there is lots of noise in the background compared to quiet indoors conditions.

If an activity in real world context has to be accomplished at the same time when the application is used (e.g. navigating to a place or some kind of assembly task), the used device should adjust to the task accomplishment. If the application is used in a handheld tablet computer when both hands should be free for accomplishing the task, data glasses would be a better choice for device platform.

References:

- Billinghurst et al. 2005, 18
- Bowman et al. 2001, 98, 102
- Stanney et al. 2003, 465–466
- Ko et al. 2013, 507
- Wang & Dunston 2009, 18–19, 21–23, 33–37, 97
- Pribeanu et al. 2009, 180

AR6 Physical comfort of the use

The usage device should not be too heavy, difficult to handle or cause any kind of physical load or discomfort. The user should not be in difficult or uncomfortable positions while using the application. The usage of the application should not cause nausea, headache, eye pain or other physical symptoms, which might appear in the usage situation or afterwards. The usage time must be appropriate — too long usage time may cause tiredness and other harmful side effects.

References:

- Dünser et al. 2007, 39
- Sutcliffe & Gault 2004, 833
- Stanney et al. 2003, 467
- Pribeanu et al. 2009, 180
- Kaufmann & Dünser 2007, 667–668
- Ko et al. 2013, 507
- Wang & Dunston 2009, 97–98
- Vallino 2008, 20
- Azuma et al. 2001, 40

The developed AR heuristics are still not ready to be used, and they would require experimental validation and after that most probably, further refinement. The heuristics need to be tested with different kinds of AR applications and enough feedback must be gained to be able to conclude anything about their applicability in practice.

6 DISCUSSION

The literature review is based on writer's experience in the domain of AR as a researcher, and no systematically reported literature search was carried out in this work. Still, because literature searches were accomplished in the reference databases such as Web of Science and Scopus using the keywords related to the research topic, it can be assumed that the most important literature has been reviewed. Literature studied consists mainly on theoretical references (even though many actual research reports are referred), and the main problems concerning usability issues of AR are likely to be covered. Another approach might have been to review research reports about the use of AR applications.

The most obvious limitation of this work has been the small amount and the heterogeneity of the evaluators. It was difficult to find evaluators with the background in AR and usability, especially at the limited amount of time which was to accomplish the evaluations, since this thesis was somehow connected to a research project, and the heuristics needed to be completed in a very fast time schedule. The experience of the evaluators in usability evaluation issues also varied, which may have been slightly visible in the evaluation results concerning some fine-grained AR issues. Also the limited amount of time the evaluators themselves had in evaluating the AR heuristics may have caused some hastiness in the evaluation results.

Stemming from the issue of small amount of evaluators, it was not possible to use any statistical methods as a help while analysing the evaluation results. The systematicality suffered and the analysis was partly based on the experience, impressions and decisions of the researcher. Still, this issue was probably not crucial, since also a posteriori validation would be used to validate the heuristics. The experience from real evaluations will be valuable in further developing the heuristics. Also, the original goal of this work was to be a first step for developing a generic evaluation heuristics for AR applications.

The structure of the developed heuristics was modular to allow generic evaluation of any AR applications. Nielsen's heuristics are used as another module with AR heuristics. As already mentioned, it has been speculated that some of the criteria in Nielsen's heuristics (like the criteria concerning the efficiency of use) would not be appropriate for each AR application (Dünser & Billinghurst 2011, 297), since some of them are developed purely for providing entertaining experiences for the users. Thus, Nielsen's heuristics should be carefully inspected before its use with some AR applications. After all, it would be a good idea to check all of the heuristics carefully before evaluating an application with it, because it is possible that some of the criteria are not relevant, even though the heuristics are meant to be generic and applicable to all kinds of AR applications.

An issue that is also connected to the modularity issue is the amount of separate heuristics. The amount of them is 15 in this case, when Nielsen's heuristics and AR heuristics are combined. If the evaluators find it difficult to keep that many heuristics in mind at the same time while evaluating an application (especially from the new domain area, if they are not familiar with AR), one solution would be to run two evaluations separately, first with Nielsen's and then with AR heuristics. This kind of approach was suggested by Jaferian et al. (2014, 344). Also the descriptions of AR heuristics were quite long, since they were actually formed of many separate items and collected under a common category. It might be advisable to go through the descriptions and raise the abstraction level, which would help to shorten them.

It might also be a good idea to commit a quick analysis of the usage environment of the application before starting to evaluate its usability. In this way, these issues would be fresher in the minds of the evaluators. The known limitations and requirements which are related to the user, usage environment and task which is accomplished affect very much to the selection of the devices and technologies and evaluation of them. Of course this kind of analysis should be carried out already in the application development phase, but there may be different usability requirements for different user groups, and an application may be usable for one group and unusable for another.

Categorising items considering static and dynamic registration errors to different criteria (AR2 and AR3) according to the feedback from the evaluators required considerations. Still, it may prove to be a right solution, since the usability is evaluated from the different viewpoint than technological considerations. Even though static and dynamic registration errors are closely connected with tracking, they may be connected with different issues from the viewpoint of usability. This was probably a proof of how much the AR usability issues are connected and dependable on each other. As Wang & Dunston (2009, 16) point out considering the selection of a input mechanism for MR system, the more intuitive input mechanism needs more computational power which can cause system lag for example with registration.

Rapid application development (RAD) has gained a lot of attention and is worth a few thoughts. Rapid development cycles, prototypes and iterative development processes may have in many respects replaced traditional design methods. Mackay et al. (2000, 740) point out that RAD will be *"most suitable for applications with a strong element of user interface and with a lack of complexity regarding both requirements and computation"*. In the respect of complexity, AR may not be the best possible candidate. On the other hand, de Sá & Churchill (2013, 139–164) have tested the use of different kind of prototypes for mobile AR. To create a realistic experiences for evaluators, the prototype does not need to be high-fidelity AR prototype, for example video prototypes proved to be the best option for rapid prototyping and thus, saving the costs and also involving the users in early phases of the development process. Since AR applications consist of many different technologies (device, display, input, output) it may be challenging to take all of them into account when designing and evaluating applications, since one should also be aware of the effects caused by their different combinations, not to forget the many different usage contexts. (Bowman et al. 2005, 367.) But are heuristics still needed? It might be assumed that development of generic heuristics will not be made obsolete, since they will probably be used as general level design guidelines. On the other hand, specific heuristics may not be viable when technologies are getting more diverse and complex.

Finally, an interesting thought has somewhat bothered the writer while accomplishing this thesis, because it could shake the whole grounds the work is based on. An issue which provides a completely new way to look at the usability of AR is to think about AR as a user interface connecting real and digital environments (Olsson 2012, 32–33, Bowman et al. 2005, 388). Should, then, the usability of AR applications be evaluated by seeing the applications as interfaces to the physical environment, which can reveal issues which are invisible for our senses as such? This idea would raise the abstraction one level higher. The applications should be evaluated on the grounds of how well they succeed in giving user information about the physical environment. This would bring AR close to its very essence and most potential usage area. This also comes near to Rekimoto & Nagao's idea about augmented interaction with real world environments, where the goal is to reduce the amount of computer manipulations by using environmental information as implicit input. At the same time, people are not bothered by computer operations while accomplishing a task related to real world. (Rekimoto & Nagao 1995, 29–30.) The criteria in AR heuristics concerning the connectedness to the physical environment is also quite near to this idea.

7 SUMMARY

Usability is an important part of application development. Some attempts to develop usability heuristics for AR applications have been made, but they are mostly application-specific, and no generic AR heuristics exists. Because of the diversity of AR applications, platforms and devices, generic heuristics would prove to be a better solution than heuristics attempting to cover all of the application and low-level usability specific concerns. An attempt to develop such a heuristics was made in this thesis.

Different kinds of methods have been used to develop usability evaluation heuristics. In this work, a framework developed by Rusu et al. (2011) was used as a basis, but insights of other heuristics development cases were combined to it and it was slightly modified A priori validation phase was added to the original framework, and a posteriori validation phase was left out of the scope of this work.

A literature review of AR as a technology, its typical features and known problems was accomplished. Also application specific heuristics for AR and heuristics developed in its near fields such as VEs and 3D user interfaces were studied. Based on them, a preliminary set of heuristics was developed, and it was further modified based the feedback of three experts. Four experts evaluated the relevance and cohesion of the next version of the heuristics. Some modifications were made and the final version of the heuristics was formed with references to the literature they were based on (Table 10).

Table 10. AR heuristics.

Code	Heuristic
AR1	Interaction methods and controls
AR2	Presentation of virtual objects
AR3	Relationship between virtual objects and real world
AR4	Information related to virtual objects
AR5	Suitability for the usage context
AR6	Physical comfort of use

AR heuristics should be used in combination with generic usability heuristics such as Nielsen's (1995). The heuristics are not ready to be used yet, since they will need a validation. Validation can be carried out by evaluating different kinds of AR applications.

REFERENCES

1. Azuma, R.T. 1997. A Survey of Augmented Reality. *In Presence: Teleoperators and Virtual Environments*, vol. 6, no. 4, pp. 355–385.
2. Azuma, R.T., Baillot, Y., Behringer, R., Feiner, S., Julier, S. & MacIntyre, B. 2001. Recent Advances in Augmented Reality. *Computers & Graphics*, vol. 21, no. 6, pp. 34–47.
3. Beck, C.T. & Gable, R.K. 2001. Ensuring Content Validity: An Illustration of the Process. *Journal of Nursing Measurement*, vol. 9, no. 2, pp. 201–215.
4. Billinghurst, M., Grasset, R. & Looser, J. 2005. Designing Augmented Reality Interfaces. *ACM SIGGRAPH Computer Graphics*, vol. 39, no. 1, pp. 17–22.
5. BlindSquare. [WWW pages]. [Retrieved 20.11.2014]. From: http://blindsquare.com/
6. Bowman, D.A., Gabbard, J.L. & Hix, D. 2002. A Survey of Usability Evaluation in Virtual Environments: Classification and Comparison of Methods. *Presence*, vol. 11, no. 4, pp. 404–424.
7. Bowman, D.A., Krujiff, E., LaViola Jr., J.J. & Poupyrev, I. 2001. An Introduction to 3-D User Interface Design. *Presence*, vol. 10, no. 1, pp. 96–108.
8. Bowman, D.A., Kruijff, E., LaViola jr, J.J. & Poupyrev, I. 2005. 3D User Interfaces. Theory and Practice. Boston: Addison-Wesley.
9. Carmichael, G., Biddle, R. & Mould, D. 2012. Understanding the Power of Augmented Reality for Learning. Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education, Oct 09, 2012 in Montréal, Quebec, Canada, vol. 2012, no. 1, pp. 1761–1771.
10. Carmigniani J. & Furht, B. 2011. Augmented Reality: An Overview. In Furht, B. (ed.) Handbook of Augmented Reality. New York: Springer.
11. DeVon, H.A., Block, M.E., Moyle-Wright, P., Ernst, D.M., Hayden, S.J., Lazzara, D.J., Savoy, S.M., Kostas-Polston, E. 2007. A Psychometric Toolbox for Testing Validity and Reliability. *Journal of Nursing Scholarship*, vol. 39, no. 2, pp. 155–164.
12. Dinosaurs — Live! [In Apple iTunes WWW pages]. [Retrieved 28.12.2014]. From: https://itunes.apple.com/us/app/dinosaurs-live!/id463198136?mt=8
13. Dubois, E., Scapin, D.L., Charfi, S. & Bortolaso, C. 2013. Usability

Recommendations for Mixed Interactive Systems: Extraction and Integration in a Design Process. In Huang, W., Alem, L. & Livingston, M.A. (eds.) Human Factors in Augmented Reality Environments. New York: Springer.

14. Dünser, A. & Billinghurst, M. 2011. Evaluating Augmented Reality Systems. In Furht, B. (ed.) Handbook of Augmented Reality. New York: Springer.

15. Dünser, A., Grasset, R. & Billinghurst, M. 2008. A Survey of Evaluation Techniques Used in Augmented Reality Studies. Technical Report TR-2008-02. University of Canterbury, Human Interface Technology Laboratory New Zealand.

16. Dünser, A., Grasset, R., Seichter, H. & Billinghurst, M. 2007. Applying HCI principles to AR systems design. Mixed Reality User Interfaces: Specification, Authoring, Adaptation. MRUI'07: 2nd International Workshop at the IEEE Virtual Reality 2007 Conference, 11 March 2007, Charlotte, NC, USA, pp. 37–42.

17. Engels, M.C. 2013. Validation procedure MILQ. Research report [WWW document]. Utrecht University, Center for Textbook Studies (CLU). [Retrieved 20.11.2014]. From: http://clu.nl/wp-content/uploads/2014/05/Validation-procedure-MILQ-Maaike-Engels.pdf

18. Gabbard, J.L. & Hix, D. 2001. Researching Usability Design and Evaluation Guidelines for Augmented Reality (AR) Systems [WWW document]. Systems Research Center, Virginia Tech, USA. [Retrieved 22.11.2014]. From: http://www.sv.vt.edu/classes/ESM4714/Student_Proj/class00/gabbard/index.html

19. Ganapathy, S. 2013. Design Guidelines for Mobile Augmented Reality: User Experience. In Huang, W., Alem, L. & Livingston, M.A. (eds.) Human Factors in Augmented Reality Environments. New York: Springer.

20. Gartner. 2014. Gartner's 2014 Hype Cycle for Emerging Technologies Maps the Journey to Digital Business [WWW document]. [Retrieved 19.11.2014]. From:: http://www.gartner.com/newsroom/id/2819918

21. Google Glass. [WWW pages]. [Retrieved 14.12.2014]. From: https://www.google.com/glass/start/

22. Gutiérrez, J.M., Saorín, J.L., Contero, M., Alcañiz, M., Pérez-López, D.C. & Ortega, M. 2010. Design and validation of an augmented book for spatial abilities development in engineering students. *Computers & Graphics*, vol. 34, no. 1, pp. 77–91.

23. ISO 9241–11. 1998. Ergonomic requirements for office work with visual display terminals (VDTs) — Part 11: Guidance on usability. Geneva: International Organization for Standardization (ISO).

24. Jaferian, P., Hawkey, K., Sotirakopoulos, A., Veles-Rojas, M. & Beznosov, K. 2014. Heuristics for Evaluating IT Security Management Tools. *Human-Computer Interaction*, vol. 29, no. 4, pp. 311–350.

25. Jiménez, C., Rusu, C., Roncagliolo, S., Inostroza, R. & Rusu, V. 2012. Evaluating a Methodology To Establish Usability Heuristics. 31st International Conference of the Chilean Computer Science Society, November 12–16, Valparaiso, Chile, pp. 51–59.

26. Kaufmann, H. & Dünser, A. 2007. Summary of Usability Evaluations of an Educational Augmented Reality Application. In Shumaker, R. (ed.) Virtual Reality, HCII 2007, LNCS 4563, Berlin: Springer-Verlag.

27. Ko, S.M., Chang, W.S. & Ji, Y.G. 2013. Usability Principles for Augmented Reality Applications in a Smartphone Environment. *International Journal of Human-Computer Interaction*, vol. 29, no. 8, pp. 505–515.

28. Koivisto, E.M.I. & Korhonen, H. 2006. Mobile Game Playability Heuristics. Version 1.0. Nokia.

29. Korhonen H. & Koivisto E.M.I. 2006. Playability Heuristics for Mobile Games. 8th international conference on Human-Computer Interaction with mobile devices and services, MobileHCI'06, September 12–15, Espoo, Finland, pp. 9–16.

30. Kruijff, E., Swan, J.E., Feiner, S. 2010. Perceptual Issues in Augmented Reality Revisited. Science and Technolgy Proceedings, IEEE International Symposium on Mixed and Augmented Reality (ISMAR) 2010, 13–16 October, Seoul, Korea, pp. 3–12.

31. Li, N. & Duh, H.B-L. 2013. Cognitive Issues in Mobile Augmented Reality: An Embodied Perspective. In Huang, W., Alem, L. & Livingston, M.A. (eds.) Human Factors in Augmented Reality Environments. New York: Springer.

32. Livingston, M.A. 2005. Evaluating Human Factors in Augmented Reality Systems. Projects in VR. IEEE Computer Graphics and Applications, vol. 25, no. 6, pp. 6–9.

33. Mackay, H., Carne, C., Beynon-Davies, P. & Tudhope, D. 2000. Reconfiguring the User: Using Rapid Application Development. *Social Studies of Science*, vol. 30, no.

5, pp. 737–757.

34. Mariette, N. 2013. Human Factors Research in Audio Augmented Reality. In Huang, W., Alem, L. & Livingston, M.A. (eds.) Human Factors in Augmented Reality Environments. New York: Springer.

35. Martín-Gutiérrez, J., Contero, M. & Alcañiz, M. 2010. Evaluating the Usability of an Augmented Reality Based Educational Application. In Aleven, V., Kay, J. & Mostow, J. (eds.) ITS 2010, Part I, LNCS 6094. Berlin: Springer.

36. McCall, R., Wetzel, R., Löschner, J. & Braun, A-K. 2011. Using presence to evaluate an augmented reality location aware game. *Personal and Ubiquitous Computing*, vol. 15, no. 1, pp. 25–35.

37. Milgram, P. & Kishino, F. 1994. A Taxonomy of Mixed Reality Visual Displays. *IE-ICE Transactions on Information Systems*, vol. E77-D, no. 12, pp. 1321–1329.

38. Muñoz, R., Barcelos, T. & Chalegre, V. 2011. Defining and Validating Virtual Worlds Usability Heuristics. 30th International Conference of the Chilean Computer Science Society, pp. 171–178.

39. Neumann, U. & Majoros, A. 1998. Cognitive, Performance, and Systems Issues for Augmented Reality Applications in Manufacturing and Maintenance. In Proceedings of IEEE Virtual Reality Annual International Symposium VRAIS '98, 14–18 March, 1998, Atlanta, GA, USA, pp. 4–11.

40. Nielsen, J. 1993. Usability Engineering. Amsterdam: Morgan Kaufmann.

41. Nielsen, J. 1995. 10 Heuristics for User Interface Design [WWW document].1.1.1995, NN/g Nielsen Norman Group, [Retrieved 7.9.2014], From: http://www.nngroup.com/articles/ten-usability-heuristics/

42. Nielsen, J. & Molich, R. 1990. Heuristic Evaluation of User Interfaces. Proceedings of ACM CHI'90 Conference, 1–5 April 1990, Seattle, WA, USA, pp. 249–256.

43. Olsson, T. 2012. User Expectations and Experiences of Mobile Augmented Reality Services. Doctoral dissertation. Tampere: Tampere University of Technology.

44. Olsson, T. 2013. Concepts and Subjective Measures for Evaluating User Experience of Mobile Augmented Reality Services. In Huang, W., Alem, L. & Livingston, M.A. (eds.) Human Factors in Augmented Reality Environments. New York: Springer.

45. Polit, D.F., Beck, C.T. & Owen, S.V. 2007. Is the CVI an Acceptable Indicator of

Content Validity? Appraisal and Recommendations. *Research in Nursing & Health*, vol. 30, no. 4, pp. 459–467.

46. Pribeanu, C., Balog, A. & Iordache, D.D. 2009. Measuring the usability of augmented reality e-learning systems — a user-centered evaluation approach. In Cordeiro, J., Shishkov, B., Ranchordas, A. K. Helfert, M. (eds.) Software and Data Technologies. Third International Conference, ICSOFT 2008, Porto, Portugal, July 22–24, 2008, Revised Selected Papers. Berlin: Springer.

47. Rekimoto, J. & Nagao, K. 1995. The World through the Computer: Computer Augmented Interaction with Real World Environments. In UIST '95 Proceedings of the 8th annual ACM symposium on User interface and software technology, Pittsburgh, PA, USA. New York: ACM.

48. Rusu, C., Roncagliolo, S., Rusu, V. & Collazos, C. 2011. A Methodology to Establish Usability Heuristics. ACHI 2011: The Fourth International Conference on Advances in Computer-Human Interactions, February 23–28, Gosier, Guadeloupe, France.

49. Sá, M. (de) & Churchill, E.F. 2013. Mobile Augmented Reality: A Design Perspective. In Huang, W., Alem, L. & Livingston, M.A. (eds.) Human Factors in Augmented Reality Environments. New York: Springer.

50. Specht, M., Ternier, S. & Greller, W. 2011. Dimensions of Mobile Augmented Reality for Learning: A First Inventory. *Journal of the Research Center for Educational Technology (RCET)*, vol. 7, no. 1, pp. 117–127.

51. Stanney, K.M., Mollaghasemi, M., Reeves, L., Breaux, R. & Graeber, D.A. 2003. Usability engineering of virtual environments (VEs): identifying multiple criteria that drive effective VE system design. *International Journal of Human-Computer Studies*, vol. 58, no. 4, pp. 447–481.

52. Sutcliffe, A. & Gault, B. 2004. Heuristic evaluation of virtual reality applications. *Interacting with Computers*, vol. 16, no. 5, pp. 831–849.

53. Sutcliffe, A.G & Kaur, K.D. 2000. Evaluating the usability of virtual reality user interfaces. *Behaviour & Information Technology*, vol. 19, no. 6, pp. 415–426.

54. Tian, Y., Guan, T. & Wang, C. 2010. Real-Time Occlusion Handling in Augmented Reality Based on an Object Tracking Approach. *Sensors*, vol. 10, no. 4, pp. 2885–2900.

55. TryLive. TryLive for Eyewear Online Demo. [In TryLive WWW pages]. [Retrieved 28.12.2014]. From: http://www.trylive.com/demos/trylive-eyewear/player-fitting-room
56. Träskbäck, M. 2004. User requirements and usability of mixed reality applications. Licentiate thesis. Helsinki: Helsinki University of Technology.
57. Vallino, J.R. 1998. Interactive Augmented Reality. Doctoral dissertation. Rochester, New York: University of Rochester.
58. Viking Shoe. [Viking Shoe in Labratian WWW pages]. Updated October 7, 2011. [Retrieved 20.11.2014]. From: https://labratian.wordpress.com/2011/10/07/viking-shoes-in-ar/
59. Wang, X. & Dunston, P.S. 2009. Mixed Reality-based Collaborative Virtual Environments. Specification, Prototype, and Evaluation in Design and Construction. Saarbrücken: VDM Verlag.
60. Wagner, D. 2013. History of Mobile Augmented Reality. [WWW document]. [Retrieved 29.9.2013]. From: https://www.icg.tugraz.at/~daniel/HistoryOfMobileAR/Wikitude 2014
61. Wikitude. [WWW pages]. [Retrieved 20.11.2014]. From: http://www.wikitude.com
62. Wynd, C.A., Schmidt, B. & Schaefer, M.A. 2003. Two Quantitative Approaches for Estimating Content Validity. *Western Journal of Nursing Research*, vol. 25, no. 5, pp. 508–518.

APPENDIX 1. The second version of the heuristics

Heuristics	Items / description
1. Interaction methods and controls	- The interaction methods and controls and their functionalities should be easily recognisable by the user. - If different interaction methods and controls are used in the same application (for example traditional mouse and keypad controls) and newer touch displays or gesture control) the co-usage of them should be clear. - The user should have a possibility to choose between different interaction methods, if some of them are better applicable to the usage context. - The manipulation of the virtual objects should be carried out in ways which support natural, activity supporting ways. For example, it is possible to "touch" a three-dimensional virtual object through a touch-screen and for example to rotate it using finger instead of manipulating it with traditional menu commands.
2. Virtual objects	- Virtual objects should be accurately aligned with the real world objects linked with them. If, for example, the application offers information about a nearby building, it should be obvious to which building the information is connected. If the application visualises a part which is located inside a device by presenting a virtual image of it augmented on top of the physical device, the virtual image should be aligned accurately to the part of the physical device in which it is actually located inside the device. - If the user moves while using the application, virtual objects should stay where they are supposed to be situated, not move around. - Virtual objects should adjust user's movements and changed viewpoints (for example, the user should be able to walk around a three-dimensional object in a similar way if the object was real). - It should be possible to study the virtual objects from different viewpoints. In a location-based application, the user should typically be able to change between a camera view, map view and list view. When the user is presented with 3D virtual geometrical objects, the user should be able to zoom the object, watch it from different viewpoints and rotate it in order to study it from different angles. - Virtual objects should not occlude each other or physical real world objects in a way that disturbs the usage of the application (for example concealing other important information which should be visible). - Virtual objects should adjust to the physical environment and other visible virtual objects in a way that they seem natural and believable in respect to the distance and location. - There should not be too many virtual objects visible at the same time, so that the view would not be too crowded and messy. - The size of a virtual object should be appropriate — not too small, when studying the object is difficult, and not too large, when the object may conceal a too big part of the view and it may be difficult to get an idea of the object. - Virtual objects should be easily separated from their background in different usage sitations, e.g. be appropriate by their brightness and contrast in order to be separable. - If virtual objects contain text, it should be legible in respect of its size, font, location, colour and how well it can be separated from its background. - It should be possible to identify the purpose of virtual action or symbol icons based on their appearance, without a need to read textual explanations. All of them should also be separable from other icons.

(continues)

APPENDIX 1. (continues)

3. Relationship between virtual objects and real world	- The virtual objects of the application should correspond the expectations of the user of the real world objects and their behaviour (for example what can be done with an object, is it possible to study an object like a corresponding real world object, do the objects react to the procedures user has accomplished like corresponding objects in physical real world environment would react). - The relationship between real and virtual objects should be understandable in a way that the user is not confused, for example, trying to manipulate a physical real world object instead of a virtual object. - The user should be aware on his / her location, the virtual objects around him / her, the spatial relations between him- / herself and the expectations of the future state of the environment. - Switching the attention between the application and the physical real environment should be smooth and easy. - The basis for using the application should be physical real world, e.g. the application should naturally integrate on it — it should not be unnatural or artificial. - The application must benefit from the physical real environment — an appropriate basis is for example presenting some additional information about the real world which would be otherwise invisible.
4. Information related to the virtual objects	- The information offered by the application should be organised and grouped clearly (for example different categories and taxonomies visualising different kinds of information). - The user should be able to filter the information offered by the application based on his / her own interests and needs — for example, if the application offers lots of information, the user should be able to filter the information that is unnecessary for him- / herself. - Important information requiring actions from the user should be easily recognisable (for example by highlighting or using other ways to separate it from other information). - If some information is emphasized for quick identification, it should still be possible to separate information affecting the actions or decisions from the background. - The application should make it faster and easier to retrieve information from the physical environment.
5. Usability of the device	- The device and application used should match the usage context — for example the user should be able to see what is presented on the display or hear the audio related to the application without any effort. - If a task in physical real world needs to be accomplished simultaneously while using the application (e.g. going to a place or an assembly task), the device used must be appropriate for the task (cf. data glasses would be a better device choice than an application in a handheld device when both hands should be free for the task). - The used device should not be too heavy, difficult to handle or cause depressions on the body disturbing the user.
6. Physical comfort of the use	- Using the application should not cause nausea, headache, eye pain or other physical symptoms. - The usage time of the application should be appropriate — too long usage time may cause for example fatigue or unwanted side-effects. - The user should not end up in difficult and unpleasant positions while using the application .

(continues)

APPENDIX 1. (continues)

7. Usage of the application	- Starting to use the application should be straightforward and should not demand different kinds of procedures that hast to be carried out before the device is ready to use, such as calibration of the device, adjusting usage settings etc. - The application should be stable and not stagnate or stop working while it is being used.
8. Social usage of the application	- Using the application with other users (in physically same place or from distance) should be easy. - Using the application with other users (in physically same place or from distance) should offer added value (for example make the use of application more fun and engaging or help in accomplishing different tasks).
9. Possibility for content production	- If it is possible for the user to create content in the application, the content creation and attaching it to the application should be easy.

APPENDIX 2. First page of evaluation sheet

Usability heuristics for Augmented Reality applications										
MARK ON THIS COLUMN HOW RELEVANT THE CRITERIA ON THE RIGHT SIDE COLUMN IS (e.g. Should it be included in the heuristics): **1 = Not relevant at all** **2** **3** **4 = Very relevant**	**MARK ON THE INTERSECTION OF EACH CRITERIA IN COLUMN B AND CATEGORY ON ROW 3 ONE OF THE FOLLOWING NUMBER VALUES:** **1 = Not related at all** **2** **3** **4 = Strongly related** **NOTE THAT THE CATEGORIES ARE ALSO EVALUATED AGAINST EACH OTHER (THE CATEGORIES ARE ALSO FOUND LOWER IN COLUMN B**	**1. Physical comfort of the use**	**2. Possibility for content production**	**3. Usage of the application**	**4. Usability of the device**	**5. Social usage of the application**	**6. Virtual objects**	**7. Information related to the virtual objects**	**8. Relationship between virtual objects and real world**	**9. Interaction methods and controls**
	1. Switching the attention between the application and the real environment should be smooth and easy.									
	2. If different interaction methods and controls are used in the same application (for example traditional mouse and keypad controls and newer touch displays or gesture control) the co-usage of them should be clear.									
	3. If the user moves while using the application, virtual objects should stay where they are supposed to be situated, not move around.									
	4. If it is possible for the user to create content in the application, the content creation and attaching it to the application should be easy.									
	5. The used device should not be too heavy, difficult to handle or cause depressions on the body disturbing the user.									
	6. The device and application used should match the usage context — for example the user should be able to see what is presented on the display or hear the audio related to the application without any effort.									
	7. The user should have a possibility to choose between different interaction methods, if some of them are better applicable to the usage context.									

APPENDIX 3. Relevance evaluations (CVI)

The criterion (red font and yellow backgrounds on the rows where values 1 or 2 were given)	Evaluator 1	Evaluator 2	Evaluator 3	Evaluator 4	Average value	CVI for scale item (e.g. how many of the 4 evaluators gave a value of 3 or 4 for the criterion)	How many of the 4 evaluators gave a value of 1 or 2 for the criterion
1. Switching the attention between the application and the real environment should be smooth and easy.	4	3	2	3	3	0.75	1
2. If different interaction methods and controls are used in the same application (for example traditional mouse and keypad controls and newer touch displays or gesture control) the co-usage of them should be clear.	3	3	2	3	2.75	0.75	1
3. If the user moves while using the application, virtual objects should stay where they are supposed to be situated, not move around.	4	2	4	4	3.5	0.75	1
4. If it is possible for the user to create content in the application, the content creation and attaching it to the application should be easy.	4	4	2	4	3.5	0.75	1
5. The used device should not be too heavy, difficult to handle or cause depressions on the body disturbing the user.	4	3	4	2	3.25	0.75	1
6. The device and application used should match the usage context — for example the user should be able to see what is presented on the display or hear the audio related to the application without any effort.	4	3	3	3	3.25	1	
7. The user should have a possibility to choose between different interaction methods, if some of them are better applicable to the usage context.	3	3	2	3	2.75	0.75	1
8. The user should not end up in difficult and unpleasant positions while using the application.	4	3	4	2	3.25	0.75	1
9. The user should be aware on his / her location, the virtual objects around him / her, the spatial relations between him- / herself and the expectations of the future state of the environment.	4	3	4	3	3.5	1	

(continues)

APPENDIX 3. (continues)

10. The user should be able to filter the information offered by the application based on his / her own interests and needs — for example, if the application offers lots of information, the user should be able to filter out the information that is unnecessary for him- / herself.	3	3	3	4	3.25	1	
11. There should not be too many virtual objects visible at the same time, so that the view would not be too crowded and messy.	4	3	4	3	3.5	1	
12. If some information is emphasized for quick identification, it should still be possible to separate information affecting the actions or decisions from the background.	3	3	2	3	2.75	0.75	1
13. If a task in real world needs to be accomplished simultaneously while using the application (e.g. going to a place or an assembly task), the device used must be appropriate for the task (cf. data glasses would be a better device choice than an application in a handheld device when both hands should be free for the task).	4	3	4	3	3.5	1	
14. If virtual objects contain text, it should be legible in respect of its size, font, location, colour and how well it can be separated from its background.	4	3	4	3	3.5	1	
15. Starting to use the application should be straightforward and should not demand different kinds of procedures that have to be carried out before the device is ready to use, such as calibration of the device, adjusting usage settings etc.	2	3	4	3	3	0.75	1
16. Using the application should not cause nausea, headache, eye pain or other physical symptoms.	4	4	4	2	3.5	0.75	1
17. Using the application with other users (in physically same place or from distance) should be easy.	4	3	3	3	3.25	1	
18. Using the application with other users (in physically same place or from distance) should offer added value (for example make the use of application more fun and engaging or help in accomplishing different tasks).	4	3	4	2	3.25	0.75	1

(continues)

APPENDIX 3. (continues)

19. The usage time of the application should be appropriate — too long usage time may cause for example fatigue or unwanted side-effects.	3	3	4	2	3	0.75	1
20. The information offered by the application should be organised and grouped clearly (for example different categories and taxonomies visualising different kinds of information).	3	3	3	4	3.25	1	
21. The basis for using the application should be physical real world, e.g. the application should naturally integrate on it — it should not be unnatural or artificial. The application must benefit from the physical real environment - an appropriate basis is for example presenting some additional information about the real world which would be otherwise invisible.	4	3	4	2	3.25	0.75	1
22. The application should make it faster and easier to retrieve information from the physical environment.	3	3	4	3	3.25	1	
23. The application should be stable and not stagnate or stop working while it is being used.	4	3	3	4	3.5	1	
24. The virtual objects of the application should correspond the expectations of the user of the real world objects and their behaviour (for example what can be done with an object, is it possible to study an object like a corresponding real world object, do the objects react to the procedures user has accomplished like corresponding objects in real world environment would react).	4	3	4	3	3.5	1	
25. The interaction methods and controls and their functionalities should be easily recognisable by the user.	4	3	4	3	3.5	1	
26. The relationship between real and virtual objects should be understandable in a way that the user is not confused, for example, trying to manipulate a physical real world object instead of a virtual object.	4	3	4	4	3.75	1	
27. Important information requiring actions from the user should be easily recognisable (for example by highlighting or using other ways to separate it from other information).	4	3	3	3	3.25	1	

(continues)

APPENDIX 3. (continues)

28. Virtual objects should not occlude each other or physical real world objects in a way that disturbs the usage of the application (for example concealing other important information which should be visible).	4	3	4	3	3.5	1	
29. Virtual objects should be accurately aligned with the real world objects linked with them. If, for example, the application offers information about a nearby building, it should be obvious to which building the information is connected. If the application visualises a part which is located inside a device by presenting a virtual image of it augmented on top of the physical device, the virtual image should be aligned accurately to the part of the physical device in which it is actually located inside the device.	4	2	4	4	3.5	0.75	1
30. It should be possible to study the virtual objects from different viewpoints. In a location-based application, the user should typically be able to change between a camera view, map view and list view. When the user is presented with 3D virtual geometrical objects, the user should be able to zoom the object, watch it from different viewpoints and rotate it in order to study it from different angles.	4	3	4	4	3.75	1	
31. The size of a virtual object should be appropriate — not too small, when studying the object is difficult, and not too large, when the object may conceal a too big part of the view and it may be difficult to get an idea of the object.	4	3	4	2	3.25	0.75	1
32. The manipulation of the virtual objects should be carried out in ways which support natural, activity supporting ways. For example, it is possible to "touch" a three-dimensional virtual object through a touch-screen and for example to rotate it using finger instead of manipulating it with traditional menu commands.	4	3	4	4	3.75	1	
33. Virtual objects should be easily separated from their background in different usage situations, e.g. be appropriate by their brightness and contrast in order to be separable.	4	3	4	3	3.5	1	

(continues)

APPENDIX 3. (continues)

34. Virtual objects should adjust to user's movements and changed viewpoints (for example, the user should be able to walk around a three-dimensional object in a similar way if the object was real).	4	3	4	4	3.75	1	
35. Virtual objects should adjust to the physical environment and other visible virtual objects in a way that they seem natural and believable in respect to the distance and location.	4	3	4	2	3.25	0.75	1
36. It should be possible to identify the purpose of virtual action or symbol icons based on their appearance, without a need to read textual explanations. All of them should also be separable from other icons.	4	3	4	3	3.5	1	
37. The application should be tailored for different device platforms (for example, a computer-based application should function in more than one operation systems of Windows, Mac or Linux, in the case of a mobile application, it should be function on more than one of Android, iOS and Windows).	4	2	1	4	2.75	0.5	2
38. Physical comfort of the use	4	3	4	4	3.75	1	
39. Possibility for content production	3	2	2	4	2.75	0.5	2
40. Usage of the application	4	3	4	3	3.5	1	
41. Usability of the device	4	2	4	3	3.25	0.75	1
42. Social usage of the application	4	3	3	2	3	0.75	1
43. Virtual objects	4	3	4	4	3.75	1	
44. Information related to the virtual objects	2	3	4	4	3.25	0.75	1
45. Relationship between virtual objects and real world	4	4	4	4	4	1	
46. Interaction methods and controls	4	3	4	4	3.75	1	

APPENDIX 4. Cohesion evaluations of items

MARK ON THE INTERSECTION OF EACH CRITERIA IN COLUMN B AND CATEGORY ON ROW 3 ONE OF THE FOLLOWING NUMBER VALUES: **1 = Not related at all** **2** **3** **4 = Strongly related** **NOTE THAT THE CATEGORIES ARE ALSO EVALUATED AGAINST EACH OTHER (THE CATEGORIES ARE ALSO FOUND LOWER IN COLUMN B)**	**1. Physical comfort of the use**	**2. Possibility for content production**	**3. Usage of the application**	**4. Usability of the device**	**5. Social usage of the application**	**6. Virtual objects**	**7. Information related to the virtual objects**	**8. Relationship between virtual objects and real world**	**9. Interaction methods and controls**
1. Switching the attention between the application and the real environment should be smooth and easy.	1. Physical comfort of the use	3. Usage of the application	8. Relationship between virtual objects and real world	4. Usability of the device	5. Social usage of the application	6. Virtual objects	9. Interaction methods and controls	7. Information related to the virtual objects	2. Possibility for content production
	3	3	3	2.25	2	2	2	1.75	1
2. If different interaction methods and controls are used in the same application (for example traditional mouse and keypad controls and newer touch displays or gesture control) the co-usage of them should be clear.	9. Interaction methods and controls	3. Usage of the application	1. Physical comfort of the use	4. Usability of the device	7. Information related to the virtual objects	8. Relationship between virtual objects and real world	2. Possibility for content production	5. Social usage of the application	6. Virtual objects
	3.75	3.5	2.75	2.75	1.5	1.5	1.25	1.25	1.25

(continues)

APPENDIX 4. (continues)

4. If it is possible for the user to create content in the application, the content creation and attaching it to the application should be easy.	2. Possibility for content production	1. Physical comfort of the use	3. Usage of the application	4. Usability of the device	9. Interaction methods and controls	5. Social usage of the application	6. Virtual objects	7. Information related to the virtual objects	8. Relationship between virtual objects and real world
	4	2.5	2.25	2.25	2	1.5	1.5	1.5	1.5
5. The used device should not be too heavy, difficult to handle or cause depressions on the body disturbing the user.	1. Physical comfort of the use	4. Usability of the device	9. Interaction methods and controls	3. Usage of the application	8. Relationship between virtual objects and real world	2. Possibility for content production	5. Social usage of the application	6. Virtual objects	7. Information related to the virtual objects
	3.75	3	2.25	1.75	1.5	1.25	1.25	1.25	1.25
6. The device and application used should match the usage context — for example the user should be able to see what is presented on the display or hear the audio related to the application without any effort.	3. Usage of the application	4. Usability of the device	1. Physical comfort of the use	8. Relationship between virtual objects and real world	9. Interaction methods and controls	5. Social usage of the application	7. Information related to the virtual objects	6. Virtual objects	2. Possibility for content production
	3.25	3.25	3	2.75	2.25	2	1.75	1.5	1.25
7. The user should have a possibility to choose between different interaction methods, if some of them are better applicable to the usage context.	9. Interaction methods and controls	3. Usage of the application	1. Physical comfort of the use	5. Social usage of the application	4. Usability of the device	8. Relationship between virtual objects and real world	7. Information related to the virtual objects	2. Possibility for content production	6. Virtual objects
	3.75	2.75	2.5	2.25	2	1.75	1.5	1.25	1.25
8. The user should not end up in difficult and unpleasant positions while using the application.	1. Physical comfort of the use	4. Usability of the device	3. Usage of the application	9. Interaction methods and controls	5. Social usage of the application	2. Possibility for content production	6. Virtual objects	7. Information related to the virtual objects	8. Relationship between virtual objects and real world
	4	3	2.5	2	1.5	1.25	1.25	1.25	1.25

(continues)

APPENDIX 4. (continues)

9. The user should be aware on his / her location, the virtual objects around him / her, the spatial relations between him- / herself and the expectations of the future state of the environment.	8. Relationship between virtual objects and real world	3. Usage of the application	1. Physical comfort of the use	6. Virtual objects	5. Social usage of the application	7. Information related to the virtual objects	9. Interaction methods and controls	4. Usability of the device	2. Possibility for content production
	3.5	2.5	2.25	2.25	2	2	1.75	1.5	1.25
10. The user should be able to filter the information offered by the application based on his / her own interests and needs — for example, if the application offers lots of information, the user should be able to filter out the information that is unnecessary for him- / herself.	3. Usage of the application	1. Physical comfort of the use	7. Information related to the virtual objects	9. Interaction methods and controls	6. Virtual objects	8. Relationship between virtual objects and real world	4. Usability of the device	2. Possibility for content production	5. Social usage of the application
	3	2.75	2.5	2.25	2	2	1.5	1.25	1.25
11. There should not be too many virtual objects visible at the same time, so that the view would not be too crowded and messy.	6. Virtual objects	3. Usage of the application	1. Physical comfort of the use	7. Information related to the virtual objects	8. Relationship between virtual objects and real world	4. Usability of the device	9. Interaction methods and controls	2. Possibility for content production	5. Social usage of the application
	3.75	2.75	2.5	2.5	2.25	1.5	1.5	1.25	1.25
12. If some information is emphasized for quick identification, it should still be possible to separate information affecting the actions or decisions from the background.	1. Physical comfort of the use	7. Information related to the virtual objects	3. Usage of the application	6. Virtual objects	9. Interaction methods and controls	8. Relationship between virtual objects and real world	4. Usability of the device	2. Possibility for content production	5. Social usage of the application
	3	3	2.75	2.25	2.25	1.75	1.33	1.25	1.25

(continues)

APPENDIX 4. (continues)

13. If a task in real world needs to be accomplished simultaneously while using the application (e.g. going to a place or an assembly task), the device used must be appropriate for the task (cf. data glasses would be a better device choice than an application in a handheld device when both hands should be free for the task).	4. Usability of the device	1. Physical comfort of the use	8. Relationship between virtual objects and real world	3. Usage of the application	5. Social usage of the application	9. Interaction methods and controls	6. Virtual objects	7. Information related to the virtual objects	2. Possibility for content production
	3.5	3	2.25	2	2	2	1.5	1.5	1.25
14. If virtual objects contain text, it should be legible in respect of its size, font, location, colour and how well it can be separated from its background.	7. Information related to the virtual objects	1. Pleasantness of the usage	6. Virtual objects	3. Usage of the application	8. Relationship between virtual objects and real world	9. Interaction methods and controls	4. Usability of the device	2. Possibility for content production	5. Social usage of the application
	3.75	3	3	2	1.75	1.75	1.5	1.25	1.25
15. Starting to use the application should be straightforward and should not demand different kinds of procedures that have to be carried out before the device is ready to use, such as calibration of the device, adjusting usage settings etc.	3. Usage of the application	1. Physical comfort of the use	4. Usability of the device	9. Interaction methods and controls	5. Social usage of the application	2. Possibility for content production	6. Virtual objects	7. Information related to the virtual objects	8. Relationship between virtual objects and real world
	3.5	3.25	2.5	2.25	1.5	1.25	1.25	1.25	1.25
16. Using the application should not cause nausea, headache, eye pain or other physical symptoms.	1. Physical comfort of the use	4. Usability of the device	3. Usage of the application	9. Interaction methods and controls	8. Relationship between virtual objects and real world	2. Possibility for content production	5. Social usage of the application	6. Virtual objects	7. Information related to the virtual objects
	4	2.75	2.5	1.75	1.5	1.25	1.25	1.25	1.25
17. Using the application with other users (in physically same place or from distance) should be easy.	5. Social usage of the application	3. Usage of the application	1. Physical comfort of the use	9. Interaction methods and controls	4. Usability of the device	2. Possibility for content production	6. Virtual objects	7. Information related to the virtual objects	8. Relationship between virtual objects and real world
	4	2.75	2.25	2	1.75	1.5	1.5	1.5	1.5

(continues)

APPENDIX 4. (continues)

18. Using the application with other users (in physically same place or from distance) should offer added value (for example make the use of application more fun and engaging or help in accomplishing different tasks).	5. Social usage of the application	3. Usage of the application	1. Physical comfort of the use	9. Interaction methods and controls	8. Relationship between virtual objects and real world	2. Possibility for content production	4. Usability of the device	6. Virtual objects	7. Information related to the virtual objects
	4	2.75	1.75	1.75	1.5	1.25	1.25	1.25	1.25
19. The usage time of the application should be appropriate — too long usage time may cause for example fatigue or unwanted side-effects.	1. Physical comfort of the use	3. Usage of the application	4. Usability of the device	9. Interaction methods and controls	2. Possibility for content production	5. Social usage of the application	6. Virtual objects	7. Information related to the virtual objects	8. Relationship between virtual objects and real world
	4	3	2	1.75	1.5	1.5	1.5	1.5	1.5
20. The information offered by the application should be organised and grouped clearly (for example different categories and taxonomies visualising different kinds of information).	3. Usage of the application	1. Physical comfort of the use	7. Information related to the virtual objects	6. Virtual objects	9. Interaction methods and controls	4. Usability of the device	8. Relationship between virtual objects and real world	2. Possibility for content production	5. Social usage of the application
	2.75	2.5	2	1.75	1.75	1.5	1.5	1.25	1.25
21. The basis for using the application should be physical real world, e.g. the application should naturally integrate on it — it should not be unnatural or artificial. The application must benefit from the physical real environment - an appropriate basis is for example presenting some additional information about the real world which would be otherwise invisible.	8. Relationship between virtual objects and real world	3. Usage of the application	6. Virtual objects	7. Information related to the virtual objects	1. Physical comfort of the use	4. Usability of the device	5. Social usage of the application	9. Interaction methods and controls	2. Possibility for content production
	3.75	2.25	2	2	1.5	1.5	1.5	1.5	1.25

(continues)

APPENDIX 4. (continues)

22. The application should make it faster and easier to retrieve information from the physical environment.	8. Relationship between virtual objects and real world 3.75	1. Physical comfort of the use 2.25	7. Information related to the virtual objects 2.25	3. Usage of the application 1.75	6. Virtual objects 1.75	9. Interaction methods and controls 1.75	2. Possibility for content production 1.5	4. Usability of the device 1.5	5. Social usage of the application 1.25
23. The application should be stable and not stagnate or stop working while it is being used.	1. Physical comfort of the use 3.75	3. Usage of the application 2.75	4. Usability of the device 2.5	9. Interaction methods and controls 2	2. Possibility for content production 1.25	5. Social usage of the application 1.25	6. Virtual objects 1.25	7. Information related to the virtual objects 1.25	8. Relationship between virtual objects and real world 1.25
24. The virtual objects of the application should correspond the expectations of the user of the real world objects and their behaviour (for example what can be done with an object, is it possible to study an object like a corresponding real world object, do the objects react to the procedures user has accomplished like corresponding objects in real world environment would react).	8. Relationship between virtual objects and real world 4	6. Virtual objects 3	7. Information related to the virtual objects 2.75	9. Interaction methods and controls 2.75	3. Usage of the application 2	1. Physical comfort of the use 1.75	4. Usability of the device 1.5	2. Possibility for content production 1.25	5. Social usage of the application 1.25
25. The interaction methods and controls and their functionalities should be easily recognisable by the user.	9. Interaction methods and controls 4	3. Usage of the application 2.75	1. Physical comfort of the use 2	4. Usability of the device 1.75	7. Information related to the virtual objects 1.75	5. Social usage of the application 1.5	6. Virtual objects 1.5	8. Relationship between virtual objects and real world 1.5	2. Possibility for content production 1.25

(continues)

APPENDIX 4. (continues)

26. The relationship between real and virtual objects should be understandable in a way that the user is not confused, for example, trying to manipulate a physical real world object instead of a virtual object.	8. Relationship between virtual objects and real world 3.75	7. Information related to the virtual objects 3	6. Virtual objects 2.5	1. Physical comfort of the use 2.25	9. Interaction methods and controls 2.25	3. Usage of the application 2	4. Usability of the device 1.75	2. Possibility for content production 1.25	5. Social usage of the application 1.25
27. Important information requiring actions from the user should be easily recognisable (for example by highlighting or using other ways to separate it from other information).	7. Information related to the virtual objects 2.75	9. Interaction methods and controls 2.75	1. Physical comfort of the use 2.25	3. Usage of the application 2.25	6. Virtual objects 2	8. Relationship between virtual objects and real world 1.75	4. Usability of the device 1.5	2. Possibility for content production 1.25	5. Social usage of the application 1.25
28. Virtual objects should not occlude each other or physical real world objects in a way that disturbs the usage of the application (for example concealing other important information which should be visible).	6. Virtual objects 3.75	8. Relationship between virtual objects and real world 3	1. Physical comfort of the use 2.75	7. Information related to the virtual objects 2.75	3. Usage of the application 2.5	9. Interaction methods and controls 1.5	2. Possibility for content production 1.25	4. Usability of the device 1.25	5. Social usage of the application 1.25
29. Virtual objects should be accurately aligned with the real world objects linked with them. If, for example, the application offers information about a nearby building, it should be obvious to which building the information is connected. If the application visualises a part which is located inside a device by presenting a virtual image of it augmented on top of the physical device, the virtual image should be aligned accurately to the part of the physical device in which it is actually located inside the device.	8. Relationship between virtual objects and real world 3.75	6. Virtual objects 3.25	7. Information related to the virtual objects 2.75	1. Physical comfort of the use 2.25	3. Usage of the application 2.25	9. Interaction methods and controls 1.5	2. Possibility for content production 1.25	4. Usability of the device 1.25	5. Social usage of the application 1.25

(continues)

APPENDIX 4. (continues)

30. It should be possible to study the virtual objects from different viewpoints. In a location-based application, the user should typically be able to change between a camera view, map view and list view. When the user is presented with 3D virtual geometrical objects, the user should be able to zoom the object, watch it from different viewpoints and rotate it in order to study it from different angles.	6. Virtual objects	9. Interaction methods and controls	3. Usage of the application	7. Information related to the virtual objects	8. Relationship between virtual objects and real world	1. Physical comfort of the use	5. Social usage of the application	2. Possibility for content production	4. Usability of the device
	3.75	3	2.5	2.5	2.5	2.25	1.5	1.25	1.25
31. The size of a virtual object should be appropriate — not too small, when studying the object is difficult, and not too large, when the object may conceal a too big part of the view and it may be difficult to get an idea of the object.	6. Virtual objects	1. Physical comfort of the use	7. Information related to the virtual objects	8. Relationship between virtual objects and real world	3. Usage of the application	5. Social usage of the application	9. Interaction methods and controls	2. Possibility for content production	4. Usability of the device
	3.75	2.75	2.5	2.5	2.25	1.5	1.5	1.25	1.25
32. The manipulation of the virtual objects should be carried out in ways which support natural, activity supporting ways. For example, it is possible to "touch" a three-dimensional virtual object through a touch-screen and for example to rotate it using finger instead of manipulating it with traditional menu commands.	9. Interaction methods and controls	6. Virtual objects	1. Physical comfort of the use	3. Usage of the application	8. Relationship between virtual objects and real world	7. Information related to the virtual objects	4. Usability of the device	2. Possibility for content production	5. Social usage of the application
	4	3.5	2.75	2.25	2.25	2	1.75	1.5	1.25
33. Virtual objects should be easily separated from their background in different usage situations, e.g. be appropriate by their brightness and contrast in order to be separable.	6. Virtual objects	1. Physical comfort of the use	3. Usage of the application	7. Information related to the virtual objects	8. Relationship between virtual objects and real world	4. Usability of the device	9. Interaction methods and controls	5. Social usage of the application	2. Possibility for content production
	3.75	2.75	2.75	2.5	2.5	2	1.75	1.5	1.25

(continues)

APPENDIX 4. (continues)

34. Virtual objects should adjust to user's movements and changed viewpoints (for example, the user should be able to walk around a three-dimensional object in a similar way if the object was real).	6. Virtual objects	8. Relationship between virtual objects and real world	3. Usage of the application	9. Interaction methods and controls	1. Physical comfort of the use	7. Information related to the virtual objects	4. Usability of the device	5. Social usage of the application	2. Possibility for content production
	3.75	3.25	2.5	2.5	2	2	1.75	1.5	1.25
35. Virtual objects should adjust to the physical environment and other visible virtual objects in a way that they seem natural and believable in respect to the distance and location.	8. Relationship between virtual objects and real world	6. Virtual objects	3. Usage of the application	1. Physical comfort of the use	7. Information related to the virtual objects	4. Usability of the device	9. Interaction methods and controls	2. Possibility for content production	5. Social usage of the application
	3.75	3.25	2.25	2	2	1.75	1.5	1.25	1.25
36. It should be possible to identify the purpose of virtual action or symbol icons based on their appearance, without a need to read textual explanations. All of them should also be separable from other icons.	7. Information related to the virtual objects	9. Interaction methods and controls	3. Usage of the application	6. Virtual objects	1. Physical comfort of the use	8. Relationship between virtual objects and real world	4. Usability of the device	2. Possibility for content production	5. Social usage of the application
	3.5	3	2.75	2.75	2.5	1.75	1.5	1.25	1.25
37. The application should be tailored for different device platforms (for example, a computer-based application should function in more than one operation systems of Windows, Mac or Linux, in the case of a mobile application, it should be function on more than one of Android, iOS and Windows).	4. Usability of the device	3. Usage of the application	1. Physical comfort of the use	9. Interaction methods and controls	5. Social usage of the application	6. Virtual objects	2. Possibility for content production	7. Information related to the virtual objects	8. Relationship between virtual objects and real world
	3	2.75	2	1.75	1.5	1.5	1.25	1.25	1.25

APPENDIX 5. Cohesion evaluations of categories

MARK ON THE INTERSECTION OF EACH CRITERIA IN COLUMN B AND CATEGORY ON ROW 3 ONE OF THE FOLLOWING NUMBER VALUES: 1 = Not related at all 2 3 4 = Strongly related NOTE THAT THE CATEGORIES ARE ALSO EVALUATED AGAINST EACH OTHER (THE CATEGORIES ARE ALSO FOUND LOWER IN COLUMN B)	1. Physical comfort of the use	2. Possibility for content production	3. Usage of the application	4. Usability of the device	5. Social usage of the application	6. Virtual objects	7. Information connected to the virtual objects	8. Relationship between virtual objects and real world	9. Interaction methods and controls
1. Physical comfort of the use									
2. Possibility for content production	2								
3. Usage of the application	3.25	2							
4. Usability of the device	2.5	1.5	2.75						
5. Social usage of the application	2.5	1.75	2.75	1.75					
6. Virtual objects	1.5	1.5	2	1.75	1.5				
7. Information related to the virtual objects	1.5	2	2	1.25	1.5	3.5			
8. Relationship between virtual objects and real world	1.5	1.75	2	1.75	1.5	3.5	3.5		
9. Interaction methods and controls	3	2.5	2.75	2.25	2	2	2.25	2.25	

APPENDIX 6. References for AR heuristics

Heuristics	References
AR1 Interaction methods and controls	AR allows integrating different input and output devices and interaction techniques, and different tasks may benefit of them. None of them is best in itself, since their performance is task and environment dependent. User also has more choice based on his / her preferences. Different devices and techniques may also complement each other, since each of them have advantages and disadvantages. Selection of appropriate displays, using headsets with portable devices and haptic output while other senses cannot allow operating in different kinds of environments and usage situations. Multimodality can increase performance. Inappropriate input and output devices and interaction techniques may cause cognitive overhead which distracts user's focus. If interaction with the application is fluid and intuitive, user can concentrate on the task (s)he is accomplishing. New methods may allow this and make the computer vanish into familiar objects. Interaction should also correspond to the user's expectations of interaction in the real world as much as possible. Interaction designers need to understand the ergonomics, advantages and limitations of them in order to make natural mappings between interaction techniques and hardware. The integration of 2D and 3D interaction techniques is important, since it should not be difficult for the user to switch between them while using the application. Manipulations should not be beyond the physical capabilities of the users. User should be able to easily identify the selection and manipulation points in the virtual objects, and objects should not occlude each other, which makes manipulation difficult. Even though these issues are challenging, some principles are available. On the other hand, since the devices and interaction modalities of AR may be new for the user, getting used to them may require time. (Dünser et al. 2007, 38–40; Billinghurst et al. 2005, 17–18; Bowman et al. 2001, 97–98, 102–103; Sutcliffe & Gault 2004, 833; Sutcliffe & Kaur 2000, 419–420; Stanney et al. 2003, 463–466; Pribeanu et al. 2009, 180; Martín-Gutiérrez et al. 2010, 303; Ko et al. 2013, 507; Wang & Dunston 2009, 15; Kaufmann & Dünser 2007, 663; Ganapathy 2013, 179; Li & Duh 2013, 118–121; Azuma et al. 2001, 37-38.)
AR2 Presentation of virtual objects	Presentation of virtual objects combined in real environment cause certain requirements, which are crucial to avoid causing the user cognitive load and frustration. To identify the objects and their attributes, sufficient acuity and contrast in relation to their surroundings is needed. The contrast ratio between the text and the drawing style and between the drawing style and the background may affect text legibility. Transparency of the objects can help in avoiding occlusion and concealing other objects or items of interest in the real environment. The amount of information is also crucial, since especially in small mobile devices the screen becomes easily cluttered. On the other hand, too little information is not good either. Objects that are part of the environment and necessary to carry out actions should be visible, easily located and clearly presented, easily separable from other objects. Objects should also be designed properly, and the rendering quality should be good enough to capture the environmental illumination and reflectance information. Necessary objects may also be highlighted and rendered in more detail. AR presentation schemes such as arrows for navigation guidance help the users interpret the distance. The user should be able to examine objects from multiple perspectives and viewpoints, for example, from egocentric and exocentric, depending on the situation. For example, exocentric viewpoint serves as providing better overview of the surrounding context, while egocentric viewpoint present information for local guidance. When virtual information is attached to the real world surroundings, poor and slow tracking performance may interfere the user and task performance. Error estimations, environmental conditions like changing light and human errors may cause inaccuracies. Virtual information may jump, jitter or suddenly disappear. While user moves, virtual objects should follow the changed viewpoint and position. If tracking fails, dynamic registration errors, which have no effect until

(continues)

APPENDIX 6. (continues)

	the viewpoint or objects move, are caused. There is knowledge available about the visual update rates which are appropriate in avoiding dynamic registration errors. One solution may be using hybrid tracking techniques, or at least selecting a technique which is suitable for the application and its usage. (Livingston 2005, 8; Dünser et al. 2007, 40; Stanney et al. 2003, 464–466; Sutcliffe & Gault: 2004, 833; Sutcliffe & Kaur 2000, 419; Martín-Gutiérrez et al. 2010, 303; Azuma 1997, 373; Gabbard & Swan 2008, 523; Ko et al. 2013, 507; Vallino 2008, 19; Wang & Dunston 2009, 97; Ganapathy 2013, 177–179; Li & Duh 2013, 112–117; Azuma et al. 2001, 36, 39–40.)
AR3 Relationship between virtual objects and real world	If there is a clear and meaningful relationship between virtual objects and the real world, AR may prove to be a good choice. If there is no reason to associate the virtual object with an aspect of reality, AR may not be the best choice of technology. At its best, when cognitive components are integrated as a part of manual work, the accomplishment of tasks may be enhanced and speed up. Items that should be remembered may also be recalled better while they are virtually situated in real-world locations. AR may, for example, display real objects with virtual cues. In order to perceive these properly, the virtual cues must be accurately aligned, or registered, to the real environment so that the user perceives them as attached to each other. Static registration errors are caused when the user's viewpoint and the objects in the environments remain still, but the physical and virtual object are not properly aligned in relation to each other. It is task-dependent how much registration error can be allowed. Even small amounts of registration error may cause too much cognitive load for the user while accomplishing certain tasks. How much user is able to enjoy some AR applications may depend on the suspension of disbelieve, and in these cases, registration errors may also be a breaking point for natural interaction. Photorealistic rendering and realistic lighting of the augmented objects may improve the illusion of virtual objects appearing as a part of physical reality, also visual occlusions between virtual and real objects and appropriate depth cues should be carried out correctly. Interaction should correspond to user's expectation of the interaction in the real world, and same applies to the behaviour of the virtual objects and affordances. The effect of user's actions on virtual objects should conform the laws of physics. While location-based AR applications are used, user should receive help in wayfinding, not feeling disoriented. The user should also be spatially aware of surrounding objects and spatial relationships between the objects and self. If the user changes viewpoint between the application and real world, the transitions between different perspectives should be smooth in order to avoid extraneous cognitive load. (Carmichael et al. 2012, 1768; Livingston 2005, 7–8; Billinghurst et al. 2005, 17–18; Dünser et al. 2007, 38–39; Sutcliffe & Gault 2004, 833; Stanney et al. 2003, 463; Pribeanu et al. 2009, 180; Wang and Dunston 2009, 26–28, 97; Neumann & Majoros 2008, 4–6; Vallino 2008, 19–20; Azuma 1997, 372–373; Li & Duh 2013, 113–117.)

continues)

APPENDIX 6. (continues)

AR4 Information related to virtual objects	Augmented virtual objects often present additional information to the user. Organisation of the information is important, so that the user can identify it easily and filter the information (s)he specifically needs, at the same time limiting the amount of information presented. For example, in an outdoor application, information filter based on the distance of an object should be available. The icon may also indicate whether the labelled item is visible (for example, a building may be behind another building). It should be easy to identify different information based on object's appearance, without needing to read the text label. Information that may need action, especially in the case of training or medical applications, should be highlighted. The placement of virtual information is important, because it may affect the understandability of the information. Appropriate arrangement may help the user to connect the meaning of virtual information. Information may be presented at two stages, allowing the user browse general information first and then after selecting favourite spots, additional and detailed information may be shown. Visualisation techniques may be used to increase the efficiency of searching information. (Ganapathy 2013, 178–179; Li & Duh 2013, 114–117.)
AR5 Suitability for the usage context	Technology selections in case of an AR application may have a crucial impact how usable the application is. For example, the device choices and interaction method choices and altogether the preliminary choice of AR as a technology all affect to the usability. If user must have his / her hands free at the same time while using an AR application on mobile phone, usability suffers. The display size and possibility to get audio or haptic feedback may increase the usability of an application in some circumstances. User interface should be designed considering different usage environments. User scenarios should be mapped to tasks. Task mental requirements (perceptual, like object recognition), cognitive (like discrimination and comparison), working environments (factors involved in the working environment (like situational awareness requirements, indoor / outdoor location, noise level, work area hazards, working volume), physical disposition (motion, body position) and hand occupation should be taken as a starting point for designing a system. Different tracking technologies work in a different way in different environments. (Billinghurst et al. 2005, 18; Bowman et al. 2001, 98, 102; Stanney et al. 2003, 465–466; Ko et al. 2013, 507; Wang & Dunston 2009, 18–20, 21–23, 33–37, 97–99; Pribeanu et al. 2009, 180.)
AR6 Physical comfort of use	Ergonomic issues are an important aspect of the usability of an AR application. The application and device should cause low physical effort for the user, and the task should be accomplished with a minimum of interaction steps, without unnecessary interventions. The system should react to user's actions efficiently reducing the likelihood of fatigue. If user has to wear parts of the system, they should be as lightweight and comfortable as possible. Simulator sickness can diminish usability of a system. The response time between user movement and update of the display should be as small as possible to avoid motion sickness. Image quality of HMDs may also cause unwanted effects. Different symptoms which are possible are general discomfort, headaches, eye-pain, increased salivation, sweating, high levels of body fatigue, nausea and burping. Also aftereffects may be caused, like blurred vision, dizziness, nausea, difficulty focusing, difficulty with coordinated tasks, visual or coordination aftereffects and vertigo. Usage times of the applications should also be considered. (Dünser et al. 2007, 39; Sutcliffe & Gault 2004, 833; Stanney et al. 2003, 467; Pribeanu et al. 2009, 180; Kaufmann & Dünser 2007, 667–668; Ko et al. 2013, 507; Wang & Dunston 2009, 97–98; Vallino 2008, 20; Azuma et al. 2001, 40.)