A Survey of Usability Issues in Mobile Map-based Systems

Sadegh Karampanah

Supervisor: Prof. Dr. Christian Kray

Co-supervisors: Dr. Morin Ostkamp and Dr. Sven Casteleyn

Submitted in partial fulfilment of the requirements for the degree of

Master of Science in Geospatial Technologies

Institute for Geoinformatics (ifgi)

University of Münster

February 2019

ACKNOWLEDGEMENTS

First, thanks to GOD, the main source of love and energy, then thanks to Prof. Dr. Christian Kray for teaching us how to convey our love and energy to the knowledge and for his kindness that always during his supervision was helping and even during holidays and the times he was travelling.

Thanks to my co-supervisor, Dr. Morin Ostkamp for his constructive criticisms always was helpful and many thanks to my second co-supervisor Dr. Sven Casteleyn from University of Jaime I.

My thanks to all academic and administrative staffs of institute for geoinformatics (ifgi) in University of Muenster.

ABSTRACT

How geospatial information could be represented in map or other forms of communication to display in mobile phones to convey spatial knowledge to users more effective and efficient with less ambiguity? This triggering question stems from the usability problems available in mobile map-based systems, that made using mobile navigation services and applications for pedestrians, a tedious and complicated task which is rather confusing to be helpful. Problems such as; losing the spatial overview of the area, overload of information in small screens of mobile phones, visibility issue of off-screen entities, weaknesses in orienting users with real environment, too much engagement of users with interface which causes environment distraction and so on. There are a lot of solutions have proposed to mitigate these available issues in mobile map-based systems, but each one has its pros and cons that is not complete enough to tackle above mentioned issues alone, and most of the time a combination of them is proposing. We tried with systematic literature review (SLR) that is more reliable, replicable and valid [1], find the most frequently applied usability evaluation method in the available studies to detect the usability issues in mobile map-based systems (MMSs), then find the most frequently usability issues that detected among the reviewed literatures and how to categorize them, in what contexts they mostly happened and what solutions proposed so far to resolve them.

We operated tree iterations of systematic literature review (SLR) with totally 8667 identified publications (within 6 relevant databases and a search engine with priority of 4 most prominent and relevant journals and conferences in the field of mobile HCI and

location based services), that 196 one of them included in first screening[1], were thoroughly read in order to check with predefined inclusion criteria and overall, 56 papers (between those 196 papers) that qualified with our well-defined and updated inclusion criteria properties read in-depth at least two times to extract the data. In the first iteration 25 papers have reviewed and relevant data with considering our research questions has extracted and reflected in the first iteration table. In the second iteration, 24 papers which had adjusted inclusion criteria parameters have included to data extraction for filling the updated table. The last iteration according to the scarcity of publications in this realm and time limitation, has operated only with 7 literatures and relevant data extracted to fill in the last updated table.

Results of the SLR showed the most frequently usability evaluation method was "*Questionnaire*" to achieve *effectiveness* and *efficiency* of the system, and the most frequently usability issue that detected within available literatures was "*losing the spatial overview*" which followed by "*too much zooming and panning operations by users*" that stems from the same problem; small screen size of mobile devices. We categorized the issues into two main groups of technological and spatial issues, which we only here focused on the usability issues relevant to map interfaces in mobile phones (spatial issues), not the technological problems relevant to the server or the hardware perspective (sensors, connectivity, battery drainage, GPS accuracy etc.). We have noticed the most frequently usability issue has happened in the mobile phone with average screen size of 3.83 inches, 87% of the cases in the laboratory environment, with users (not experts) with average age of 26 years old that 64.2% of them had relevant knowledge (GI[2] knowledge).

[1] First screening only done by scanning the title, keywords, abstract and in some cases the conclusion section

[2] Geospatial information (recruited from students, alumni or authorities of GIScience field that had at least a basic geographical knowledge)

The low amount of field-based studies highlights the lack of considering real context in available case studies that in usability evaluation of location based mobile systems is highly important. Some traditional solutions have proposed to address the most frequently occurred usability problem in mobile map-based systems such as the techniques for visualizing the off-screen objects (such as Overview&Detail, Scaled Arrows, Wedge etc.) and some techniques for enhancing the zoom and pan operations (such as vario-scale maps, semi-automatic zooming (SAZ), tilt zooming, content zooming, anchored zoom etc.) that none of them were not completely suitable enough to be applied in these systems and the most famous systems such as Google Maps still working without taking advantage of such approaches, techniques and widgets, with a lot of usability issues.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

UCD	User-centered design
UX	User Experience
HCI	Human Computer Interaction
SLR	Systematic Literature Review
LBS	Location Based Services
GIS	Geographic Information System
MMSs	Mobile Map-based Systems

LIST OF FIGURES

LIST OF TABLES

Chapter 1: Introduction

1.1 Introduction

Today, the use of mobile devices is growing, and mobile phones have become an important inseparable part of the people's life. People are using their mobile phones to do many daily tasks and sometimes they are facing some problems in working with applications or websites, especially when their tasks are spatial. According to some limitations and difficulties that these touch-based and small screen devices have, performing some tasks that related to maps can be challenging for the mobile users. Some of these common challenges are for example successive zooming and panning interactions which arising from smallness of the screen of such mobile devices, bothers users and confuse them in term of acquiring the spatial knowledge of the geographic area and also some issues around map representations such as the level of details that should be represented to mobile users and landmarks and other representation hints that need to be reviewed carefully. According to Jiamsanguanwong et al. [2] usability test is an evaluation method to identify user experiences and errors from the interface design. They believed that, with usability test not only the problems can identify, but also the high concern problems can be separated. They added, without usability test, the applications would have a complexity. This complexity in mobile touch-based interfaces and especially in map services, might cause avoidance of use of such devices and services by old users or people with low technology affinity or low "*Sense-of-Direction*" [3]. Despite a tons of studies in developing and implementing mobile map-based systems, there is not enough attention paying to the map-based usability evaluation in industry and academia in context of mobile devices to address them and most of the available mobile map-based systems (MMSs) still have some usability issues that these problems might be the reason that interacting and using them are not easy to everyone (mostly for people with poor technological affinity).

1.1.1 Aim and Objective

We are conducting a systematic literature-based review in order to overview the usability issues that have detected in map-based mobile systems (MMSs) and reported in the scientific publications to achieve a deeper insight and be noticed of some available trends through carefully studying and analysing the empirical works have done so far that reflected in those reviewed literatures and find some possible gaps and shortages in their studies. The outcomes of this review can contribute in providing producers, designers and researchers in this realm, a broader view about the most common map interaction and technological issues related to the concept of map, the available methods for detecting these issues (usability evaluation methods), the solutions that proposed to tackle them so far, and might also looking for the reasons behind them to occur to finally with a deep knowledge that we are gaining from the available issues and barriers in the way of representing the spatial information in mobile devices, to have some useful recommendations for designers and researchers to enhance the usability of mobile map-based systems (MMSs).

All in all, these struggles lead the designing, producing and evaluating the mobile map-based systems, according to user-centered design (UCD) principles, to a direction that could help mobile users, which because of ubiquitousness nature of mobile devices, almost are novice in GIS[3], with minimum time and effort, easily achieve better spatial

[3] Today the "USER" role in GIS has changed in comparison to previous decades. Before, users of these systems were only GIS experts, but now, GIS has become ubiquitous, and a wide range of the people in society is dealing with maps (e.g. mobile maps) that has an explicit effect on people's daily tasks (e.g. navigation).

understanding to execute their spatial tasks more effective (e.g. one of the most common spatial tasks is navigation).

1.1.2 Theoretical Framework and Background

1.1.2.1 Mobile Map-based Systems (MMSs)

According to Elzakker et al. [4] Mobile Map-based Systems contain *Positioning*, *GeoData* and *Mobile Maps* that differentiated this realm from other GI systems. The first component, *Positioning*, refers to the way that the position of the mobile device (user) is representing on a coordinate system by some technologies such as RFID, Bluetooth, Laser, Ultrasound, Global Positioning System (GPS) etc. The position of the mobile device (user) is representing by means of the second component, *GeoData* as a 2-dimension or 3-dimension or with considering time, that could be 4-dimension, usually representing with geographical features (spatial entities) with different formats. The last component (*Mobile Maps*) makes the domain exclusively different than desktop GIS. The model of reality needs to represent on a small screen of mobile devices in a form of Augmented Reality, Photorealistic or Panorama views, textual or verbal guidance, Vibro/gaze-based interactions or Cartographic map displays, which the latter one is the most prominent form of representation on MMSs. The Geospatial information usually is showing in a static or dynamic form with raster or vector formats. But something that making the cartographic design for such systems completely different than paper maps or desktop GIS is the limitations that such systems have such as small screens, which induces users to do a lot of zooming and panning (scrolling in desktop applications) operations to acquire overall and detail understanding of Geospatial information at the same time. Here the representing map needs special sophistication in design with using some cartographic techniques such as generalization, colour codding, size, form, and taking advantages of some important entities to link between reality, mobile maps and mental map.

1.1.2.2 Usability evaluation

ISO [4] (International Organization for Standardization) defines usability as the "extend to which a product can be used by specific users to achieve specified goals with effectiveness, efficiency and satisfaction in a specific context of use".

Usability evaluation of mobile devices should be different than the way of evaluating desktop systems. In evaluation of Mobile Map-based Systems (MMSs), since users mostly are in moving in the real environment with real-time positioning and exposing to the natural outdoor conditions[5] with small screen sizes, the user's context plays an important role. Although, according to Elzakker et al. [5], [4] most of the studies (81%) on the usability evaluation of mobile geo-applications are executed in the laboratory, which a big part of contextual information cannot be investigated and real behavior of user and activities may not sufficiently be understood [5]. They argument that the reason for executing a greater number of the user studies in lab might be the high cost of human and material resources that need for operating field studies. It is not easy to categorize the usability evaluation methods, for example in the lab or in the field or by end users or experts (which latter one calls heuristic evaluation) and in which stages of system development they are conducting. As shows in figure 1-1, the usability evaluation in the software development procedure can be held at the last stage of requirement analysis that [4] believes most often in this stage quantitative methods are using and qualitative research will be executed more in the earlier stages of UCD process, although this important stage of human-centred design has an iterative manner in the ISO's

[4] ISO 9241-11 (1998)

[5] The different context that interaction with such devices has such as weather situation (daily sunlight, precipitation etc.), environmental distractions in crowed cities, incoming calls and messages etc. making some interruptions.

ergonomics, i.e. whenever the design solutions does not meet the user requirements, this stage is going back to the first stage.

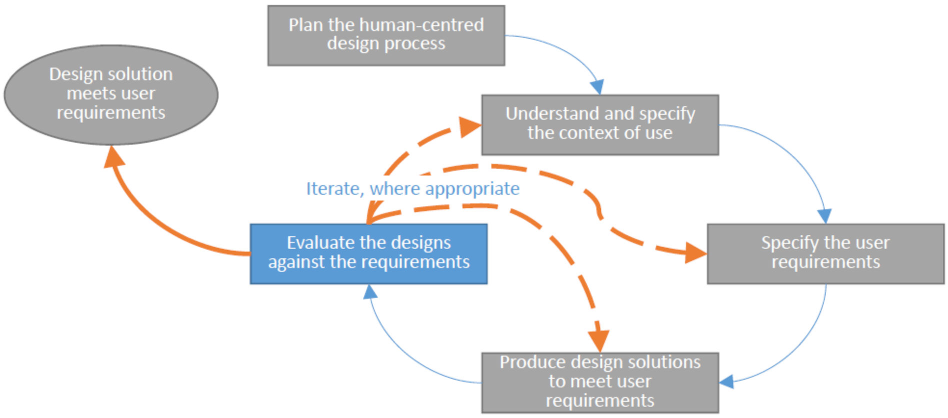

Figure 1-1: software development process of ergonomics of human-cantred design (ISO 9241-210)

Elzakker in 2004, categorized the evaluation tools for collecting qualitative and quantitative data from representative users in four groups; interview, questionnaire, observation and product analysis [5].

Interview can be in-depth or unstructured that questions are formulated spontaneously, albeit within an interview framework. The advantage of this kinds of interview is a lot of in-depth information can be achieved but comparing the answers of different respondents is difficult [5] [4].

Questionnaire, according to Wikipedia is one of the most frequently used method for subjective usability evaluation that is cheap, without a need for verbal or other efforts, with standardized answer that is simple to compile and compare and analyse. It has also some drawbacks that has too few options to answer (users are limited to questions, except open-ended), people might have really positive or really negative viewpoint or who are most likely unbiased, typically don’t respond because they might think it is not worth their time. The most usable types of questionnaire are; NASA TLX (measuring workload), USE (measuring ease of use, learnability, satisfaction, usefulness), SUS (measuring effectiveness, efficiency and satisfaction).

In observation method, investigator might in simple cases watch the subjects and take some notes or with some equipment record the observations or with "logged data" (screen recording through some injected proxies to the system) or "eye tracking" might record some data.

When users need to execute task(s) with an existing application or a prototype, it calls product analysis.

Flink et al. [6] claimed with thinking aloud and questionnaire they were achieving results for concrete input for the design process of a map service. Think aloud is a usability evaluation method that when users performing designated tasks with the system, all the time verbalizing their thoughts out loud, and the evaluator is recording the voice for the data analysis.

1.1.2.3 Systematic Literature Review (SLR)

Literature review is a kind of secondary study (i.e. studies that are based on analysing previous research) [7] that overviewing some primary studies in order to achieve some insights, statistics, results, trends and gaps out of aggregating the results from those conducted studies. Conducting literature review systematically, can better lead researchers to achieve the outcomes of the literature review and every stage of the review should be documented transparently. As Xiao et al. [1] said, with systematic literature review the quality, replicability, reliability and validity of review can enhance. The process of literature review can be iterative. During conducting the review, unforeseeable problems may appear that needs modifying the research questions and even the topic and consequently the inclusion criteria to find relevant studies [1], therefore our approach is conducting systematic literature review (SLR) in an iterative manner. In fact, the literature review using to aggregate the experiences gained from different studies (that such studies may employ very different experimental forms and contexts) in order to answering the research question(s) [1].

Conducting SLR in different realms are following different approaches, for example, in the field of medicine, medical guidelines for performing SLRs recommend a kind of broad search procedures including automated searching which includes any relevant grey literatures[6] that is different than the approach that the researchers follow for example in software engineering (SE) [8].

Schoen et al. [9] executed a SLR with 27 papers within 10 months in order to derive deep insights to some aspects of requirement engineering (RE) of AGILE software development. They used some specific places to search for their literature with considering some inclusion and exclusion criteria.

Yusop et al. [10] conducted a SLR with 57 papers (published from the year 2000 to 2016) in the domain of software engineering to make some recommendations to improve usability defect reporting. They used 5 electronic database resources with some search strings that in their first screening only the title and abstract were analysing and the second stage of analysis were done by reading full papers that were considering some inclusion and exclusion criteria.

Lacerda et al. [11] performed a SLR proposed by Kitchenham et al with totally 15 papers (published from the year 1993 to 2017) that found them in only two defined repositories (Google Scholar and SCOPUS) in usability engineering. They ordered the results of Google Scholar by relevance and only screened the first 150 results. Their first

[6] Gray literature refers to papers that have not been published in a source with full peer review process that includes technical reports and thesis too.

screening was according to quickly reviewing the title, abstract and keywords to identify if the papers matched the inclusion/exclusion criteria.

1.1.3 Research Questions

After conducting some preliminary researches in the field of mobile map-based applications and achieve the necessity and importance of the research in this realm[7], an approximate broad view of the topic attained, and the research questions formulated. Investigating the problems that are available in maps that are presenting in mobile devices, first needs a deep evaluation of available methods that researchers applied so far to detect them, which methods have used more frequently, which methods are suitable for detecting a special kinds of usability problems (the most reoccurring ones) and so on. How these usability problems can be categorized in terms of their importance and their nature. How, where and when they might happen and what possible factors might provoke them to happen. In reviewing the available empirical scientific works, we can recognize what solutions have applied to tackle these usability issues in mobile map-based systems (MMSs) and how much they have been successful so far. To address these ambiguities, we have formulated the following research questions that the study is trying to answer them with operating an iterative systematic literature review (SLR).

- RQ1: What usability evaluation method is more frequently used to detect usability defects (issues) in mobile map-based systems (MMSs) according to available studies?
- RQ2: What are the most frequent usability issues in mobile map-based systems (MMSs) that reported in the relevant literature?

[7] There are a few numbers of works have done in this realm in comparison to works in the GI desktop systems or mobile systems without map aspects.

a. How to categorize them?

b. In what contexts they are happening?

c. What methods have developed so far for resolving them according to the available literature?

1.2 Methodology

According to Xiao et al. [1] “Literature review is an essential feature of academic research.” With literature review the researchers can understand the “breadth and depth” of the existing body of work and also be familiar with their methodologies and identify the gaps and then according to those works, can come up with new methodologies to operate their research [1]. According to Kitchenham et al. [7] a successful review involves three major stages: planning the review, conducting the review and reporting the review. In planning stage, researchers first identify the need for a review, then specify research questions and finally develop a review protocol. Here, before start to conduct the review (as described in the previous section), with some preliminary studies some primary keywords extracted to input to the first stage of our SLR that we called this stage, *searching*. In conducting stage, after identifying primary studies to review they should extract, analyse and synthesize data (Analysing stage). Here, we constructed a big table (Appendix B) with a primary list of criteria to extract the data in the analysing stage. The last but not the least, within reporting stage, researchers write the report to publicize their findings from the literature review (we call it reflecting stage here) [7]. In the last stage we here, reflecting the results of the SLR. These stages in our work has an iterative manner, which means after fulfilling those above mentioned three stages, the next iteration will be started with same structure (searching, analysing and reflecting) again (with doing calibration of the search terms, inclusion and extraction criteria) and so on. Whenever no new (or repetitive) results achieved, or the time schedule limited us to continue, the process of iteration can be stopped. The reason why this iterative method applied is, with gaining new knowledge about the topic after the first iteration, the criteria for extracting new data will be updated to extract more relevant information to achieve the more relevant goals and objectives. This flow is also repeating for the next iterations. Here, we have conducted systematic literature review with three iterations according to our time limitation (from October 2018, the first round of iteration to February 2019 the last round of the iteration) for reviewing 56 papers (25 papers for first and 24 papers for the second iteration, and 7 papers for the last one). All of the studies retrieved from relevant and valid sources (one search engine and five databases with a priority of selecting the papers from four of the most prominent journals and conferences in location based services and mobileHCI fields). Figure 1-2 shows the outline of the procedure of the systematic literature review.

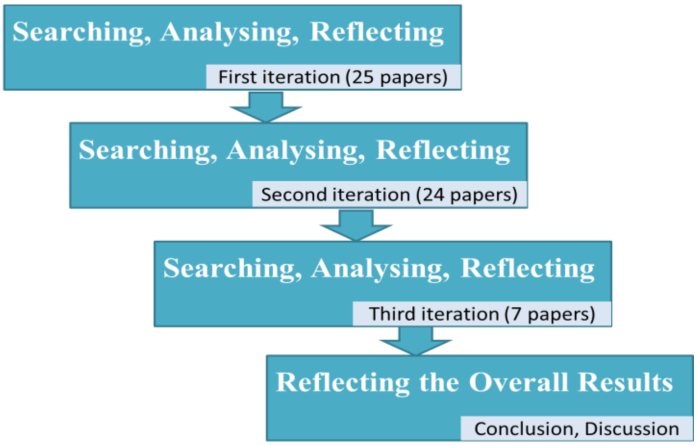

Figure 1-2: outline of the methodology

1.3 Thesis Outline

This thesis is organized as follow; In the next chapter (chapter 2) the process of the first iteration of the SLR with the achieving results and some initial conclusions will be presented. Chapter 3 discusses about the second iteration of the SLR with corresponding results and conclusions. Chapter 4 is about the last (third) round of the SLR iteration and the reflected results and conclusions of that. In chapter 5 we have discussion section and finally chapter 6 draws the overall conclusions.

Chapter 2: First iteration

The procedure of the our SLR in each iteration involves three main stages; searching, analysing and reflecting.

2.1 Searching

According to the research questions (RQs) that has formulated before which described in previous chapter and the knowledge that achieved through preliminary studies (with studying some SLR studies in usability evaluation of software/requirement engineering and the material of the two courses[8] that have passed in the University of Muenster at the previous semester and the previous experiences and educations of the author), some keywords with their corresponding synonyms have been extracted through reviewing the available works in our first round of the iteration (Table 2-1). Initially, some of the most common usability evaluation methods such as "*think aloud*" and "*SUS" (System Usability Scale) questionnaire* etc. also inputted to usability evaluation method keywords to achieve at least a few usability evaluation methods in the expected results[9].

Table 2-1: Keywords

Core concepts	Synonyms and related phrases
Usability	UX, user experience, user-centered design, usage-centered design, UCD, human-centered design, HCD, human computer interaction, HCI, Mobile HCI, mobile user interfaces, usability engineering
Usability defects	Usability issues, Usability problems, usability flaws, usability mistakes
Usability evaluation	Automated usability evaluation, Remote usability evaluation, Usability test, Usability testing, Automated usability test, Automated usability testing, Remote usability testing, usability inspection, usability heuristics, heuristic evaluation, usability inspection

[8] Location Based Services and Usage Centered Design courses
[9] In the next iterations we have not added any usability evaluation method for avoiding bias in our results.

Mobile GIS	Map-based Mobile applications, Mobile map applications, MMAs, Mobile maps, Mobile devices, Mobile phones, haptic systems, Location Based Services
Usability evaluation method	Automated usability evaluation method, Usability testing method, Automated usability testing method, Usability inspection method, Usability heuristics method, Heuristic evaluation method, Usability inspection method, User study, Field study, Elicitation study, Think aloud, NASA TLX, SUS, USE

According to these comprehensive keywords, the search strings have calibrated for executing the search in the first iteration. The selection of the search terms for the search was in a systematic way of excerption of some of the combinations that whenever we got a huge number of papers, we have tried to narrowing down the search terms to achieve more specific papers (fewer) that would be more relevant to mobile map-based systems (MMSs). Search string number one is the most probability state of search strings, which other search strings are the systematic excerpts of that (Appendix A). Putting this huge list of the search terms in the search engines and databases is not possible since most of them accepting a limit number of the search strings. We bring them as an example of combining the search strings with "AND" and "OR" conditions. But obviously every search engine or database has a specific search strategy that same search string is not necessarily working well everywhere.

2.1.1 Search priorities

Other search strings created and the search for the first iteration has conducted in the October of 2018 in the following 4 databases and search engine. The search engine and databases that used for searching were **Google scholar**, **Scopus**, **ACM digital library**, **dblp** and **Science direct**. In addition, some of the most relevant and prominent journals and conferences in domains of human-computer interaction (HCI), mobile GIS and location based services such as **Conference on Human Factors in Computing Systems (CHI)**, **International journal of Mobile HCI**, **Journal of Location Based**

Services, **conference on Human-computer Interaction with Mobile Devices and Services (MobileHCI)** have considered as a priority in executing the search.

2.1.2 Search Strings and their Results

Here, there are search strings and their corresponding results (Table 2-2). Because of the formula function possibility of the Microsoft excel that we used in the table, here we avoid plus signs at the beginning of some search strings, that the software might recognize them as a formula.

This is one of the random search strings sample that can be extracted from the main search string: "user-centered design" OR ucd OR "usability engineering" AND "usability issues" OR "usability flaws" AND "remote usability evaluation" AND "Map-based mobile applications" OR "mobile map applications" OR MMAs AND "Automated usability evaluation method" OR "Automated usability testing method"

Table 2-2: Search strings and their corresponding results – First iteration

Search string	Database	Filtered by	Result	Included in first screen	Included for data extraction
map +location based services +user study	dblp	(-)	1	1	1
map +location based services +user study	ACM	MobileHCI and 2008 to 2018	138	1	0
map +location based services +usability	ACM	CHI and 2008 to 2018	40	3	1
map AND location based services AND usability	Scopus	2008 to 2018	41	12	4
ux +map-based mobile applications +usability evaluation	Google scholar	2008 to 2018	5,100	18	6
ux OR user experience OR mobileHCI +usability issues OR usability problems OR usability defects +usability evaluation OR usability heuristics OR usability inspection +map-based mobile applications OR mobile map applications OR MMAs OR mobile maps +usability evaluation method OR user study OR elicitation study OR field study OR think aloud OR TLS OR SUS OR USE	ACM	CHI and MobileHCI and 2008 to 2018	371	22	1
usability issues OR usability problems OR usability defects +map-based mobile applications OR mobile map applications OR MMAs OR mobile maps OR map +user study OR elicitation study OR field study OR think aloud OR TLS OR SUS OR USE	Google scholar	2008 to 2018	2,270	39	12
Total number			7961	96	25

In the first round of the search 7961 papers found, between them, 96 papers with the first screening selected. The first initial screening procedure have done according to reading the title, keywords, abstract and in some cases the conclusion section. The procedure of the first screening was in this way that at first, if the title seemed relevant (with the experience that the author had in the field of GIS, subjectively if the title had a common point with mobile maps and was relevant to usability evaluation and LBS, proceeded to the next stage, otherwise excluded) the keywords was checking, if the keywords have terms such as Mobile, Map, LBS, user study, user experience or a kind of usability evaluation methods or terms, then the abstract section was checking and if the abstract seemed not completely relevant, the conclusion section was reading carefully to try to as possible as not missing any relevant paper. We also used Web Mendeley in order to manage the papers and citation.

2.1.3 Inclusion Criteria

For quality and eligibility assessment, we include papers for data extraction and analysis that:

1. Have at least one usability evaluation method
2. Evaluated a mobile map-based application
3. Within the recent eleven years (from 2008 to 2018)
4. Written in English

In order to answer the first RQ we defined the first inclusion criterion to qualify the included papers. The studies that had usability evaluation method for mobile devices, but without map or GIS aspects (there are a lot of studies that only evaluated the mobile-user interactions that are irrelevant to map and geo applications) and also the literatures with usability evaluation method in GI Systems or map-based desktop systems, which are not about mobile devices, have excluded from the review, since the usability evaluation conditions of mobile map-based applications are different than other above-mentioned domains. We also excluded grey literatures. For the third criterion, we decided to include the papers from last 11 years that before this period the mobile devices were different than today's tough based form and these years were coincide with the first iPhon's inauguration[10] that mobile phones have changed drastically.

In order to weed out the papers that do not have the specified inclusion criteria, in the second screening, the entire of each paper (full content) studied (reading stage) with considering inclusion criteria.

The number of 29 papers for the first data extraction and analysis have been selected. During the data extraction, we have found that four more papers should be excluded because of the lack of enough quality and repetitiveness. Therefore, finally 25 papers inputted in the first round of data extraction that 72% of them was achieved in Google Scholar search engine (advanced search), since this search engine covered some of the results of some of the defined databases too.

2.2 Analysing

The initial table with initial criteria (columns) for extracting the data for those 25 papers has created. The table includes 20 columns (criteria or field) and 25 rows (each paper is one record in the table) has showed in Appendix B. The papers in table have ordered according to the date that had published (descending). Figure 2-1 shows the time distribution of the included papers, which the lack of enough studies during the recent years is noticeable, but since there is not enough papers in the review so far, is too soon to deduce any conclusion. Most of the papers (24% of the papers) in the first round of the iteration between these 11 years, are from the year 2010, and surprisingly, in the year 2017 and 2018 (2 recent years), there is not any study to review.

[10] The first iPhone released on June 29, 2007

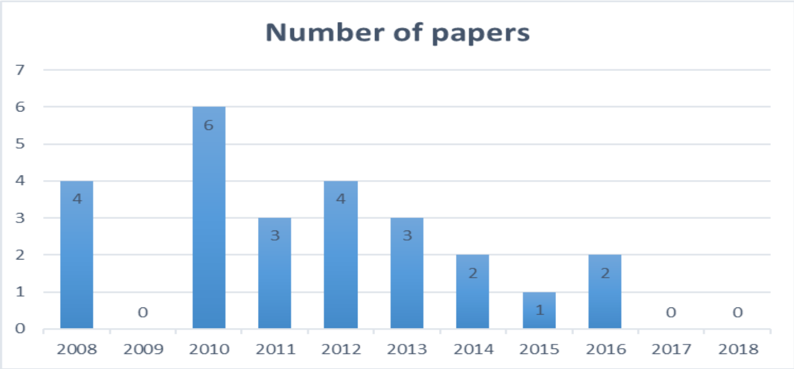

Figure 2-1: The time distribution of the included papers in first iteration

At the end of each analysing stage, the search strings should be modified (calibration) according to the knowledge that gain from the review, and the analysis must be updated for the next iterations to extract new desired data. After reflecting the findings, insights and results from the first analysis, in second round of the search, the first-round papers (25 papers for the first iteration) might need to be review again in order to extract new and more data and reflect the updated analysis (same for each round of iteration).

2.3 Reflecting the Results

In order to answer the first research question (RQ 1), the column "**Usability evaluation method**" in the table (Appendix B) has analysed. Totally 12 distinct method was recognized in the first round of the review. These methods ordered according to their higher frequency occurrences (descending order):

1- Questionnaire (all the different kinds) (31 times)

2- Synchronized video and audio recording (10 times)

3- Think aloud (8 times)

4- Interview (post-session, post-task or post-test) (7 times)

5- Logged data (screen logging and so on) (6 times)

6- Experimenter observations

7- Spatial memory task (least frequency)

8- User comments (least frequency)

9- Average traversal speed (least frequency)

10- Users' task performance (user workload) (least frequency)

11- Multicamera recording (least frequency)

12- User interface actions and accuracy (least frequency)

Between these 12 extracted usability evaluation methods, "*questionnaire*" was the most frequently used method in the first round of the review, followed by "*synchronized video and audio recording*" which was only 13%, and "*thinking aloud*" method with 10% of the all methods that were used in the available studies of the first round of iteration to detect usability issues in mobile map-based systems.

According to [12], for detection of critical and serious problems, the "*thinking aloud*" method is most effective, and then they came up with a conclusion that a combination of "*thinking aloud*" method and *video recording* with eyewear are most suitable for the evaluation of mobile devices in the field.

Usability evaluation methods, in term of the environment that they operate, can be categorized in two major groups; field-based or laboratory-based methods. For the domain of mobile map-based systems, this criterion is extremely important since the user's physical environment context in detecting the usability issues is playing a significant role when mobile users are not always sitting behind the desk like desktop users, and in interacting with map-based systems (e.g. way finding tasks) most of the time are in moving status. In the reviewed papers (first round of iteration), approximately both kinds of these two group of methods have used equally (53% field-based methods). Burghardt et al. [12] subdivided field-based usability methods to observation methods and survey methods. According to them, observation methods are *thinking aloud* and *audio and video recording*, while survey methods are *questionnaires* and *interviews*.

According to ISO 9241 standard, part 11, *context* in usability evaluation is related to users, tasks, equipment (hardware, software and materials), and the physical and social environments in which a product is used. Screen size, wireless data transfer, daylight exposure, touch interaction etc. are some equipment differences in mobile domain with desktop computer environment, but there are other user characteristic's such as their information needs and their age and computer literacy, that are also very important in consideration of design process of the geo-applications [13] and after implementation phase, especially mobile map-based applications, since more wider range of users using them[11].

Here according to extracted data in the first iteration table in appendix B, the size of the devices' screen can be derived from the "Tested device" column, and the age and computer literacy or relevant GI knowledge of participants (i.e. Geo students), ("TPs number with relevant knowledge" column) of the users in the studies has already identified and reflected in the table.

About the age groups of the participants in a user study, Burghardt et al. [12] believed that the senior test persons identify in an empiric usability analysis the most critical problems and can say the high-quality mistakes, when the middle age group is suitable for the refinement of a product and young group seems to be inappropriate for this kinds of evaluation, because they are difficult to recruit, they also tolerate a lot of errors and are often not self-confident enough to "blame" the tested device.

---

[11] Some of the commercially current available mobile map-based systems such as Google Maps still have some usability problems, that have not resolved yet.

Beul-Leusmann et al. [14] believed that their participants were young with high-technology-affinity, that might overlook the usability problems.

Here, opposite to previous thinking about the age, the most frequently usability issue (*losing the overview*) was detected by the age group 20 to 26 which is belong to young age group[12]. According to extracted data from the table in Appendix B, this usability issue was occurred among the age groups of 16 to 37, 17 to 27[13], 18 to 26, 18 to 60, 19 to 47, 20 to 32 and 20 to 59, which the pick range was the above-mentioned range (20 to 26).

64% of the participants that referred to this issue, had relevant knowledge (but not experts) and the issue in 83% of the user studies has found in the laboratory environment which is not a proper method to expose subjects to the real environment of real mobile users that most of the time they are not behind a desk like desktop users.

The most frequent usability issue (the problem of *losing the overview*) has detected in studies that operated with mobile phones with the average screen size of 3.62 inch, which in comparison to mobile phone screens' today, is too small. One reason for that can be, since most of the studies in the first round of the review operated on the year 2010, which the mobile technology has characteristically changed during these recent years that none of the papers were from these two recent years (2017 and 2018), the size of the screens that the issue has detected is too small. Another reason can be, the screen sizes of iPhone mobile phones until September 2014, were smaller than 4 inches (which we had only 3 papers to review after this time). It needs more investigation that the issue of *losing*

[12] According to UN, persons between 15 to 24 year consider as young.
[13] They reported only the college students had recruited for their user study, which according to Statistics Canada, Postsecondary Student Information System, over 75% of students were between 17 and 27 years of age in the college. (https://www150.statcan.gc.ca/n1/en/subjects?MM=1)

the overview of the map in mobile map-based systems is still remained with the new generation of mobile phones (with bigger screens) or not.

The issue of *losing the overview* has mainly detected by "*questionnaire*" method. Therefore, we can conclude that *questionnaire* could possibly be a suitable method to detect these kinds of reoccurring issues in mobile map-based systems.

Therefore, for the sub Research Question 2.b we have found the most frequent usability defects in mobile map-based application among our studies that reviewed, in the context that showed in the figure 2-2.

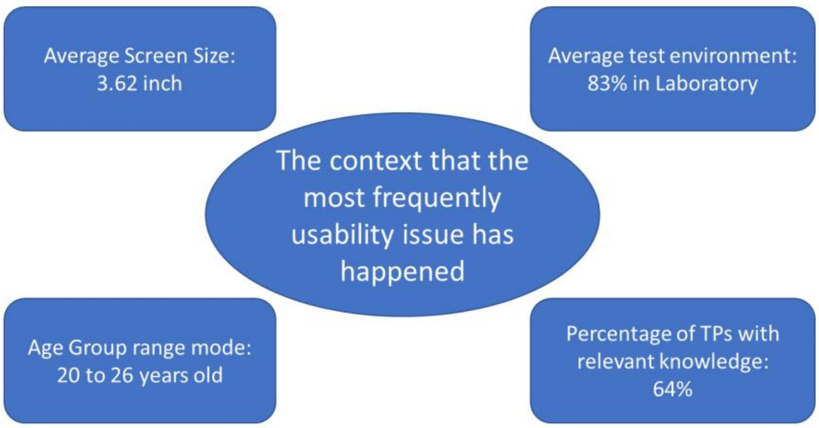

Figure 2-2: The context that the most frequently usability issue has happened in the first iteration

About the gender differences in user studies, Coluccia et al. [15] found that in wayfinding tasks, males generally outperform females. There is well stablished research about gender differences in spatial ability and navigation behaviour. Males and females employ different strategies in navigation and spatial orientation. For example, men use more directions and distances in navigation, females on the other hand use landmarks to orient themselves in navigation tasks [16]. In their study, they have found that male participants performed better than females. In addition, women relayed more on landmark knowledge than the overview that provided by map.

In the first round of the SLR, 58% of all the participants, in all the studies were male, but for the group of the participants that referred to the most frequently issue, the distribution on the gender was 50% for male and 50% for female. Therefore, we cannot deduce any conclusion about the gender differences in our first round of the iteration.

Since the number of the studies that has reviewed is not enough in the first round of the iteration, we cannot deduce any conclusion yet.

Among all the literature that have reviewed in the first round of the iteration only 36% of them evaluated the real mobile map services or online map services that are available in the market which the most well-known and the most famous one of them is Google Maps, and others evaluated only their prototypes or some software that they implemented for mobile users and tried to do usability evaluation for them. Actually, with studying the mobile map-based applications that are not completely designed or implemented also can notify some usability issues that are available in map interactions in mobile devices, but we tried to achieve the papers that evaluated the usability issues of the commercially current available mobile map-based systems that still have a lot of problems and not ease to use for so many people.

Flink et al. [6] categorized the usability issues in mobile map-based applications in an interesting way. They grouped the results of evaluation of the mobile map application into three main groups:

1. Hardware
2. Contents
3. User interface

The hardware group contains all the technical issues such as internet connectivity and the issues with the iPhone itself and downloading the maps.

The ‘contents’ includes various observations pertaining to visualizing the content of the maps, such as text and icons. E.g. Users argued that the usefulness of different background maps and additional features such as sound, video and photo landscapes on the mobile maps need to be clarified and geotagged.

The ‘User interface’ group is that the number of actions the user needs to perform a task should be minimized. Meaning that for example the map applications should have a search field or a function to choose from a list that this would be more usable and time saving for users. Our focus here, is around these two last categories.

For the second research question (RQ 2.b), first we can categorize the usability defects in mobile map-based systems in two major groups:

1- Spatial

2- Technological

The spatial issues are all the usability problems which are related to mobile map itself and mobile map interactions that make using of maps, difficult for the users, such as the algorithms that are behind the navigation services (such as routing algorithms), the functionalities that the different maps used, and the different kinds of map interactions and displays and so on.

The technological problems are all the issues relevant to the technology itself, such as sensor inaccuracies, battery drainage, internet connectivity etc. which is out of our focus in this work.

2.3.1 Spatial issues

Delikostidis et al. [61] in 2010 and 2016 [18] referred to lack of automatic rotation of the North-up map in Google Maps to the actual direction of the user. Wen et al. [17] also detected this problem in directional orientation with simple north-up map. Elzakker et al. [5] pointed out the inability of the mobile map to be oriented toward the actual view point of the user.

For the third sub research question (RQ 2.c), there are some solutions that proposed in the first round of reviewed literatures. Wen et al. [17] proposed a forward-up map, which shows the direction of the device during the navigation. Delikostidis et al. [18, 61] and Elzakker et al. [13] to overcome the problem of the user direction of the North-up map, proposed a rotating map and a compass-based heading-up (rotating) map.

Delikostidis et al. [61], detected icon overlapping in particular zoom levels. They proposed some methods such as: landmark pop-up information, multi-perceptive photos and landmark symbology for dealing with the above-mentioned issue and some other solutions such as; vertical scale bar with the combination of distance and time needed, landmark filtering and dual map to enhance the usability of the mobile map-based systems.

Rehrl et al. [18] found difficulties in readability of the name of the streets on the map (upside down), because they used standard OSM[14] tiles that were align to the north. For tackling this issue, they proposed different map tiles for the four cardinal directions that were rendered in their prototype. Ramsay et al. [19] also found significant delay

[14] Open StreetMap

during scrolling while new map tiles were downloaded from the remote server. In order to overcome this issue, they propose catching the tiles covering the area.

One of the most significant and well-known issues in mobile map-based systems, that arise from smallness of the screen size of mobile devices, and in our first round of systematic literature review (SLR) was also the most frequently issue[15] that happened in the user studies, is *losing the overview* of the area and orientation because of continues user's *zooming and panning*. When users zoom and pan continuously, it might be confusing to get the overall view of the region they are viewing. When they zoom in, they only can see a small area of the region, and when they try to have an overall view of the region, they usually doing zoom-out operation, but at the same time they can't distinguish some detail information [20] (such as the street names or other spatial details on the map that might be generalized or might be because of the smallness of the screen). That is why mobile map users doing a lot of zooming and panning that causes the losing overview problem.

Typically, when the spatial information is displayed in the mobile screens entirety, users obtain an overview without sufficient detail (e.g., they are unable to read the texts). By zooming in, users may obtain needed details but at the same time lost the overall view of the spatial information that display the area outside of the screen. If the essential point of interest (or any entity of the map) located in the off-screen area, users need further zooming and panning to see them. These extra struggles for visualizing the desire detail level make the experience of the user of mobile devices more difficult and time consuming and decrease user satisfaction of work with the map applications and also according to [21] hinders the creation of cognitive map of the explored spatial area. In

[15] Research question number 2

addition, since the mobile users most of the time are not in a stable situation because of mobility feature, it is very important for them to gain suitable spatial information awareness by glancing at the screen [22]. Burigat et al. [23], Delikostidis et al. [24], Burigat et al. [22], Polino et al. [20] Hooten et al. [25] Burigat et al. [26], Bouwer et al. [21] all referred to this issue in their works.

Dünser et al. [16] operated a user study to compare navigation task with augmented reality (AR) interface and a simple digital map and a combined map and AR condition. They found no overall difference in task completion time, but they found evidence that AR browsers are less useful in navigation at some environmental conditions. One of the usability issues that they detected was the losing of the overview in AR interface, which users didn't recognize the dead ends routs.

In terms of the available methods and solutions to deal with this issue, there are several approaches that one of the common and prominent ones provides users both overview and detail simultaneously, calls "*Overview&Detail*" approach. In this approach when users zoom to a specific level of detail on the map, one or multiple overviews of the space (usually with smaller scales) are representing in a small portion of the screen, around 10 percent of the full screen size (in a thumbnail) [26].

This method has proposed by Burigat et al. [23] as a solution to avoid users' extra zooming and panning when they lose the overview of the space during the navigation. Burigat et al. [26] in another study proposed ZEN (Zoom-Enhanced Navigator), which is an adaptation of *Overview&Detail* approaches to mobile devices. In this method panning and zooming is integrated in a same interaction and only an outline of the overview is showing to user, thus the screen space can be saved in comparison to other methods of *Overview&Detail* which a new smaller window (with usually smaller scale) occupying a part of the screen space.

Hooten et al. [25] suggested the "paper maps" when users *losing the overview* of the space. Paolino et al. [20] proposed a new method for off-screen visualization which called *Framy*. The off-screen visualization approach followed by CityLights concept, which proposed by Mackinlay et al in 2003 [20], that compact graphical representations such as points, lines or arcs which are located along the margins of the screen to hint off-screen objects located in their direction [26]. *Framy* is a kind of off-screen visualization method that uses a cornice semi-transparent shape that resembles the off-screen objects (POIs) according to their distance to the map focus with using colour intensity [20]. This method provides a situation for user to simultaneously with focusing of a subset of selected data, getting an insight on what is around too [20].

Burigat et al. [22] compared the effectiveness of tree off-screen visualization techniques (*Scaled arrows*, *Wedge*, and *Overview + Detail*) in their experimental evaluation. *Wedge* that proposed by Gustafson et al in 2008 is a visualization technique to convey the location of off-screen objects through triangles, that the base and partially two legs of the triangles are shown on the screen which that two legs point towards the off-screen object. Users should estimate the location of that off-screen objects (POIs) according to the direction and the size of those triangles. *Scaled Arrows* that proposed by Burigat et al in 2006, is another technique for visualization of off-screen objects with using the different size arrows that the larger the arrow, the closer to the screen is the off-screen object. Therefore, users can estimate the distance and direction of the off-screen objects when they are in a specific zoom level of the map. In *Overview&Detail* visualization, the overview of the space is showing as a small thumbnail that cover about 10% of the screen at the bottom right corner of the detail view. Figure 2-3 shows the tree visualization techniques that they considered in their study.

Figure 2-3: (a) Scaled Arrows, (b) Wedge, (c) Overview + Detail [22].

They found, totally, there is no single best solution to support users in carrying out different spatial tasks on mobile devices when relevant objects are off-screen, but in some of their tests results, the *Overview&Detail* technique showed superiority.

Delikostidis et al. [24] and Elzakker et al. [13] proposed *reversed Overview&Detail* that is a new approach which is opposite of the *Overview&Detail* approach in showing the overview and detail views, in order to reduce continues zooming and panning, which is presenting the detail view of the space in a smaller window (thumbnail) inside the overview map in the full screen (figure 2-4). In previous approach (*Overview&Detail),* we had the detail view of the region on full screen and the overall view on the small thumbnail, but in the *reversed Overview&Detail* approach, this is reverse.

There are not enough researches about the effects of the limitations of mobile devices on design and use of overview + detail visualizations. Most of the time the overview thumbnail is too small to users' eyes to recognize the spatial information [22] and since the overview map usually is in a small scale that add more difficulty to recognize spatial information.

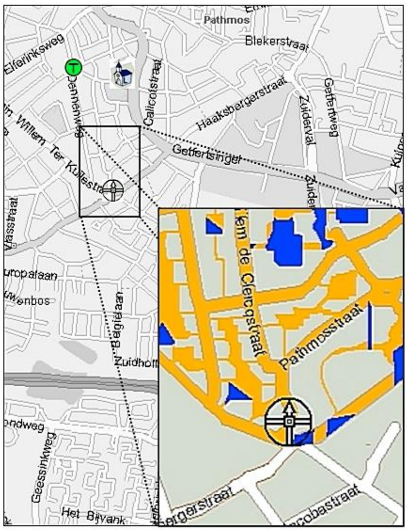

Figure 2-4: Reversed overview + detail mobile map [13].

Elzakker et al. [13] proposed two possible solution to overcome the problem of disorientation and spatial confusion of mobile users in navigation, when they have to zoom and pan a lot, because of smallness of the screen. Their solutions were to keep particular landmarks visible in all zoom levels and applying smooth zooming instead of discrete zooming.

Flink et al [6] found that interpretation of some background maps were difficult for the participants and they proposed legend to overcome this issue. This issue can be an important topic in map interfaces of mobile map-based systems (MMSs), where designers and producers should take it to account at the initial stages of the design procedure. And also, they have noticed with their study that the map interface doesn't have search function and to address it, they proposed search field or choosing from a list.

Another study in 2010 by Van Tonder et al. [27] proposed a new way of user interaction with map-based applications which calls "*Tilt interaction*". This technique has proposed to tackle the issue of big finger in mobile phone interactions that the display can be obscured by user's hand, specifically in map interactions. They compared two interaction methods, namely: tilt and keypad on a prototype mobile map-based application (MapExplorer) in a laboratory to measure the user performances of the three tasks (locating, navigating and checking tasks), and they have found that "*Tilt interaction*" only in navigation tasks performed better.

According to Nielsen (1989) and Virzi (1992), minimum required number of the test persons are between 5 to 6 that reveals the approximately 80 percent of usability problems [12]. In our first round of the literature review, more than a half of the user studies have conducted with 18 to 24 participants.

Thanachan et al. [2] did an interesting usability test to compare two map applications (NOSTRA map and Google map) with only 5 novice participants (around 21% of the user studies have operated by less than 10 test persons) in laboratory to measure 5 usability attributes of Nielsen and ISO 9241-11 such as *learnability*, *efficiency*, *effectiveness*, *memorability* and *satisfaction* through *video recording* and *Post-Study System Usability Questionnaire (PSSUQ)* methods. They have found a lot of user-software interaction usability problems encounter of NOSTRA Map application and Google Maps from a path analysis task with iPone 5 that most of the issues were related to the design of the icons and their location in both applications. The usability problems were: words used on interface were misinterpreted by users, users cannot find category and need to search, icon sub-category was not easily noticeable, cannot save the favourite places, function finding route was complicated, cannot open the list of favourite places, unable to show detail result page, didn't see current location button, cannot chose to Hybrid Map, get lost into Measurement Tools function, the overall problems founded were related to the design of the icons and their location in both apps which inappropriately presented. And they also proposed some solutions to dealing with those issues such as: redesign of icons and change their location on screen, words use in apps should be minimized and according to individual user, comprehension should be confirmed, the more menu should be integrated into main menu to reduce user confusions.

They have found Google Maps had better design in term of learnability (the time duration to work successfully for the first time) than NOSTRA application.

Beul-Leusmann et al. [14] investigated a usability evaluation of an intermodal passenger information system (a prototype) and tested in comparison with the leading mobility application in Germany (DB Navigator). They detected several usability issues by the subjective comments of the participants such as: Lack of auto-completing in the text fields and lack of overall view and small screen, lack of automatically selection of surrounding bus stops, lack of arrival/departure time, problem in deleting the texts in the text fields, lack of information about the overall progress during the trip, absence of properly placed landmarks, color code and some technical problems of the location-based service.

Burigat et al. [26] noticed occupying the screen with hand or stylus when users interact with mobile map application, and to solve this problem, proposed DoubleScrollBar technique. This method allows users to perform scrolling operation by using separate horizontal and vertical scrollbars. And provide users some predefined zoom levels to choose a specific zoom level directly.

Noguera et al. [28] in evaluating a prototype found that the map should have the possibility to switch from the 3D to the 2D interface and vice versa. And the map also can be integrated with social networks.

Delikostidis et al. [29] in their evaluation study found stacking in the previous position in Google Maps when user has moved to a new position and they also noticed some big landmarks in reality that didn’t represent in Google Maps.

Elzakker et al. [13] pointed to simplicity of the map that should not be overloaded with many symbols or 3D landmarks. They have found that landmark photos that pop-up when clicking on them, were more preferable. And also, by presenting landmarks in successive scales, the frequent zooming and panning can be avoided by users. They deduce that the spatial information on the map should be represented in a way that users spend more time to observe surrounding to develop mental maps than looking at the mobile map. They pointed out that more choices for pedestrians should be provided to freely select any possible routes (flow channels) to the destination. They proposed automatic landmark recognition with using integrated digital camera with GPS position and heading information with landmark visibility map data on image recognition algorithm.

Kuparinen et al. [30] measured the suitability of a domain specific heuristic evaluation (HE) for mobile map applications (MMAs) in comparison to Nielsen's Heuristic, and found that more usability problems were found with the proposed HE for MMA. They found that the applicability of Nielsen's heuristics (1994) are not only too general, but also limited to be applicable for evaluating MMAs [30].

In the first round of iteration of our SLR, only 8% of the studies hired experts to operate heuristic evaluation (HE) in order to evaluate the usability of mobile map-based systems. The main difference between heuristic evaluation and empirical user testing is, with HE, identifying the errors is on the centre of the focus, since user testing is determined by *effectiveness*, *efficiency* and *user satisfaction* [30].

These three components (*effectiveness, efficiency* and *satisfaction*) that mentioned in the definition of usability by ISO 9241-11 in 1998, were the basis of most of the reviewed literatures in our first iteration to evaluate usability of mobile map-based systems. A lot of them, [4, 5, 7, 9, 64, 13, 14, 20, 21, 23, 25] measured the *task completion* *time* in their user studies to evaluate *effectiveness, efficiency* or *overall efficiency* and *effectiveness* of the mobile map-based systems (MMSs). Some of them, to evaluate *satisfaction* and *ease of use*, measured *error rates* in their empirical user studies [13, 25]. Other studies to evaluate *satisfaction* and *ease of use*, operated *questionnaire* and *interview* [6, 7, 9, 13, 14, 15, 16, 18, 19, 20, 22, 23, 25, 26, 27].

Thanachan et al. [2] measured The time duration to work successfully for the first time to evaluate *learnability*, and the time duration to work successfully after avoid using system for 5 days in order to evaluate *memorability* (rememberability) of the evaluated the usability of two mobile map-based services in their comparative user study. Kratz et al. [31] used *USE questionnaire* in order to evaluate perceived *learnability, ease, satisfaction* and *usefulness* of users.

Technology acceptance model (TAM) introduced by Davis and his colleagues at 1989, "is widely regarded as the most successful model for explaining how people form their opinions, use and accept particular services or technologies" [32].

Park et al. [32] in addition to use two psychological beliefs of TAM (perceived usefulness and ease-of-use) in their study, also have measured five usability variables such as perceived location accuracy (PLA), satisfaction, service and display quality (SDQ), perceived mobility and flow state[16] by *in-depth interview* (they believed these factors may significantly contribute in users' intention to use) with two groups of individuals: a user group (users of mobile map services) and a professional expert group (developers, engineers, and designers in the field of mobile map services) and conducted an online survey (questionnaire) about three mobile application sites and three mobile

[16] "Flow state" is a mental state of user, when he/she is fully immersed in something that he/she is doing, specified by an energetic focus and full involved manner that not enough aware of his/her surroundings and success in the operation of the task [27].

services sites by 1109 respondents who had at least 6 months of experience with mobile map services. Their study contributed in deeper and more comprehensive insight about users' behaviour toward using mobile map services.

They have found that service and display quality (SDQ) and perceived location accuracy (PLA) are the notable determinants of attitude toward mobile map services acceptance and also SDQ had a more powerful effect on attitude than PLA did, showed that users are more affected by factors related to the user interface than by technical factors.

Service and display quality (SDQ), is defined as "the degree of general performance of an information system and related services" [32].

Perceived location accuracy (PLA), which explained by Park et al. [32] is the degree of awareness of mobile map services' users of their exact locations on the displayed maps on their screen.

Perceived mobility (PM) is the degree of user's awareness, satisfaction and perceived usefulness of the portability of the services and systems. The questions such as; it is convenient to access mobile map services anywhere at any time [27].

2.3.2 Technological issues

Since the aim of our work in this study is not investigating the impact of issues that resulting from hardware limitations (from the mobile devices and servers to the positioning systems), these kinds of technological problems are not in the centre of the focus.

Delikostidis et al. [66] and Dünser et al. [16] referred to GPS and compass inaccuracy and Rehrl et al. [18] pointed to sensor inaccuracies in augmented reality (AR). Park et al. [32] found some problems such as location accuracy, processing speed, display and service quality. Some other studies [16] talked about some of the other issues that are relevant to the technology such as; screen brightness, shakiness and compass input, but in this study the usability issues that are related to the map is in the centre of our focus and technological part of this story is only at the marginal section of the analysis.

2.4 Conclusion on the first iteration

In the first round of the iteration, the problem of *losing the overview* of the spatial information on the small screens of mobile phone devices in map-based systems, were the most reoccurring usability issue between reviewed papers. There are some detected gaps in the reviewed studies in the first round of the iteration about some usability issues such as lack of update in maps and information available in mobile map-based services and the problem of internet connectivity (for example without the internet connection, navigation task in online map services such as Google Maps is not possible) that have not pointed out yet. Actually, there is no recent studies in our reviewed literature that investigated the usability issues in user-mobile map interaction that stem from individual user experiences in the real scenarios that might happen in different contexts of use.

And also, in our first round of review, only 8% of the usability studies had operated with "*experts*" that calls heuristic evaluation (HE) or usability inspection, and also, only 36% of the studies evaluated the real, available and current online map services such as Google Maps and other widespread services.

According to Park et al. [32] Over 150 million users have activated Google Maps (until 2011) on their mobile phones, despite this wide spread use of such mobile map services, there is a little search that has focused on users' acceptance and behaviour regarding these kinds of services. There might be three possible reasons for this shortage; first, since Google Maps is available by majority of companies in mobile market [27], on the most types of the mobile phones as a default application, and even most of iPhone users use to use that instead of Apple Maps (which is the default map system of iOS), there is not any serious competitor for it to push them to try to evaluate and make it better than the others. Second, the rapid changes in the mobile technology, might be a reason that there is not a lot of researches around the ease of use and usability studies of mobile map-based systems in terms of user interface that would be lag behind the new changes of the device properties, and the last but not the least possible reason might be the searching strategy should be in a different way than something we have done here, to lead us to the studies that evaluated Google Maps or other commercially current available services.

There are some intended usability issues about the current commercial mobile map services that there is not enough publication of the reviewed studies which properly referred to them. One of the important problems of usability of such services is internet dependency, means while user is connected to the internet can use the service, otherwise he/she cannot navigate to a destination or doing other map-based tasks. Imagine when user is in a place with a Wi-Fi internet connection, and in order to navigate to another place, use an online navigation service, while he/she is moving to the destination and becoming far from the source of the Wi-Fi signals, the navigation system stops to work (i.e. in the context of tourist users, when they want to go out of their hotel without any tour leader to see their surroundings or buy something, they lost the Wi-Fi internet connection while their mobile data connection only works in their source country, or if they want to use their own internet data from their original country they should pay extra money for roaming, that might be too expensive, therefore they cannot use the mobile map-based services outdoor easily, it can be a reason that today despite the availability of mobile phones, most of the tourists still are using paper maps). It could be much more usable when users don’t have access to network data connection on their mobile phones and especially in some emergency situations, still be able to use the system to do the navigation or other related spatial tasks. Sometimes the spatial information is too much for some mobile devices with low hardware configurations and the overload of the data make the navigation task too slow and tedious, especially in some cases the user is in hurry (which most of the times users use navigation services when they are in rush). Therefore, the mobile map services in order to be more useful to everybody, need to be more customized according to users' context, not only the technological context but also in social context such as; the age and the level of technological affinity and literacy and local languages and also according to the internet speed of some countries that is not similar to the developed countries. In addition, the map interaction should be very easy, without complexity, that with minimum hints, navigates user to the destination in a navigation task. Because on one hand, users of mobile map services most of the time are in an unstable position to be able to pay full attention their mobile devices. on the other hand, the system shouldn't make them completely flow in the virtual space, that distract them from the reality which might be dangerous for the pedestrians, cyclists or drivers from some unexpected things that might happen in their surrounding environments (although, some new user interactions such as GazeNav have introduced in mobile navigation system that more or less tackled the issue of distracting the user from his environment by mobile devices, but not ubiquitous until now, like traditional turn-by-turn pedestrian navigation systems). Therefore, it is not strange when today, at the age of information and communications technology (ICT) and ubiquity of mobile phones, still there are some people instead of using their mobile phones, using paper maps to find some places on the map.

Other problems of the mobile map-based systems which none of the reviewed literatures referred to that, are lack of up-to-datedness of the information, traffic information or for example the topology of the area (i.e. when Google Maps consider 8 minutes by walking from point A to point B, doesn't consider for example the slope or other topographic properties of the region between A and B).

Reviewed literature pointed to a general users' behavior that when they are interacting with mobile maps, most of the time users going to a specific zoom level and start to do "*panning*" operator to see off screen objects on the map that has represented on their mobile phones. One possible solution that comes to mind is the system according to the distance that user queried to see on the map (for example the route length between destination and current location of user in navigation tasks) can provide an abstract map[17] with a suitable scale, that make user independent of zooming and panning operations.

Dünser et al. [16] found that users preferred the combined map+AR condition and felt that there would be a significant problem with using the AR view alone for navigation. A simple 2D map interaction can be combined with some augmented reality (AR) interfaces, especially in the initial orientation of the users (or when they arrive at the destinations) which according to Elzakker et al. [5] and other studies, initial misunderstanding of users' location was a frequently reoccurring problem in the some user studies' navigational tasks (orientation), and also using landmarks in combination with such maps as Beul-Leusmann et al. [14] referred to the highly importance of them (In pedestrian navigation, landmarks are the most valuable navigation cues and they might be more important than street's names or distance information) can be useful when they could be appear in each zoom level to give the user a better sense of overview and the orientation of the area.

[17] With a perfect generalization of the information with taking advantage of using some salient landmarks that easily could be seen in a decision points (such as intersections).

There are other alternatives instead of usual mobile map interaction for navigation tasks such as auditory navigation systems, that are not actually an adequate solution since for example the noise of the urban environment around users might be disturbing [33] or users might feel alienate from their surroundings.

In some user contexts the lack of compass or accelerometer on some types of mobile devices makes the user orientation difficult, in such cases, the system should be intelligent enough to use some techniques to help them to perform spatial tasks easier and more successful (for example with using some prominent and salient landmarks and/or a North-up map can orient users toward the right direction of their destination in reality).

The mobile map-based systems shouldn't be intrusive and provide users too much information (sometimes with advertisement) that bother them. Imagine when a user is going to an important job interview and he is late. The navigation system reminding him several times that he is late or is in the wrong path. This too much intrusive information sometimes is annoying and making the user anxious and stressful. Or in some cases that users have more free time such as tourist case, users most of the time like to wander and a little bit even stray in the environment rather than only strictly follow the optimal shortest path since they have more time to enjoy the environment. The warnings should be simple and let users to pay more attention to the environment surrounding them and enjoy their visits rather than strictly alarming them going to a certain path.

Chapter 3: Second iteration

We need more iterations to find out that the outcomes of the next round of the review, reinforce the results of the first one or give us new knowledge about usability issues and their corresponding proposed solutions in mobile map-based systems.

3.1 Searching

In the new iteration, we need to refine the search strategy. First, we should narrow down the search to pedestrian navigation systems, since first of all car drivers usually don't use their personal mobile phones in order to navigate (today, most of the cars have their own navigation systems with completely different interaction than mobile phone interaction), second, there are a lot of differences between the navigation needs of pedestrians and drivers' [34] that make evaluation of these two systems different such as; in car navigation systems, drivers are limited to some specific routes (e.g. the kind of the way, since cars are limit to go to any ways like one way or two ways streets or streets with steps), since pedestrians have more choices to select their optimized route (e.g. parks, pedestrian malls, grasslands etc.) to their destination and also they would be more lost in terms of orientation and need the map information usually in larger scales with more details [24]. Ohm et al. [54] stated in contrast to navigation mode in car navigation systems, pedestrians prefer using landmarks in their route orientation. Günther et al. [35] pointed out some differences between vehicle and pedestrian navigation. They categorized these differences in data availability, degrees of freedom (which we referred here, such as pedestrians can go indoor and outdoor and to most kinds of streets with less limitations), hardware, positioning, interaction, human focus, navigation instructions.

We also excluded indoor systems from our review not only for the reason that there are enough publications that dedicated to evaluate map interface of outdoor systems, but also there are several differences between those two systems that makes the usability evaluation of those two systems different such as; the scale and dimension of indoor navigation systems are much more smaller than usual outdoor pedestrian navigation systems in mobile map based systems [36], the usual mobile map-based systems using GPS or other positioning satellite based technologies in order to navigate, that indoor navigation systems only use Wi-Fi or RFID technologies to navigate [31], most of the time Indoor navigation systems deal with multi-layer areas such as the floors of a building, that verbal directions like “go up to the 6th floor” are used, that never happen in the mobile map-based systems for outdoor navigation tasks [31].

The studies that evaluated merely augmented reality (AR) also excluded from the review because the interface of such systems is different than map-based systems and there are a huge number of studies in this field that surveying them is out of the scoop of our work.

There are some keywords that extracted from the first round of the iteration that can be added to our keywords such as; internet map services, mobile passenger information systems, pedestrian navigation systems, online mobile map services, mobile map services, technology acceptance model (TAM), technology acceptance concept and mobile map-based tasking interface.

There are some lessons learned from the first iteration searches that can help us to enhance the search skills for the next iterations.

3.1.1 Search priorities

New keywords have revised and shown in table 3-1 according to some knowledge that extracted from the first review. Some of the keywords have eliminated and replaced with new ones. Some excerptions of them systematically have used in search terms that used in the searching stage in those databases and search engine (with using one more database; Taylor & Francis database) that have used in the first round of the iteration with priority of those 4 most prominent outlets in this field (same with the first round of the review).

For this iteration we have updated our 4 inclusion criteria with adding one more criterion. The included paper should have at least one usability issue or one solution for a usability problem of mobile map-based systems.

We a little bit incline the focus of our review from usability studies before and during the design process of map-based applications (that concluded 64% of the reviewed literatures in the first iteration) and services, to outdoor pedestrian navigation system, which is one of the common location based services and most common spatial tasks that users are engaging with it in the most of mobile map services (e.g. Google Maps) and also after-design usability evaluation of available apps and services. The usability issues of such available services and applications are more deserving to review because their usability issues also can be the usability issues of each prototype or self-implemented application that performed by researchers or designers and they have also exposed to real users that according to Elzakker et al. [64] existing mobile navigation systems available on the market do not meet the user requirements in a suitable way.

Table 3-1: keywords for the second iteration

Core concepts	Synonyms and related phrases
Usability	UX, user experience, user-centered design, UCD, Mobile HCI, mobile user interfaces
Usability defects	Usability issues, Usability problems, Usability flaws
Usability evaluation	Remote usability evaluation, Usability test, Usability inspection, Usability heuristics, Heuristic evaluation, Usability inspection, Usability engineering
Mobile GIS	Mobile Map-based applications, Mobile Map-based systems, Mobile map applications, MMAs, Mobile maps, Location Based Services, Mobile internet map services, Mobile passenger information systems, Pedestrian navigation systems, Online mobile map services, Mobile map services, Mobile map-based tasking interface
Usability evaluation method	Automated usability evaluation method, Usability testing method, Automated usability testing method, Usability inspection method, Usability heuristics method, Heuristic evaluation method, Usability inspection method, User study, Field study, Elicitation study, Technology acceptance model, TAM, Technology acceptance concept

This search string (+map +usability +mobile "location based services" OR evaluation "pedestrian navigation systems") that operated in Google Scholar search engine, linked us to a bunch of good resources that within 316 results, in first screening, 60 papers have chosen. The probable reason might be the depend of the phrase “pedestrian navigation system” which is a kind of usual and common location based services (LBS) that wide range of users dealing with it. The search strings and their correspondence results with the search engines and databases have shown in table 2-2 (because of the limitation of the Microsoft excel in showing the plus sign in the first character of the cells that the software considered it as a formula, the plus signs are omitted in the table, but in front of each search strings in searching, there were a plus sign).

3.1.2 Search Strings and their Results

The searching section of the second round of the iteration has operated in December 2018. Between 527 results, in first screening, 86 papers have selected according to their title, abstract and keywords and in some cases the conclusion section that were comparing with our inclusion criteria that we had from the first round of the iteration with one more criterion that the paper should refer to some usability issues or solutions (table 3-2).

Some of the search results have excluded because commonly didn't refer to any mobile map-based usability issues and solutions. Some of them only achieved some subjective comparisons between two systems. In most of the cases they only evaluated some haptic or auditory interactions that were not relevant to mobile map-based interaction. One paper was repetitive in the first iteration. We excluded them from our analysis.

Table 3-2: Search strings and their corresponding results - second iteration

Search string	Database	Filtered by	Result	Included in first screen	Included for data extraction
"mobile map-based applications" +"user study"	Google scholar	2008 to 2018	14	4	3
"mobile map applications" OR "mobile map services" AND "usability test"	Taylor & Francis	2008 to 2018	5	2	1
"mobile maps" +"user study"	dblp	(-)	2	2	1
"mobile maps" +"user study"	ACM	MobileHCI and 2008 to 2018	162	2	0
"mobile maps" +"user study"	Scopus	2008 to 2018	12	6	4
"mobile maps" +"user study"	Science direct	2008 to 2018	11	8	3
"technology acceptance model" +"mobile map-based systems" or "mobile map applications" or pedestrian navigation system"	Science direct	2008 to 2018	3	1	1
map +usability +mobile +"location based services" OR evaluation +"pedestrian navigation systems"	Google scholar	2008 to 2018	316	60	11
map +usability +mobile +"location based services" OR evaluation +"pedestrian navigation systems"	Scopus	2009 to 2018	2	1	0
Total number			527	86	24

Finally, we have included 24 publications for data extraction in second round of the iteration that figure 3-1, shows their distribution during the 11-year period that has considered. In comparison to the first iteration, there are more publications in the recent years in our resources to review.

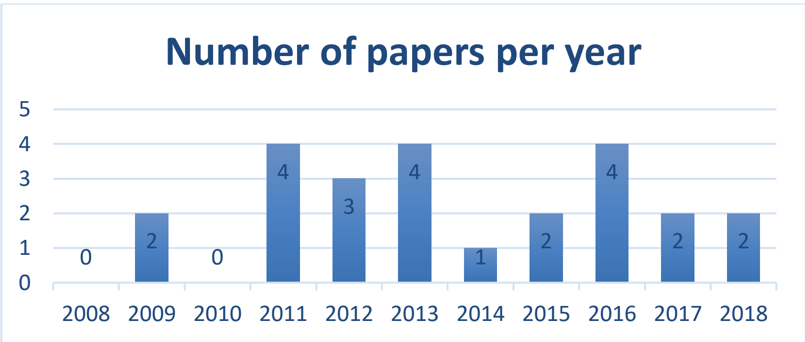

Figure 3-1: The time distribution of the second iteration publications.

3.2 Analysing

We have updated the table of data extraction (Appendix C) for the second round of the iteration. Some of the criteria have omitted from the table such as, the type of the studies that in the first iteration, most of the studies were comparative and the statistical methods that have used in analysing the results of the evaluations. Therefore, in the second table we have 18 columns (criteria) with 24 rows for the number of the papers that reviewed. In the analysing stage, we have extracted the data from 24 included papers through entirely reading the papers and filling in the table.

3.3 Reflecting the Results

For answering the first research question (RQ 1) which is the usability evaluation method that most frequently used for detecting the usability issues in mobile map-based systems, the column "usability evaluation method(s)" has been analysed.

The results of the analysis of the usability evaluation method criterion were interesting in some points. First, in all of the studies only 9 distinct methods were used, which in comparison to the first iteration (12 methods) less variety of methods were implemented for the evaluation. Second, the result strongly reinforced the outcomes of the first iteration. "Questionnaire" method was using in most of the reviewed papers of the second iteration (88% of all the papers). Different kinds of questionnaire have used to detect usability issues and evaluated the experiments in those studies in different aspects. Methods such as: user experience questionnaire (UEQ) (which measures 6 categories such as; attractiveness, perspicuity, efficiency, dependability, stimulation and novelty), usefulness, satisfaction, and ease of use (USE), NASA TLX (which is an index that measures user's performance, frustration, effort and mental, physical and temporal task load), users' sense of direction (SoD) (which Santa Barbara is a standardized scale to measure it), system usability scale (SUS) (which consist of three parameters; effectiveness, efficiency and user satisfaction), and Attrak Diff (pragmatic quality, hedonic-identification, hedonic stimulation, and some other parameters such as, complicated or simple, impractical or practical).

Here is the list of all the usability evaluation methods which used in the reviewed studies in the second iteration that ordered in term of their frequency using (descending):

1. Questionnaire (28 times)
2. Logged data – (9 times)
3. Interview (semi-structure and/or subjective feedback) – (8 times)
4. Experiment observations (6 times)
5. Think aloud – (5 times)
6. Video and/or audio recording – (5 times)
7. Eye tracking
8. Wizard of Oz

All the methods were used in the first-round papers too, except the last two ones that were new methods used in this iteration.

Wizard of Oz [37] is an experiment methodology to track the participants' location on the virtual environment and record some other data about the users' performance during the experiment in laboratory. In their user study [37], they provided users a mobile phone with an allocentric view of the user's location to navigate on the map and the reality was simulated with an egocentric presentation of the environment on a big screen as a virtual environment.

Logged data and *interview* were two methods that after *questionnaire* were used most frequently in the reviewed experimental studies of this iteration.

60% of the user studies conducted in the field, in this round of the review and other researchers evaluated user experience in laboratory environment, which in comparison to the first iteration (53%), there are more cases that evaluated outside. Despite a unique nature of mobile Geo-applications which user interact simultaneously with mobile and environment, 80% of the usability studies in this realm was actually executed in laboratory environment [66]. One possible reason for having less field studies than laboratory evaluations might be the higher time, effort and cost that need for operating user study in the field, otherwise field study is more similar with the real usage context of mobile devices which users are mobile that the weather condition, sun light or rain, the moving status, surrounding people or obstacles or traffic situations, egocentric view in the city environment between tall monuments, sensor inaccuracies etc. make the orientation and navigation more confusing in comparison to sitting in a quiet laboratory and according to Gkonos et al. [38] the performance of participants might be affected by these differences.

To answer the research question number 2 (RQ2) the *usability issues* column has analysed. The most frequent usability issue in this iteration was "*zooming and panning operations*" that 36% of the papers refer to this issue as a main problem in their user studies.

Zooming is a necessary and unavoidable operation in mobile map-based interactions [39]. The high need for zooming in mobile map-based interactions arises from the smallness of the screen of the mobile devices, which users induce to do a lot of zooming and panning to see the overview or more detail in map. On one hand, when they want to have an overall view of the map, they can't see some detail information such as the names or landmarks to have a better spatial pattern of the region they are looking for, in their mind, on the other hand, whenever they want to have more detail view or the off-screen objects, they need to do a lot of pan operation, since the screen is too small, and at the same time, they lose the overall view of the map. Here, in the reviewed studies (second iteration), these too many *zooming and panning operations* has noticed as a successive reflected usability issue of interaction with mobile devices in map-based services or applications and so many solutions so far have proposed to address this issue.

One of the recent solutions, that proposed at the University of electronic science and technology in China in 2018 [40] , tried to help users to deal with the problem of *touching* and *occlusion* of their interaction with such touch based devices in zooming and panning (in these two operations user need to tap-n-drag and pinch-to-zoom with direct touch of the screen [41]) based on camera with a contact-free or occlusion-free operation (CaMap, camera-based map manipulation prototype). Their method was accepted by their participants in terms of *ease of use* and *intuitive* that might be useful for some contexts such as rainy or cold winter days which touching the screen seems too difficult for doing a lot of zooming and panning operations that mobile maps need.

Konkol et al. [42] tried to solve this issue with dividing the users' attention between the mobile maps and the environment with proposing a new method that can support navigation task with assistant of the available signages and landmarks in the environment surrounding the user. Such methods might be useful that decrease too much engagement of user with mobile and also might decrease the need for too much zooming and panning.

Another interesting paper from Graz university of Austria [41], has proposed a new method that inspired from a project that the municipality of Schladming in the Austrian Alps have implemented for interacting tourists with the map of the area with engaging mobile map and a big screen that at the same time user can have the overview and detail which might help them to avoid doing too much zooming and panning operation (more detail is provided by [41]). They evaluated their method with a post hoc interview evaluation method and found that participants reported not only less *mental and physical demands* or frustration but also better *overview* for interacting with *magic lens* than usual mobile map interaction.

In year 2013, in Institute of Cartography and Geoinformation in Switzerland, Giannopoulos et al. [43] introduced a new method called *GeoGazemarks* based on a kind of generalization of the spatial information on the points that user seen before, on the mobile map during earlier interaction. They have evaluated their method in comparison with a state-of-the art mobile map interaction in terms of reducing zooming and panning or both of these interactions (with focus on people with lower spatial abilities) and noticed that not only panning operation has reduced significantly, but also their method supported users with low spatial abilities more than users with high spatial abilities.

Van Tonder et al. [44] proposed *tilt zooming* technique and found tilt interaction is particularly well-suited to mobile map-based applications. Their work was focusing on panning speed and engaging both hands of users, that can be adjusted through user's current context (walking or seated). They have compared gesture (usual touch-based zooming operation) and tilt zooming techniques with a specific user evaluation that was conducting in a 15-meter-long indoor corridor with a mixture of natural and artificial light and their analysis of log data showed that tilt zooming was more *efficient* than gesture zooming with less *perceived workload*, but *perceived workload* and *user satisfaction* ratings showed participants found gesture zooming to be *easier to use* while walking.

They compared their method with considering accelerometer-only engagement with sensor fusion tilting in another study [45] and the results of their evaluation showed that sensor fusion can be efficiently incorporated into a tilt interaction technique in mobile map-based applications.

Again, Van Tonder et al. [46] proposed a new method called IntelliTilt that addressed the shortcomings of tilt interaction and conducted a user study to compare it with basic tilt interaction incorporating SDAZ. The results of the evaluation showed better *perceived workload* (mental demand, physical demand, temporal demand, performance, effort and frustration) and higher *user satisfaction* for the IntelliTilt approach than basic tilt approach.

Kratz et al. [31] in 2010, proposed a novel approach based on semi-automatic zooming (SAZ) for manually control of zoom level in Speed-Dependent Automatic Zooming (SDAZ) approach that dealing with some problems of simple zoom interface such as; occlusion, slowness and sticky fingers problems. Their method (SAZ-based interface) contributed in quick zooming and one-hand mobile map interaction.

Cheung et al. [39] in 2009, tried to reduce the number of zoom operation by introducing *content zooming* concept which is an analogous to textual address (that in western countries is mentioning from more detail, e.g. house numbers to less detail, e.g. country names, and in eastern countries such as Iran and China is inverse). They have evaluated their technique with 20 participants and found that their approach can greatly reduce the number of zoom levels required and also is very *effective* for production of mobile-based mapping products.

The second usability problem that occurred more frequent than other problems between reviewed papers was the *complexity of the map* because of overloading with too much information.

Aditya et al. [47] have operated a user study with 18 participants that reflected the real scenario of navigation in the field with two different map interfaces; 3D map and Google Maps. Their 3D map visualization actually was level of details 1 (LOD 1), which is a simplest primitive building representation that is 2.5 dimension rather than 3D. They mentioned for mobile maps this level is enough in the term of occlusion of the map with a lot of information that makes map display complex.

They have found that selection and display of map using 3D map is highly better than simple 2D Google Maps. And also in regard to use of 3D map to support self-orientation, the responses of their user study were positive in navigation task. Even more for the case of spatial knowledge development and navigation decision, the 3D map provided *effective* and *efficient* means to accelerate test participants to go approaching the destination.

Ohm et al. [48] conducted two user studies (indoor and outdoor) to analyse different presentations (abstract design and standard map-like interface) of mobile maps. They have found that presentation modes of pedestrian navigation systems should be adaptive to users' sense of direction (SoD), which in their findings, badly oriented users tended to prefer standard map-like interfaces and well-oriented users seemed to prefer abstract designs. They also claimed that the ability to localize oneself (self-localisation) in environment may be affected by aging.

Dong et al. [49] operated a user study that was simulated in the laboratory to compare a simple 2D mobile map with a 3D photorealistic mobile map in terms of *cognitive workload*, *effectiveness* and *efficiency*. They have noticed that in map reading task, users spent more time in dealing with 3D map and also the 3D representation requires more mental effort than 2D map. But in decision making tasks, the 3D users performed better than 2D users. The method they suggested for the available usability problems in 2D and 3D mobile maps were, a combination of both for map representation. They recommended in 3D representations only the most important information should show to decrease the information density, and 2D maps, important landmarks should be included to help users locate and orient themselves.

Wither et al. [50] compared a traditional map-based navigation with panorama-based navigation. They found panorama-based navigation was more complicate to users and needs more attention of users in comparison to traditional maps. Another problem was discrete routes in panorama interfaces opposite to map interfaces that shows the entire route. They proposed switching between two interface modes combining both modes, which in City Scene (evaluated navigation application), when an overview of the entire route is required, users can switch back to map mode, although this switching by itself can cause extra workload for users.

Elzakker et al. [64] for addressing the problem of map complexity suggested that map should be simple, not overloading with many symbols or 3D buildings and must follow colour coding in a way that the size and patterns of the streets that represented on the map should properly reflect these parameters of reality [51]. They claimed that landmark photos that pop up when clicking on them are more preferable than 3D models. They emphasis that landmarks should be visible in successive scales using an algorithm to calculate landmark visibilities in any point of users' possible route on the map.

The third most frequently occurred usability issue in mobile maps between the second round of reviewed papers, was the engagement of users with mobile that can distract them from the real environment surrounding them. The mobile map should convey spatial information to users in a simple way that let them pay more attention to the real environment with minimum need to interact with the map and less cognitive (mental and physical) workload of interaction with the device. The system's warning should be simpler to let the users pay more attention to the environment and enjoy their visit [52].

There are several approaches to address this issue. Researchers tried to take advantage of vibration, sound (audio) and gaze to assist users not looking too much at the maps in navigation tasks. But these methods, by themselves have some problems, for example gaze-based approaches need more facilities such as glasses and audio-based navigation needs quiet environment that within noisy urban environment is impossible and if user want to use headset plug-ins, again he/she would be alienated from the real surrounding environment.

Gkonos et al. [38] introduced a novel pedestrian navigation approach called 'VibroGaze', a combination of a vibrotactile and a gaze-based approach and evaluated it with comparison a popular map-based turn-by-turn navigation, a vibrotactile approach and the gaze-based approach called "GazeNav" in indoor and outdoor environment. At the end they have found that their participants performed better in terms of completion time (*efficiency*) and the number of errors (*effectiveness*) through interacting with map-based navigation system. They claimed that familiarity of participants with map-based navigation, might cause bias in the results of their evaluation.

Konkol et al. [42] tried to incline users' attention to real environment (integrating real and virtual world) in navigation with using available landmarks and signages surrounding them. Whenever users reach a signage or landmark in their navigation task (in a specific threshold), the system alert them with a notification about the signage and the direction that users need to follow. There are some limitations such as sparse deployment of signages and reading the textual information about them in the interface that might engage users more. In usual turn-by-turn navigation systems, most of the time, according to sensor inaccuracies and the slow speed of users the direction of the arrow is complicate for orientation. They tried to solve this problem with using a simple big arrow and signage in the interface that complemented each other. In their evaluation, most of the users commented that the system improved the perception of the environment.

Dong et al. [49] in their user study found that 3D mobile map representation caused participants perform navigation task less effective, less efficient with higher workload requirement in comparison to 2D mobile map representation. They emphasis with the available disadvantages in both methods, a combination of them is highly recommended and in 2D maps, by showing important landmarks and in 3D maps, by reducing the number of buildings by showing only important ones, cartographers can reduce too much engagement of users during interaction with mobile maps in navigation tasks.

In 2015, Giannopoulos et al. [37] in Switzerland claimed, map interface requires visual attention with switching the users' attention several times to the navigation device. They proposed a novel approach in pedestrian navigation (*GazNav*) which help pedestrians in a way that they are more engage with the real world than the current turn-by-turn navigation systems with using eye tracking glass and vibration technology (it provides hand-free navigation which allows user to keep the visual attention to the environment). They have compared their new technique with a map-based turn-by-turn instructions approach in terms of effectiveness, efficiency, spatial learning and user experience.

Their results showed the GazNav not only outperformed the current turn-by-turn navigation in all the criteria but also performed excellent with significantly better local spatial knowledge.

Ishikawa et al. [53] in their user study found that participants looked at the device screen map with longer time with paying less attention to the surroundings, than the paper map. In using paper map, their participants had difficulty knowing their current location on the map.

Partala et al. [54] have compared three types of currently popular mobile map visualization (traditional graphical map, photorealistic satellite map, photorealistic street-level view) with nine participants in the field and understood that users need more time for look at the map while navigating with street-level map and most often preferred graphical map.

There are some other problems, have noticed by usability evaluation studies in this round of reviewed literatures. Determining the right direction when staring the navigation was reported by Vaittinen et al. [52] in a field experiment that induce users to walk and look where the GPS pointer is moving and tapping the buttons for moving between the waypoints while walking was reported to be cumbersome. For overcoming this problem, they suggest when GPS avatar on the map didn't move as expected, the panorama view in recognizing the destination might be helpful. They mentioned some problems in panorama view such as the images were not up to date and needed a long time for downloading. For the later issue they recommended seamless switching between simple map view and panorama view priority of map-based view while images are downloading.

Werkmann et al. [55] proposed a novel technique for information visualization of off-screen objects called MapCube (showing simultaneously focus and context information on the map) and evaluated it with one of the most prominent off-screen visualization techniques, Halo, and have found that their technique was better in terms of *effectiveness* and *efficiency*.

To answer the second sub research question (RQ 2.b), the column *age, male, number of TPs, number of TPs with relevant knowledge, field or lab* and *apparatus* were analysed.

The most frequent usability issue in this iteration (too much zooming and panning operations), has happened in a context within the screen size of 4.03 inch, in age group of 29 years, 87.5% in laboratory with 19 test persons which 89% of them were users (not experts) and 56% male (figure 3-2).

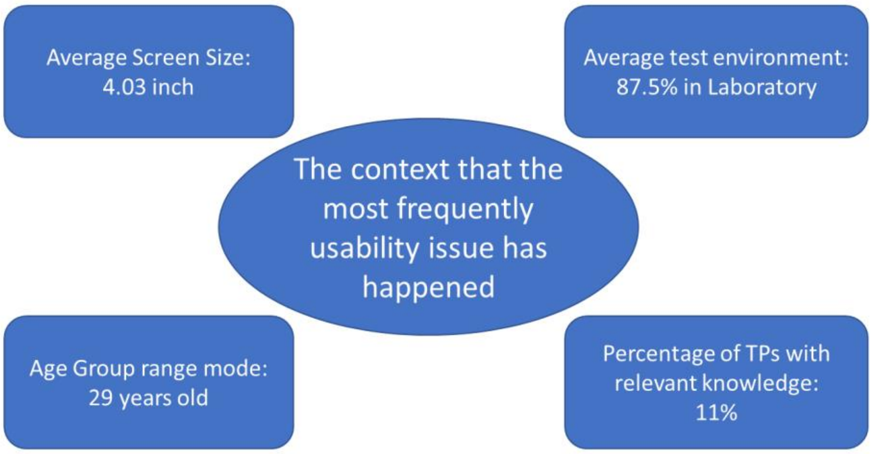

Figure 3-2: The context that the most frequently usability issue has happened in the second iteration.

3.4 Conclusion on the second iteration

The result of this iteration reinforced the outcome of previous one in term of the most frequently usability evaluation method, which in both round of the SLRs was "*questionnaire*".

The way of the evaluating mobile map-based systems in this round of the review has inclined to a direction which more measured the spatial knowledge that user can gain through interaction with mobile maps, the amount of workload that each interface needs, the engagement of user with device and environment.

In the second round of the SLR, the problem of too much *zooming and panning operations* by users of mobile map-based systems has notified. This issue is relevant to the outcome of the first iteration in a way that users, mostly because of *losing the overview* (the most frequently usability issue of the first iteration) of the region on the mobile map, need to perform a lot of *zooming and panning operations*.

In overall, only in one of the reviewed papers the evaluation had operated by "*experts*" which calls "*heuristic evaluation*".

In average in each user study, 20 participants recruited in this round of reviewed literatures with the average of 29 years old. The gender ratio between subjects in this iteration was also nearly equal (54.5% male).

In 72.2% of the studies in the second round of the iteration, prototypes were evaluated instead of real applications.

The most frequency usability issue that occurred during the second iteration mostly has detected through "*questionnaire*", followed by *semi-structured interview* and *experiment observations*.

Chapter 4: Third iteration (Last one)

According to time limitation of the Thesis, this iteration supposed to be the last one to achieve the predefined targets that we are conducting this study to reach them.

4.1 Searching

In the last iteration we have tried to use the search terms that never used before in our search strategies. We added a new search term to our keywords called; exploratory study.

We also used the search term; "*field study*" that before one time in the first iteration used (to use different search terms that might be helpful in achieving new results) and also in order to find the studies that evaluated the most globally used application for smartphones until 2013 (Google Maps) [56] inclined our search to the papers evaluated such popular services. We used "*Google Maps*" in our search terms, but unfortunately there were not enough results in our search that evaluated Google Maps in their evaluation study.

Table 4-1 shows the updated table of the keywords that we used in the last iteration.

Table 4-1: keywords for the last iteration

Core concepts	Synonyms and related phrases
Usability	UX, user experience, user-centered design, UCD, Mobile HCI, mobile user interfaces
Usability defects	Usability issues, Usability problems, Usability flaws
Usability evaluation	Remote usability evaluation, Usability test, Usability inspection, Usability heuristics, Heuristic evaluation, Usability inspection, Usability engineering

Mobile GIS	Mobile Map-based applications, Mobile Map-based systems, Mobile map applications, MMAs, Mobile maps, Location Based Services, Mobile internet map services, Mobile passenger information systems, Pedestrian navigation systems, Online mobile map services, Mobile map services, Mobile map-based tasking interface, Google Maps
Usability evaluation method	Automated usability evaluation method, Usability testing method, Automated usability testing method, Usability inspection method, Usability heuristics method, Heuristic evaluation method, Usability inspection method, User study, Field study, Elicitation study, Technology acceptance model, TAM, Technology acceptance concept, Exploratory study

4.1.1 Search Strings and their Results

The search for achieving the results of the third round of the iteration has operated in January 2019. Between 179 results, in first screening, 14 papers have selected according to their title, abstract and keywords and in some cases the conclusion section that were comparing with our inclusion criteria that we had from the previous round of the iteration (table 4-2).

Table 4-2: Search strings and their corresponding results - third iteration

Search string	Database	Filtered by	Result	Included in first screen	Included for data extraction
"mobile map" "field study" OR "exploratory study" OR "user study" -indoor	Google scholar	2008 to 2018 and Journal of Location Based Services	3	1	1
mobile map "field study" OR "exploratory study" OR "user study" -indoor	Google scholar	2008 to 2018 and chi	3	1	1
"mobile maps" "mobile map-based applications" "field study" OR "exploratory study" OR "user study" OR "usability evaluation" OR "user experience" OR "usability issue" OR "usability problem" "location based services" -indoor	Google scholar	2010 to 2018	2	1	0
"mobile maps" "field study" OR "google maps" "exploratory study" OR "user study" OR "usability evaluation" OR "user experience" OR "usability issue" OR "usability problem" "location based services" -indoor	Google scholar	2008 to 2018	81	6	3
+"mobile maps" AND "field study" OR "exploratory study" OR "user study" OR "usability evaluation" OR "user experience" OR "usability issue" OR "usability problem" AND "location based services" AND "google maps" -indoor -game -privacy	Google scholar	2008 to 2018	37	1	1
"mobile maps" and "field study" "exploratory study" and "user study" "usability evaluation" and "usability issue" "usability problem" "location based services" and "google maps"	ACM digital ibrary	2009 to 2018	53	4	1
Total number			179	14	7

The noticeable point in this iteration was there were not enough relevant results in our search and most of the results were repeatable in previous searches (according to the highlighting theme property that Google Scholar search engine is using). Actually, we have found that in the final included papers there was one paper (in search row number 4 in the table 4-2) which was reviewed in the first iteration and has excluded, that at the end we had only 7 papers to analyse. Figure 4-1 shows the time distribution of the selected papers of the third iteration.

There are some reasons that we have excluded some of the irrelevant studies:

1. Without usability evaluation (the third row)
2. Did not have usability evaluation (user study) (the forth row)
3. About desktop GIS not mobile systems (the forth row)
4. Most of them were irrelevant, some of them were thesis or a part of a book or journal that was too much expensive to buy and some others or repetitive or not in English language or not published in any journal or conference (the fifth row)
5. Evaluating Smarthwatches. Some were posters or thesis, some did not have usability issue or usability evaluation method in their contents, some were indoor navigation or repetitive

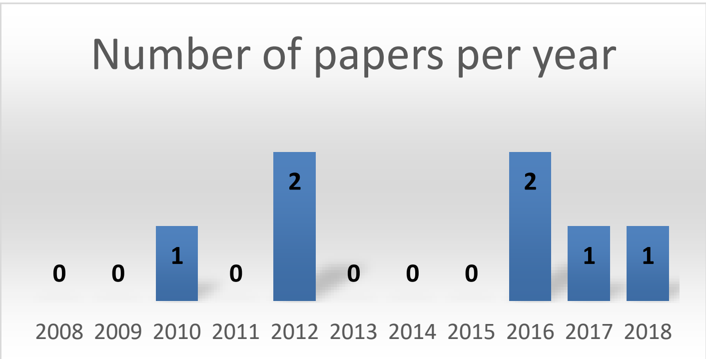

Figure 4-1: The time distribution of the third iteration publications.

4.2 Analysing

The "*Questionnaire*" decisively was the most frequently method (RQ 2) used to detect the usability issues in mobile map-based systems in both iterations, but since it is an important goal in our research, we cannot omit it from our criteria in data extraction.

We need to be sure the results of the last iteration reinforce it or not. But there are some other criteria in our table (Appendix D) that can be deleted to make the table simpler to reflect the most important criteria easier.

We omitted the "*Apparatus*" column, since the mobile devices today have not changed too much (in first iteration the size of the screen for most frequent usability issue was 3.62 inch, and in the second iteration that we had more paper for the recent years was 4.3 inch) and they are still in a certain range to be able to held in hand and this property making them different than desktop or tablet systems (Our goal is limited to mobile systems). There is not any mobile phone in a famous brand such as Apple or Sumsong recently with screen size bigger than 7 inches in the market (most of the prominent and current mobile phones are less than 6 inch) and if they would be bigger than 7 inch, they are belonging to tablet. And if they might be 7 inch, still are in a half size of a normal desktop or laptop screen size (which is around 14 inch) and still have the usability issues of 4- or 5-inch mobile phone (the different is not too much).

The column "*usability metrics*" and the "*measurable criteria*" column can also be omitted from our table since it doesn't convey any new information about the usability issues in the evaluation studies. It would be obvious that most of the studies measured the time for completing the task in order to measure the *efficiency* and the number of errors (or completing the tasks successfully) to measure the *effectiveness*, although these two criteria can be measured by self-reported questionnaire too. Figure 4-2 shows other usability metrics and corresponding measurements to measure them.

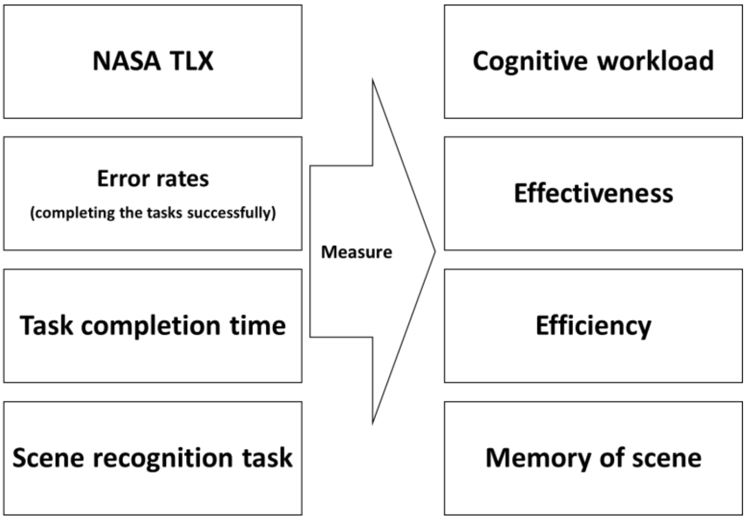

Figure 4-2: Usability metrics with measurable criteria

Therefore, we have 15 columns in our table to extract data in the last iteration with 7 papers (7 rows) to fill with the data that we are extracting.

4.3 Reflecting the Results

In order to answer research question number 1 (RQ 1), the column "**Usability evaluation method**" in the table (Appendix D) has analysed.

Here, in the third iteration, according to the small number of reviewed studies, *interview* after *questionnaire* were the most frequent usability evaluation methods that the researchers used to detect usability issues of mobile map-based systems.

Here is the list of all the usability evaluation methods which used in the reviewed studies in the third iteration that ordered in term of their frequency using (descending):

1. Questionnaire (6 times)
2. Interview (5 times)
3. Logged data (4 times)
4. Video and/or audio recording (2 times)
5. Observation (1 time)

71% of the studies conducted in the laboratory environment, which is in opposite with previous iterations (that 53% in the first round and 60% in the second round of iteration had conducted in field). One possible reason might be the solutions for zoom and pan interaction techniques that proposed by the researchers don't necessarily need to be executed in the field (the focus of this iteration was on this issue). And another possible conclusion might be for the laboratory-based usability studies, *interview* might be suitable and for field-based studies *questionnaire* might be good to detect usability issues in mobile map-based systems.

To answer the research question number 2 (RQ2) the *usability issues* column has analysed. *GPS inaccuracy* was the most frequently usability issue detected in this iteration, followed by *losing overview* and *the need for too much zooming and panning operations* issue. The possible reason for the most frequent usability issue here might be the inclination of the focus to Google Maps usability evaluation, otherwise the results of this iteration reinforce the results of last two ones (*overview* issue in the first iteration and *zooming and panning operations* issue in the second iteration).

Since this issue is belong to technological issue that is not the focus of our work here, we don't need to discuss about it.

El Ali et al. [56] evaluated Google Maps in a developing country (Lebanon) context with poor infrastructures and found out some available usability issues in that context and also investigated the navigation and direction giving strategies (solutions of users in such contexts) used by technology literature people by conducting a qualitative user study. Outdated maps, battery life, incorrect route plans, different names of streets with current names between people, irrelevant direction giving strategies by system than users' technological literacy, the problems of navigation in rural context, poor network connectivity, incorrect or missing places listing on the map, incorrect bus route plans and the generalization of the maps that only showed the main information (lack of detail).

The issue of off-screen objects, screen occlusion of information, unclearness of destination (photo-maps and AR interactions mostly were useful in the destination when users most of the time have problem to find the exact destination) and the offline features available in Google Maps (GM).

Zhou et al. [68] evaluated different interfaces of Google Maps (map view, map with route, satellite view, text view, map and street view, and street view) with two tasks; planning phase and on-path phase in the field by 10 participants and found subjects spent 90% of their time on map view in planning phase while on-path, they spent 56% of time on map, 40% on navigation and 4% on street view and when they were asked about the direction of landmarks and estimate their distance, they used map view 97% of the time. Figure 4-3 shows these views on Google Maps (GM).

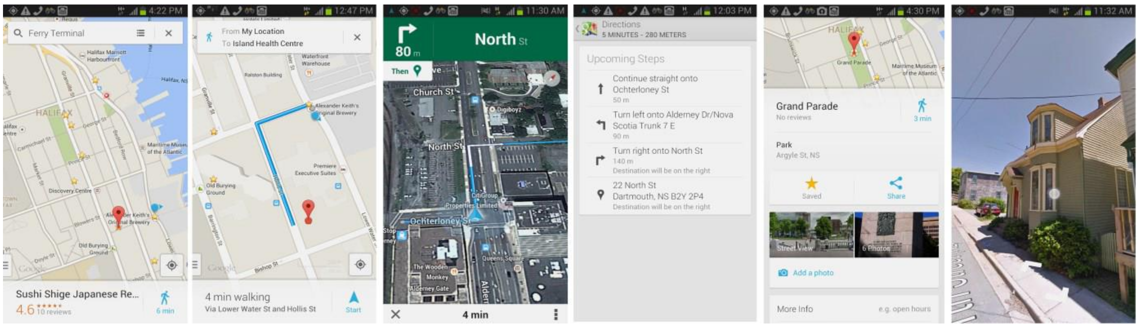

Figure 4-3: Map view, Map with route view, satellite view, text view, map and street view, street view [54].

For *off-screen object* and *losing overview* issues, Miau et al. [57] in Colombia University of United States at 2016, proposed a new method for tackling these problems by using a personalized compass that uses the natural way when people want to describe a location for a person who has a prior knowledge about the region. They first try to evaluate the familiar places such as landmarks that person knows, to give him the direction and orientation according to the relative coordinate of the target with his prior known places. This method (P-Compass) uses this strategy by collecting prior known places from personal GPS, cellular network, location history or from social network traces (facebook check-ins or Google Maps Saved Places) and inferring them from public sources (e.g. Flickr) by data mining. Then with a one or more needles communicate the direction and distance of the target to the prior user's known places (for example to the two famous cities in a low-level scale). Their user study to compare their method with "Wedge" technique (one of the famous techniques for off-screen objects) showed, the majority of the participants preferred P-compass and commented on the difficulty of using "Wedge". For orientation task (which subjects should estimate the direction of an off-screen subway station with respect to the display centre), their results showed it was more challenging for a user to estimate the direction of a distant POI than a nearby one using "Wedge" [52]. At the end they have found two techniques are complementary and offered some design recommendations. These results assert the claim that none of the proposed techniques are not completely suitable enough to apply alone as a map interface in mobile map-based systems, and always a combination of the strong properties of each method has proposed to use.

For the *overview & detail* technique, Concalves et al. [58] from Portugal in 2012, claimed that the overview & detail technique commonly used in desktop applications for visualizing spatial information and video games, but for mobile context, which the screen is small the thumbnail usually is not clearly obvious for user. They mentioned some other usability issues available for this technique such as: greater physical and mental effort for users to deal with that, reducing the available space and some information of the detail might be hidden behind the overview and the small size of the overview. They proposed a novel approach in order to enhance the *overview&detail* technique by designing a resizable overview thumbnail for Overview&Detail technique and with their user study they noticed users spent more time to do the task with resizable overview thumbnail than the classic Overview&Detail method but have less errors with the new method.

Another study operated in University of Rio de Janeiro in Brazil by Maues et al. [59] introduced a new extended version of Anchored Zoom (AZ), which is a technique to overcome the issue of switching between *zooming and panning* in mobile map interactions, by adding Anchor Management to it so called Anchored Zoom with Anchored Management (AZAM). They claimed the AZ technique has some limitations. They have compared their new method with AZ technique, and found that their new approach had superiority upon AZ in terms of *perceived satisfaction,* but their participants commented that usual pinch zoom (PZ) technique was easier to use and learn (*ease of use* and *learnability*) and also the results of time to complete the task (*efficiency*) was not significantly different than other two methods. They confessed that in a big picture usual pinch zoom (PZ) performed better than anchor-based navigation techniques.

Therefore, despite proposing several solutions, the problem of *zooming and panning* and consequently *losing the overview* still remined in mobile map-based systems.

To answer the second sub research question (RQ 2.b), the column *age, male, number of TPs, number of TPs with relevant knowledge, field or lab* and *apparatus* were analysed.

The most frequent usability issue in this iteration (GPS inaccuracy), has mentioned by participants within age group range of **20 to 35 years**, in the **field** or by **online survey** that gathered the users' comments about the available usability issues. with **15 test persons** in average which all of them were **users** (not experts) and consist of **57% male**.

4.4 Conclusion on the third iteration

The results of the search in the last iteration showed there is not any more relevant paper that we have not viewed it before, and we could not be able to operate the last iteration with more than 7 papers.

The outcomes of this iteration reinforced the results of the previous ones with discovering *losing overview* and the need for too much *zooming and panning operations* issues as the most frequently usability issues after GPS inaccuracy (which is belong to technological issues). The tree first of most frequently used usability evaluation methods (*interview, logged data* and *questionnaire*) was the same with the previous iteration and in the first iteration, *interview* ranked in 4th place and *logged data* ranked in 6th place, where *questionnaire* was the most frequently usability evaluation method.

In overall, 22 test persons in average were recruited in each user study. All of the studies evaluated by *users* that were not *experts*. In 20% of the studies, the real applications were evaluated.

Chapter 5: Results

For research question 1 (RQ 1), the overall results of analysing 56 reviewed papers show the 3 most frequently used methods for detecting usability issues in MMSs are (descending):

1. Questionnaire
2. Interview
3. Logged data (screen recording)

After these tree methods, *think aloud* and *video/audio recording* were two methods that were used more than the other usability evaluation methods.

For the second research question (RQ 2), the most frequently usability issues that were reported within all the reviewed papers, were the problem of *losing overview* followed by too much *zooming and panning operations* [7, 11, 9, 19, 17, 6, 20, 18, 16, 22, 26, 52, 53, 54, 35, 37, 36, 38, 39, 40, 41, 34]. To clarify this problem, for example, if user wanted to find the central library of the city of Muenster in Germany, first needs to search the name of it in Google (figure 5-1, a). The first initial usability issue comes to mind here, is the language (if the user is not from Germany (e.g. a tourist), it would be difficult to interact with the system). The second usability issue that comes to mind at the first glance is, the map covers a small portion of the screen that occluded by a lot of textual information (as can see in figure 5-1, b, the upper and lower parts of the screen). The third issue is, there is not enough detail information in this high-level scale, that forces user to zoom out (with two hands pinch interaction) to have a better overview of the area. When user zoomed out, there is not enough detail that user could orient the position of the library according to them (figure 5-1, c), then try to zoom in again. There is not enough clear clue such as a landmark (only the outlines of parcels are presented on the map) to link the map in the mind of user with the represented map (figure 5-1, d). This *losing the overview* and detail of the region that followed by too much *zooming and panning operations* were the most frequently reported usability issues in MMSs in our SLR.

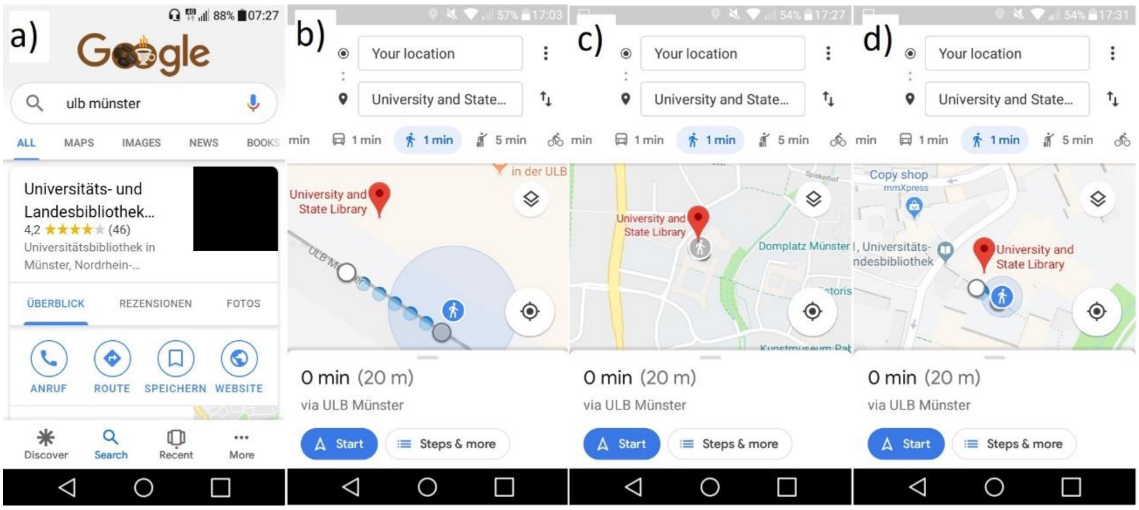

Figure 5-1: losing overview and too much zooming and panning operations

We categorized (RQ 2.a) all the usability problems in two main groups; technological and spatial. Here, we focused only on spatial problems that are available in MMSs which refer to map and map interfaces.

For research question 2.b, we found that the most frequency usability issue that occurred during the all reviewed papers (*losing overview* followed by too much *zooming and panning operations*) happened in the context that shows in figure 5-2. For gender differences, we cannot detect a trend between two sexes since overall, 54% of the test persons that the most frequent usability issue happened between them were male (researchers hired nearly equal number of genders to make their study counterbalance).

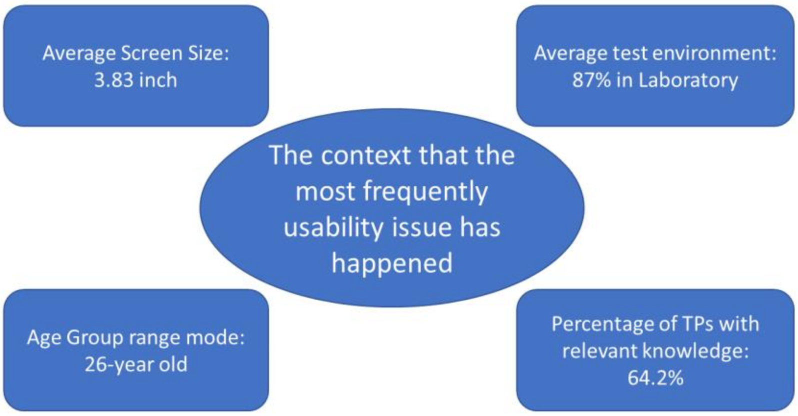

Figure 5-2: the overall context that the most frequent usability issue happened

As we have in research question 2.c, there are a lot of methods that have developed to overcome the issue of *losing overview* and too much *zooming and panning operations*.

There are some approaches, proposed that act such a cue for showing the direction and distance of off-screen objects to user to have a better overall view about the region and the POIs that are located out of the map that displayed on the screen.

Paolina et al. [16] proposed a new technique in 2008 called "*Framy*" that is a visualization method that uses a semi-transparent cornice shape with colour intensity for resembling off-screen objects on mobile maps.

"*Wedge*" proposed by Gustafson et al. in 2008, that is a visualization technique to convey the direction and distance of off-screen objects through the direction and size of triangles.

"*Scaled arrows*" is another technique proposed by Burigat et al. in 2006, which using different size arrows for visualizing off-screen objects on mobile maps.

Hooten et al. [19] proposed paper maps, when users of mobile maps losing the overview of the space.

There are other methods proposed (we have not reviewed them here) that act such a cue for visualizing off-screen objects on mobile maps that the most famous ones are; Fisheye view (Plaisant et al. 1995), Halo (Baudisch, 2003), Hop (Irani et al. 2006) etc.

Another method has proposed recently (2016) by Miau et al. [57] for treating the issue of *off-screen objects* and *losing the overview* that with a compass that contains of two noodles (personalized compass) convey the distance and direction of the off-screen objects. They compared their novel method with one of the famous methods in visualizing the off-screen objects "*Wedge*" and noticed, the majority of the participants preferred P-compass and commented on the difficulty of using "*Wedge*". For their orientation task (which subjects should estimate the direction of an off-screen subway station with respect to the display centre), their results showed it was more challenging for a user to estimate the direction of a distant POI than a nearby one using "*Wedge*".

None of the above-mentioned techniques for off-screen visualization have not applied in today's mobile map services such as Google Maps. Each one has its limitations that needs more cognitive and physical efforts of users and adds the problem of *learnability* and *ease of use*.

There is another famous method for mitigating the problem of losing the overview is "*Overview&Detail*" proposed by Plaisant et al. in 1995 that is a visualization technique that for showing the overview of the space, uses a small thumbnail that covers around 10% of the detail view map that displays on the mobile screen (usually on the bottom right corner).

This method also has some problems such as; the need for stablishing visual connection between both views, small size of overview thumbnail in mobile context [17] that is too small to read the overview map (this method proposed firs for desktop applications such as video games [53]) and the scalability issue (the material of LBS course by Christian Kray).

Elzakker and Delikostidis [6, 12] proposed a reverse technique that shows the detail view on the small thumbnail and in the main map that is represented in full screen shows the overview (figure 2-4).

Concalves et al. [53] designed a resizable thumbnail and proposed it in 2012 to enhance the *Overview&Detail* technique, but in their user study they noticed users spent more time to do the task with resizable overview thumbnail than the classic *Overview&Detail* method.

There are some techniques to enhance *zooming and panning operations* in MMSs such as semi-automatic zooming (SAZ) [26], content-based zooming [34], tilt-based zooming [39, 41], smooth zooming /panning and Vario-Scale Maps.

Burigat et al. [20] provided users some predefined zoom levels in their prototype to choose a specific zoom level directly.

Anchored Zoom (AZ) with using a reference point as a main tool try to overcome the problem of switching between zooming and panning, but Maues et al. [59] believed that it has some limitations and proposed a new method called AZAM (Anchored Zoom with Anchored Management) by adding new features that better exploit the use of anchors, improving the choice of new anchors and access to the previous ones. They have compared their new method with AZ technique, and found that their new approach had superiority upon AZ in terms of *perceived satisfaction*, but their participants commented that usual *pinch zoom* (PZ) technique was easier to use and learn (*ease of use* and *learnability*) and also the results of time to complete the task (*efficiency*) was not significantly different than other two methods. They confessed that in a big picture usual *pinch zoom* (PZ) performed better than anchor-based navigation techniques.

With proposing a lot of techniques for enhancing *zooming and panning operations* and the issue of *losing overview* on MMSs, the current most famous mobile map services such as Google Maps still using usual *pinch zoom* (PZ) (since it is easy to learn and use and also is familiar to users), and the problem of *losing overview* and too much *zooming and panning operations* remined yet.

Overall, in our reviewed literature, only 8% of studies recruited *experts* to test the usability of the systems.

Only 30.5% of the evaluated systems in all of the reviewed studies were real applications and services and around 70 percent of the evaluated systems were prototypes.

As can be noticed in figure 5-3, most of the all reviewed papers (59%), between 11-year range of our pre-defined search, are distributed in 5 years (a range 2010 to 2014), and then 16% of the papers are from the year 2016. The median value for the time distribution of all the reviewed studies was the year 2012.

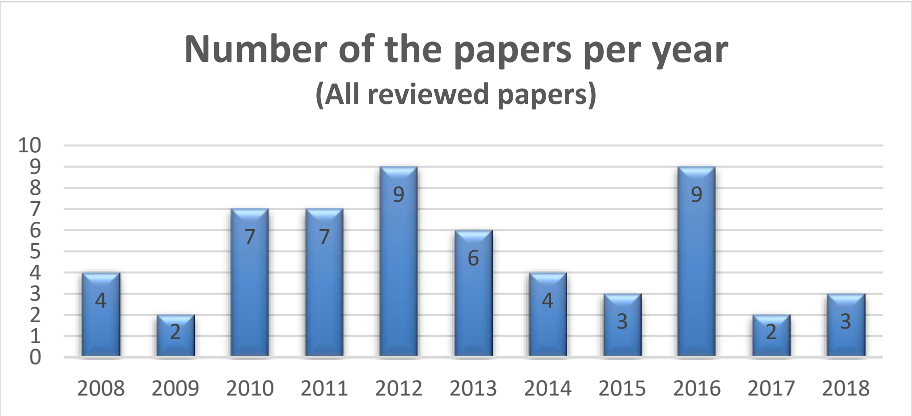

Figure 5-3: Time distribution of all 56 reviewed papers

Overall, figure 5-4 shows all the criteria that have measured within *questionnaire* method between reviewed papers to evaluate usability of mobile map-based systems.

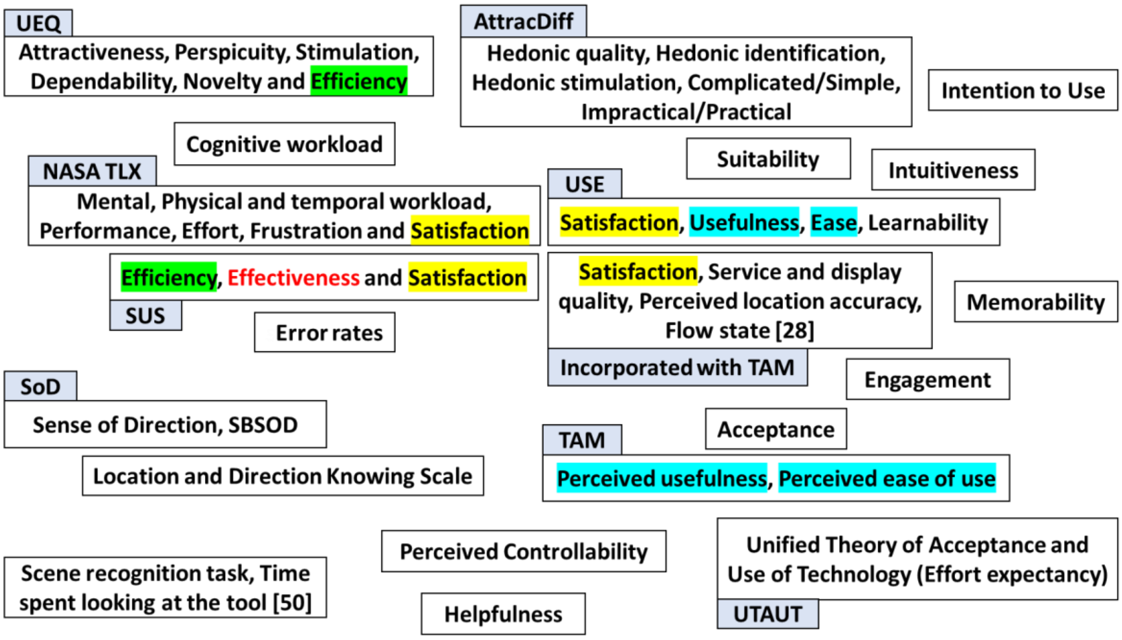

Figure 5-4: all types of questionnaire have used in the SLR. (Colourful and highlighted ones are the most frequently used and common between different methods)

As we can see in the figure above (figure 5-4) there is a trend in using questionnaire method in studies in mobile map-based systems which focused more on *satisfaction*, *effectiveness* and *efficiency*, which forming the usability criteria of *System Usability Scale* (SUS). According to Wikipedia, “in systems engineering, SUS is a simple ten-item attitude Likert scale giving a global view of subjective assessments of usability”. It was created by Brooke in 1986 ([10] claims it has presented in1996) in UK, but ISO standard (ISO 9241, part 11) claims that usability of a system can be measured only by considering the user’s *context*.

Efficiency and *effectiveness* are two criteria that not only can be measured with subjective questionnaire, but also with measuring the *task completion time* and *error rates*.

The results of the user study of Park et al. [27] indicated that *satisfaction* and *perceived usefulness* of the mobile map services were the most significant antecedents of the attitude of users to use the service.

In *questionnaire* method, that was the most frequently used in our study, after *SUS* method and its' correspondent criteria, the second most frequent measurement in our survey was *NASA TLX*, which 12 times has been used in the reviewed literature for usability evaluation in Mobile Map-based Systems (MMSs).

The interesting point is, [2] used NASA TLX in order to achieve *satisfaction*, since SUS and USE methods also measuring this criterion (figure 5-4).

Kratz et al. [26] used *NASA TLX* and *USE* questionnaire to evaluate perceived *learnability*, *ease of use*, *satisfaction* and *usefulness* of their novel technique that proposed for enhancing zoom operation.

According to table 5-1, *interview* and *logged data* methods, most frequently used together (in combination) in the reviewed studies [67, 52, 63, 64, 68, 58, 59, 14].

Table 5-1: studies that used a combination of some usability evaluation methods

Num	First Author [reference]	Questionnaire	NASA TLX	SUS	USE	SBSDS	Think aloud	Logged data	Observation	Interview	Video/Audio recording
1	Kratz [31]		✓		✓						
2	Elzakker [5, 13, 60]	✓					✓		✓	✓	✓
3	Delikostidis [24, 51, 66]	✓					✓			✓	✓
4	Delikostidis [29]				✓		✓				✓
5	Konkol [42]						✓			✓	
6	Wenig [67]			✓			✓	✓		✓	
7	Wither [50]	✓								✓	
8	Vaittinen [52]	✓						✓		✓	
9	Van Tonder [46]		✓					✓			
10	Mulloni [63]							✓		✓	✓
11	Iwata [61]	✓						✓			
12	Ranasinghe [64]	✓				✓		✓		✓	✓
13	Zhou [68]					✓		✓	✓	✓	✓
14	Concalves [58]							✓		✓	
15	Maues [59]	✓					✓	✓		✓	
16	Ramsay [19]									✓	✓
17	Burigat [22]	✓						✓			
18	Flink [6]			✓			✓				
19	Duenser [16]		✓						✓	✓	✓
20	Rehrl [18]			✓		✓					✓
21	Beul-Leusmann [14]			✓			✓	✓		✓	

Elzakker et al. [6] believed *thinking aloud* method with *screen logging* and *observation* led to the most valuable results in requirement analysis of their User Centre Design (UCD) process. Delikostidis and Elzakker combined a lot of methods in their field studies [4, 5, 13, 24, 29, 51, 60, 66] (in the table 5-3, row 2, 3 and 4) to evaluate usability problems in Google Maps that might be time and money consuming. In row 3 of the table [24, 51, 66], they used a new method which was Synchronous screen logging/multi-camera recording (recording user's actions in the field study in a multi-view manner) in combination with other usability evaluation methods. They always used *think aloud, interview* and *video/audio recording* in their user studies. *Think aloud* and *interview*, which used together in [4, 14, 43, 60], force users to talk about their view points around the system, but on one hand, as an anthropologist[18] said, "What people say, what people do, and what people say they do are entirely different things." and on the other hand, measuring *efficiency* of a system (task completion time), is impossible with the *think aloud* method.

Dong et al. [49] combined *think aloud* method with *eye tracking* (the only study that used this method in reviewed literature) and *synchronous audio and video recording* to evaluate *efficiency*, *effectiveness* and *cognitive workload* of users interacting with 2D and 3D representations of Google Maps.

Partala et al [54] combined *NASA TLX questionnaire* with *AttrakDiff* (Jassenzahl, 2003) to compare 3 kinds of mobile map visualization namely; traditional graphical map representation, photo realistic satellite map and photo realistic street-level view. They selected two scale components of *AttrakDiff* such as; complicated or simple, impractical or practical, unprofessional or professional and some hedonic stimulations and

[18] Margaret Mead

identifications and attractiveness. For example, they found photorealistic maps more *stimulating* to users than graphical maps, but less *pragmatic*. They noticed street-level view demands higher *task load.* At the end, their participants mostly preferred graphical map visualization.

In the reviewed literatures, only El Ali et al. [56] conducted the usability study in a developing country (Lebanon) with considering poor available infrastructure and the strategies local people employing in facing with those challenges when they use Google Maps[19]. They used *questionnaire* (web survey) with an open-ended and semi-structured *interview* (only with technology literate locals) in parallel. With interview they deeply went through some usability problems available in interacting with Google Maps in Lebanon such as; Multifaceted information access and direction giving strategies (people didn't only use Google Map service for navigation, using landmarks and traffic directions were not appropriate in their context etc.), Technological reliability (outdated mapping of locations, GPS inaccuracies, incorrect route plans, battery life issues and so on), Language ambiguity, conventions, and technology, Technological literacy and urban-rural divides. In their open-ended online questionnaire survey, they deliberately didn't ask questions about technological literacy, since they believed it would have been difficult to verify from a survey (they conducted an online survey), but they asked participants about the basic demographic of them, information seeking strategies for finding unfamiliar places and navigating there, the challenges they faced, and the ways the overcome them.

Flink et al. [6] combined *think aloud* method with *SUS questionnaire* to study *ease of use* and *usefulness* of a multi-publishing service (the web map, a mobile map app and

[19] Most globally used smartphone app [52]

the paper map) for hikers called MenoMaps. They believed a *questionnaire* right after the *thinking aloud* method is viable since it helps to reveal subjective opinions and user satisfaction.

One surprising point worth to mention here is, between 56 papers that reviewed in this thesis, 12 papers (most of the reviewed papers) were from Germany, followed by the Netherlands with 8 papers. A basic conclusion from these results can show the importance of *usability* in such western European countries.

Chapter 6: Discussion

Today, maps are in most of places, from a bus stop and an office (as a static traditional paper form) to an interactive and dynamic format such as the map displaying at the mobile phones. But the question is; how much people paying attention to them? The answer comes back to the main duty of the maps. They have cartographed to convey spatial knowledge to users, which by that knowledge, users could have better spatial ability to execute some tasks (mental or physical). The map should convey this knowledge in a proper way that with minimum time and effort, user could be able to execute the tasks easily more efficient and effective with satisfaction in any context. In mobile context, because of a lot of limitations that mentioned in this study, the map to be present, needs to be too much sophisticated to be able to enhance user experience. The usability issues in map representations in mobile context still have remained, although with some methods researchers and designers are detecting most of them and proposing some methods, approaches and techniques to overcome them, but each one beside addressing one issue, adding another usability issue(s) when tried to solve a problem. For example some of the famous traditional methods such as *Wedge*, *Framy* and *Scaled Arrows* and *Overview&Detail* approaches for providing better overview and visualizing off-screen objects on maps, and also new methods such as personalized compass [57] (P-compass) and some methods for enhancing zoom and pan operations such as tilt-based zooming and semi-automatic zooming (SAZ) and so on, add some problems such as *learnability* and *ease of use* to the systems and need greater physical and mental effort.

The available general methods for detecting the usability issues in mobile map-based systems (MMSs) are not completely fit to this context to lead the evaluators to specific problems around map and map interactions. Most of the time, in usability evaluation of mobile map-based systems a combination of such general usability evaluation methods was used to achieve some qualitative (such as experiment observation methods, interviews and user comments) and quantitative (such as time taken to do the tasks, questionnaire, user's workload, and error rates) data and mostly (around 80% of the cases [13, 55]) executed in the lab without considering the actual contexts, that not clearly addressed the specific map-based issues and the strategies users developing in interacting with these systems and not taking into account the cultural differences and user behaviour. Only [52] between reviewed studies operated an uncontrolled user study with considering a real context of users that were travelling to some cities as tourists to use the prototype (City Scene) during their trip with a diary for reporting experiences and activities with a *questionnaire* that followed by an *interview*. The combination of these two methods (*questionnaire* and *interview*) was using in a lot of reviewed literature to detect usability issues of MMSs. Another frequent combination of methods was a combination of *logged data* with *interview* [67, 52, 63, 64, 68, 58, 59, 14]. Another worth mentioning case is the one by Park et al. [32] that used technology acceptance model (TAM) parameters with proposing and applying five complemental behavioural intention factors to use the service (locational accuracy, satisfaction, service and display quality, mobility and flow state) through in-depth interview and online survey with two groups of users (1109 participants) and experts. Aditya et al. [47] proposed 5Es (Effective, Efficient, Engaging, Error tolerant and Easy to learn) on evaluating the effectiveness of different map displays, but those 5Es are not completely different than available criteria in questionnaire method that applied by the most of usability evaluation studies in MMSs. There is not a common approach to follow in evaluation of mobile map-based systems

(MMSs), although in our overview *questionnaire*[20], *interview* and *logged data* (screen recording) were respectively three most frequently methods used to detect usability issues in MMSs. For detecting the "*loosing overview*" and "*too much zooming and panning*" issues (that here were the most frequently usability issues), 85% of the time *questionnaire* were used followed by *logged data* (30% of the cases) and a combination of *think aloud* and *interview*[21] (25% of the cases). *Questionnaire* method is limited to questions with limit scales that evaluators provided for participants that cannot widely and deeply detect some of the important usability issues of MMSs. Only small number of cases used open-ended questionnaire that users freely can communicate their viewpoints about the system. Participants are freer to state their opinions with *interview* and if it would be semi/unstructured might detect the issues more deeply in MMSs. Some of the fundamental map interaction problems only can be detected by experts which calls *heuristic evaluation* that we didn't have enough papers to exploit it in our overview. Wenig et al. [60] combined *think aloud*, *SUS questionnaire* and *logged data* with an *interview* at the end to evaluate different combination of image-map visualization. Wither et al. [50] combined *questionnaire* and *interview* to compare traditional map-based navigation with panorama-based navigation. Vattinen et al. [52] used *questionnaire*, *logged data* with *interview* to compare map-based navigation with panorama view. Partala et al. [54] used a new kind of *questionnaire* introduced by Hassenzahl in 2003 which measuring some attributes such as; complicated or simple, practical or impractical, and some parameters to evaluate the pragmatic quality and hedonic identification and stimulation that called AttrakDiff. They found photorealistic maps were more stimulating to the users than simple graphical maps, however photorealistic maps were perceived less

[20] In questionnaire method mostly the *effectiveness*, *efficiency* and *user satisfaction* (forming SUS questionnaire) were the most frequently criteria to evaluate usability issues in MMSs
[21] Mostly unstructured or semi-structured interview

pragmatic than geographical map. There is not any paper in our reviewed studies to systematically and directly compare most of the common map interfaces (2D/3D views, satellite view, photorealistic view, panorama and street level view, AR view etc. that are the key methods to represent the spatial information) at same case study[22]. There are not compelling results in our reviewed studies to point out some vital spatial usability issues available in MMSs and some of the usability problems are underexplored in the literature. For example Google Maps that is the widely used system [61] still is working with a lot of usability problems and this system following the same approach that using in desktop applications in mobile maps. Elzakker et al. [4] believe experiences with design and producing desktop computer or paper maps and vehicle navigation systems cannot be suitable to use in developing map displays in mobile devices for pedestrian navigation. Usually, when people are giving others directions, they will frequently using landmarks to describe the route [50]. Applying landmarks in pedestrian navigation mobile maps has proposed by a lot of studies [4, 5, 13, 29, 34, 42, 49, 50, 51, 61, 62, 63]. Delikostidis et al. [51] believe landmarks foster the relationship between the real world, the mobile map and the mental map of the mobile users. Elzakker et al. [13] claimed landmark photos that pop-up when clicking them are more preferable than 3D models in MMSs and also suggested the map in mobile context should be simple, not be overload with many symbols or 3D buildings. The results of the user study of Dünser et al. [16] showed users were slower in 3D map visualization in initial orientation and route finding in comparison to 2D maps. Some types of visualization techniques for representing geospatial data in mobile maps are not merely suitable and in most of the reviewed literature, a traditional simple 2D map representation was preferred than other forms by users. But the map

[22] The comparisons only did two by two or maximum 3 or 4 interfaces were comparatively evaluated

content that is representing in mobile maps, for example Google Maps that use the simple 2 dimension map is not suitable in mobile applications which instead of displaying some salient landmarks or other generalized and abstract geographical features, sometimes in a high scale levels only shows the green lands, water bodies and outline of the buildings without any label that are meaningless entities in map interaction of mobile maps and induces user to zoom and pan to gather the spatial understanding of the map. There are some other usability issues available in Google Maps that none of the reviewed studies pointed out; continues need of internet connectivity in navigation task, lack of up-to-dateness of maps, overload of information (user needs to have a mobile phone with high configuration in terms of memory and high speed CPU and high speed internet to be able to download and display the heavy spatial information), considering user's context (age, technological affinity, literacy, language, internet speed and apparatus). Lack of taking into account the topology of the area when calculating the travel time by walking, the orientation problem (at the start and end points of the navigation, users usually have some difficulties to orient themselves in the environment according to map and the direction of the user that displays by the arrow on the map), sometimes the system is too much intrusive, it should let users have more freedom in choosing the route (sometimes such as tourist contexts, users have more time to stray and enjoy the environment). Such commercial systems might have some usability tests that not published but according to Park et al. [32] research is essential to guide the industry toward success.

Ohm et al. [48] found that presentation modes of pedestrian navigation systems should be adaptive to users' sense of direction (SoD), which in their findings, badly oriented users tended to prefer standard map-like interfaces and well-oriented users seemed to prefer abstract designs. According to Delikostidis et al. [24] traditional desktop-oriented maps are not always practicable for mobile map-based systems on small screens. We came up with the idea that current map interfaces (several scale levels) in mobile map-based systems are not completely usable for this application. They induce users to do a lot of zooming and panning operations that making the map interaction more confusing. Spatial information should be conveyed to user in a way to quickly build up an overview of the space in his mind with enough detail that makes them more independent of system with more choices to freely select any possible routes to the destination in navigation and way finding tasks. There is a need for a fundamental change in representing spatial data in MMSs. A data model should be designed with relevant specialized entities for mobile map-based systems to rebuild a system with specific features that should be necessary to display on the mobile maps to be suitable in most of the contexts.

Available usability evaluation methods are not perfectly detecting the spatial usability issues of current mobile map-based services. A suitable usability study might be like the user study of El Ali et al. [56] which they conducted two user studies in parallel (one with usual TPs with an online questionnaire and another with technology literature TPs with open-ended and semi-structure interview) in a developing country context with poor infrastructures and low technology affinity people. An ideal usability study might be in such a way that gives the experts (e.g. 8 experts) a long period of daily life experience with the system (e.g. one month) in different contexts of use and at the end with open-ended semi-structure interview try to collect their comments, feedbacks and suggestions, or with logged data (screen recording) record the real user behavior (when users feel not observing by a person) with combination of an online questionnaire.

Chapter 7: Conclusion

In mobile map-based systems (MMSs) usability is bolder than other systems, since such touch-based, small screen (which representing maps usually needs a big screen) devices need to be very precisely designed for users (that are a wide range of people with diverse technological literacy) in order to assist them in their tasks rather than confusing them. There are a lot of usability issues available in such systems which most of them have detected by some usability evaluation methods. Here, we operated a systematic literature review (SLR) with 56 papers based on three iterations to first, find the most frequently used usability evaluation methods that were used to detect usability issues of mobile map-based systems (MMSs) and then the most frequently usability problem that occurred in their usability studies and categorizing them and the context they might happen, and the solutions have developed so far for resolving them. The results of SLR show the *Questionnaire* was the most frequently usability evaluation method that used to detect the usability issues in mobile map-based systems (MMSs). Other most frequently used methods were respectively; *interview, thinking aloud, logged data* and *video recording*.

The most frequently usability issue that occurred in MMSs and detected in the first iteration was, *losing the overview*. The results of the second iteration also was interrelated to this issue which was *too much zooming and panning operations*, that users usually when losing the overview of the region on the mobile map, need to perform a lot of zooming and panning operations. In the last round of the iteration, after GPS inaccuracy (which refers to technological problems), *losing the overview* and *too much zooming and panning operations* has recognized for the most frequently reoccurred usability issue between reviewed studies.

Therefore the results of the three iterations were completely relevant to each other that first iteration showed the *losing overview* as the most frequent usability issue that the second iteration reinforced it by showing *too much zooming and panning operations* issue that is mostly because of losing the overview and detail on maps and the last iteration found both issues at the most frequently usability issue in MMSs that in all three iteration the most frequently usability evaluation method was *questionnaire*.

The outcomes of the last iteration were some solutions to overcoming the discovered issues. None of the solutions (some famous traditional solutions such as; Overview&Detail, Wedge, and Framy and some new techniques such as personalized compass (P-compass by Miau et al. [57])) completely and alone were not successful to solve the most frequently reoccurred issues (*losing overview* and *too much zooming and panning operations*) and MMSs still have a lot of usability issues.

Overall, three most frequently used usability evaluation methods in our review were respectively; *questionnaire*, *interview*, and *logged data* (*screen recording*). Although Burghardt et al [12] believed a combination of "*think aloud*" and "*video recording*" is most suitable for the evaluation of mobile devices in the field, here, these two methods were respectively the most frequently methods used in our SLR after those above-mentioned three methods. *Losing overview* followed by *too much zooming and panning operations* (related to spatial issues) were the most frequently reoccurring issues during the all reviewed papers [7, 11, 9, 19, 17, 6, 20, 18, 16, 22, 26, 52, 53, 54, 35, 37, 36, 38, 39, 40, 41, 34]. These issues have detected in a context with mobile phones with 3.83 inches in average, 87% of the cases in the laboratory, (according to Elzakker et al. [5], [4] most of the studies (81%) on the usability evaluation of mobile geo-applications are executed in the laboratory) within participants with averagely 26-year-old, that 64.2% of them had relevant knowledge. In 85% of the cases these issues detected by *questionnaire*

[24, 5, 20, 23, 26, 57, 59, 40, 41, 43, 39], 30% of all the studies noticed these issues, used *logged data* [14, 44, 58, 23, 26, 22, 59] and in 25% of the cases a combination of *think aloud* and *interview* used [14, 24, 5, 42, 59].

There is not any correlation between the number of test persons (TPs) and detected usability issues. According to Nielsen (1989) and Virzi (1992), with at least 5 or 6 number of test persons, approximately 80 percent of usability problems can be detected [12]. Overall, in the entire reviewed papers, 18 test persons in average recruited for each user study (we excluded two online surveys in the user studies as outliers, since one operated by 1109 users and the other one by 112 users).

In 30% of the reviewed studies, real systems (such as Google Maps) were evaluated, and the remained case studies operated only by prototypes.

Chapter 8: Bibliography

[1] Y. Xiao and M. Watson, "Guidance on Conducting a Systematic Literature Review," 2017.

[2] A. Jiamsanguanwong and P. Thanachan, Comparative Usability Evaluation of Mobile Map Applications, 2016.

[3] K. Rehrl, E. Häusler, R. Steinmann, S. Leitinger, D. Bell and M. Weber, "Pedestrian navigation with augmented reality, voice and digital map: Results from a field study assessing performance and user experience," in *Lecture Notes in Geoinformation and Cartography,* 2014.

[4] c. Van Elzakker and i. Delikostidis, "User-Centered Design of Mobile Geo-Applications," 2012.

[5] C. P. J. M. van Elzakker, I. Delikostidis and P. J. M. van Oosterom, "Field-Based Usability Evaluation Methodology for Mobile Geo-Applications," *The Cartographic Journal,* 2008.

[6] H.-M. Flink, J. Oksanen, U. Pyysalo, M. Rönneberg and L. T. Sarjakoski, "Usability Evaluation of a Map-Based Multi-Publishing Service," vol. 1, 2011.

[7] B. Kitchenham and S. Charters, "Guidelines for performing Systematic Literature reviews in Software Engineering Version 2.3," *Engineering,* vol. 45, no. 4ve, p. 1051, 2007.

[8] B. A. Kitchenham, P. Brereton, M. Turner, S. Linkman, R. Pretorius, D. Budgen, B. A. Kitchenham and M. K. Niazi, "Refining the systematic literature review process—two participant-observer case studies," *Empirical Software Engineering,* 2010.

[9] E. M. Schön, J. Thomaschewski and M. J. Escalona, "Agile Requirements Engineering: A systematic literature review," *Computer Standards and Interfaces,* 2017.

[10] N. S. M. Yusop, J. Grundy and R. Vasa, "Reporting Usability Defects: A Systematic Literature Review," *IEEE Transactions on Software Engineering,* 2017.

[11] T. C. Lacerda and C. G. von Wangenheim, *Systematic literature review of usability capability/maturity models,* 2018.

[12] Burghardt D and Wirth K, "CO-128 COMPARISON OF EVALUATION METHODS FOR FIELD-BASED USABILITY STUDIES OF MOBILE MAP APPLICATIONS".

[13] C. P. J. M. Van Elzakker and I. Delikostidis, Use, User, and Usability Research for Mobile Geo-Applications for Personal Orientation and Navigation.

[14] S. Beul-Leusmann, C. Samsel, M. Wiederhold, K. H. Krempels, E. M. Jakobs and M. Ziefle, "Usability evaluation of mobile passenger information systems," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics),* 2014.

[15] E. Coluccia and G. Louse, "Gender differences in spatial orientation:A review," *Journal of Environmental Psychology,* 2004.

[16] A. Dünser, M. Billinghurst, J. Wen, V. Lehtinen and A. Nurminen, “Exploring the use of handheld AR for outdoor navigation,” in *Computers and Graphics (Pergamon),* 2012.

[17] J. Wen, W. S. Helton and M. Billinghurst, “A study of user perception, interface performance, and actual usage of mobile pedestrian navigation aides,” in *Proceedings of the Human Factors and Ergonomics Society,* 2013.

[18] K. Rehrl, E. Häusler, S. Leitinger and D. Bell, “Journal of Location Based Services,” 2014.

[19] A. Ramsay, M. McGee-Lennon, G. A. Wilson, S. J. Gray, P. Gray and F. de Turenne, “Tilt and go: Exploring multimodal mobile maps in the field,” *Journal on Multimodal User Interfaces,* 2010.

[20] L. Paolino, M. Sebillo, G. Tortora and G. Vitiello, “Framy - Visualising geographic data on mobile interfaces,” *Journal of Location Based Services,* 2008.

[21] A. Bouwer, F. Nack and A. El Ali, “Lost in navigation,” in *Proceedings of the 14th ACM international conference on Multimodal interaction - ICMI '12,* 2012.

[22] S. Burigat and L. Chittaro, “Visualizing references to off-screen content on mobile devices: A comparison of Arrows, Wedge, and Overview + Detail,” *Interacting with Computers,* 2011.

[23] S. Burigat, L. Chittaro and E. Parlato, Map, Diagram, and Web Page Navigation on Mobile Devices: the Effectiveness of Zoomable User Interfaces with Overviews, 2008.

[24] I. Delikostidis, C. P. Van Elzakker and M. J. Kraak, “Overcoming challenges in developing more usable pedestrian navigation systems,” *Cartography and Geographic Information Science,* 2016.

[25] E. R. Hooten, S. T. Hayes and J. A. Adams, “A comparison of communicative modes for map-based tasking,” in *Proceedings of the Human Factors and Ergonomics Society,* 2011.

[26] S. Burigat, L. Chittaro and S. Gabrielli, “Navigation techniques for small-screen devices: An evaluation on maps and web pages,” *International Journal of Human Computer Studies,* 2008.

[27] B. van Tonder and J. Wesson, “Is tilt interaction better than keypad interaction for mobile map-based applications?,” in *Proceedings of the 2010 Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists on - SAICSIT '10,* 2010.

[28] J. M. Noguera, M. J. Barranco, R. J. Segura and L. Martínez, “A mobile 3D-GIS hybrid recommender system for tourism,” *Information Sciences,* 2012.

[29] I. Delikostidis, H. Fritze, T. Fechner and C. Kray, “Bridging the Gap Between Field-and Lab-Based User Studies for Location-Based Services,” 2015.

[30] L. Kuparinen, J. Silvennoinen and H. Isomäki, “Introducing Usability Heuristics for Mobile Map Applications”.

[31] S. Kratz, I. Brodien and M. Rohs, Semi-Automatic Zooming for Mobile Map Navigation.

[32] E. Park and J. Ohm, “Factors influencing users' employment of mobile map services,” *Telematics and Informatics,* 2014.

[33] P. K. a. M. R. Ioannis Giannopoulos, “GazeNav: Gaze-Based Pedestrian Navigation,” *MobileHCI,* 2015.

[34] I. O. B. N. U. Huseyin Zahit Selvi, “Landmark Orientation and Map Design for Pedestrians: Prototype of a Selcuk University Campus Area Pedestrian,” *Earth and Environmental Science,* p. 44, (2016).

[35] C. G. a. M. Jöst, “Indoor and Outdoor Navigation in Smart,” *Smart Mobility – Connecting Everyone,* 2017.

[36] M. W. Liping Yang, “Similarities and diferences betwen outdoor and indoor space from the perspective of navigation,” 2014.

[37] I. Giannopoulos, P. Kiefer and M. Raubal, “GazeNav,” in *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services - MobileHCI '15,* 2015.

[38] C. Gkonos, I. Giannopoulos and M. Raubal, “Maps, vibration or gaze? Comparison of novel navigation assistance in indoor and outdoor environments,” *Journal of Location Based Services,* 2017.

[39] Y. K. Cheung, Z. Li and W. Chen, “Integration of Cognition-based Content Zooming and Progressive Visualization for Mobile-based Navigation,” *The Cartographic Journal,* 2009.

[40] L. Chen and D. Chen, “CaMap: Camera-based Map Manipulation on Mobile Devices,” in *Proceedings of the 2nd International Conference on Computer Science and Application Engineering - CSAE '18,* 2018.

[41] J. Grubert, M. Pahud, R. Grasset, D. Schmalstieg and H. Seichter, “The utility of Magic Lens interfaces on handheld devices for touristic map navigation,” *Pervasive and Mobile Computing,* 2015.

[42] M. Konkol, C. Kray and M. Ostkamp, “Follow the signs—countering disengagement from the real world during city exploration,” in *Lecture Notes in Geoinformation and Cartography,* 2017.

[43] I. Giannopoulos, P. Kiefer and M. Raubal, “The influence of gaze history visualization on map interaction sequences and cognitive maps,” in *Proceedings of the 1st ACM SIGSPATIAL International Workshop on MapInteraction - MapInteract '13,* 2013.

[44] B. P. Van Tonder and J. L. Wesson, “Improving the controllability of tilt interaction for mobile map-based applications,” *International Journal of Human Computer Studies,* 2012.

[45] B. van Tonder and J. Wesson, “The Impact of Sensor Fusion on Tilt Interaction in a Mobile Map-Based Application,” no. 11, 2011.

[46] B. Van Tonder and J. Wesson, “IntelliTilt: An enhanced tilt interaction technique for mobile map-based applications,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics),* 2011.

[47] T. Aditya, D. Laksono, H. Sutanta, N. Izzahudin and F. Susanta, “A usability evaluation of a 3D map display for pedestrian navigation,” in *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives,* 2018.

[48] C. Ohm, S. Bienk, M. Kattenbeck, B. Ludwig and M. Müller, “Towards interfaces of mobile pedestrian navigation systems adapted to the user's orientation skills,” *Pervasive and Mobile Computing,* 2016.

[49] W. Dong and H. Liao, “Eye tracking to explore the impacts of photorealistic 3D representations in pedstrian navigation performance,” in *International Archives of*

the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives, 2016.

[50] J. Wither, C. E. Au, R. Rischpater and R. Grzeszczuk, “Moving beyond the map,” in *Proceedings of the 15th international conference on Human-computer interaction with mobile devices and services - MobileHCI '13,* 2013.

[51] I. Delikostidis and C. P. J. M. Van Elzakker, “2 Geo-Identifi cation and Pedestrian Navigation with Geo-Mobile Applications: How Do Users Proceed? Introduction 12.1,” 2009.

[52] T. Vaittinen, M. Salminen and T. Olsson, “City scene,” in *Proceedings of the 15th international conference on Human-computer interaction with mobile devices and services - MobileHCI '13,* 2013.

[53] K. T. Toru Ishikawa, “Relationships between Methods for Presenting Information on Navigation Tools and Users’ Wayfinding Behavior,” 2013.

[54] T. Partala and M. Salminen, “User experience of photorealistic urban pedestrian navigation,” in *Proceedings of the International Working Conference on Advanced Visual Interfaces - AVI '12,* 2012.

[55] C. B. Matthew, G. J. Thomas, R. W. Hubbard and R. P. Francesconi, “MapCube: A Mobile Focus and Context Information Visualization Technique for Geographic Maps,” *Aviation Space and Environmental Medicine,* 2016.

[56] A. El Ali, K. Bachour, W. Heuten and S. Boll, “Technology literacy in poor infrastructure environments,” in *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services - MobileHCI '16,* 2016.

[57] D. Miau and S. Feiner, “Personalized Compass,” in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16,* 2016.

[58] T. Concalves, A. Paula Afonso, M. Biatriz Carmo and P. Rombinho, “Overview "vs" Detail on mobile devices: a struggle for screen space Gonçalves, Afonso, Carmo, Pombinho,” 2012.

[59] R. De A. Maués, E. F. Nakamura and S. D. Barbosa, “Influence of anchor management on anchored navigation in mobile maps,” in *Proceedings - International Computer Software and Applications Conference,* 2012.

[60] V. Elzakker Cpjm, “CO-125 USABILITY TESTING OF A PROTOTYPE MOBILE NAVIGATION INTERFACE FOR PEDESTRIANS”.

[61] M. Iwata, T. Hara, K. Shimatani, T. Mashita, K. Kiyokawa, S. Nishio and H. Takemura, “A location-based content search system considering situations of mobile users,” in *Procedia Computer Science*, 2011.

[62] M. Kamilakis, D. Gavalas and C. Zaroliagis, “Mobile user experience in augmented reality vs. Maps interfaces: A case study in public transportation,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics),* 2016.

[63] A. Mulloni, H. Seichter and D. Schmalstieg, Enhancing Handheld Navigation Systems with Augmented Reality, 2011.

[64] C. Ranasinghe, S. Heitmann, A. Hamzin, M. Pfeiffer and C. Kray, “Pedestrian navigation and GPS deteriorations,” in *Proceedings of the 30th Australian Conference on Computer-Human Interaction - OzCHI '18,* 2018.

[65] T.-S. Lin, H. Gee and S. S. C. Young, 2010 2nd International Conforence on Education Technology and Computer (ICETC), 2010.

[66] I. Delikostidis and V. Elzakker Cpjm, “CO-125 USABILITY TESTING OF A PROTOTYPE MOBILE NAVIGATION INTERFACE FOR PEDESTRIANS”.

[67] D. Wenig, S. Brending, N. Runge and R. Malaka, “Using split screens to combine maps and images for pedestrian navigation,” *Journal of Location Based Services,* 2014.

[68] H. Zhou, A. Edrah, B. MacKay and D. Reilly, *Block party*, 2017.

Appendices

Appendix A

The most probability state of the first search strings

((usability OR ux OR "user experience" OR "user centered design" OR "usage centered design" OR ucd OR "human centered design" OR hcd OR "human computer interaction" OR hci OR "mobile hci" OR "mobile user interfaces" OR "usability engineering") AND ("usability defects" OR "usability issues" OR "usability problems" OR "usability flaws" OR "usability mistakes") AND ("usability evaluation" OR "automated usability evaluation" OR "remote usability evaluation" OR "usability test" OR "usability testing" OR "automated usability test" OR "automated usability testing" OR "remote usability testing" OR "usability inspection" OR "usability heuristics" OR "heuristic evaluation" OR "usability inspection") AND ("mobile web gis" OR "map based mobile applications" OR "mobile map applications" OR MMAs OR "mobile maps" OR "mobile devices" OR "mobile phones" OR "haptic systems" OR "location based services") AND ("usability evaluation method" OR "Automated usability evaluation method" OR "usability testing method" OR "automated usability testing method" OR "usability inspection method" OR "usability heuristics method" OR "heuristic evaluation method" OR "usability inspection method" OR "user study" OR "field study" OR "elicitation study" OR "think aloud" OR "TLX" OR "SUS" OR "USE"))

Appendix B

Extracted data from the first analysis of the first iteration of the review

Title	Author	Year	Type of study	Usability evaluation method	Evaluated application	The stage of the evaluated app	Tested device	Method used in analyzing the results	Usability metrics	measurable criteria	Test Persons (TPs) or Subjects	Number of TPs	Male	Age	TPs number with relevant knowledge	Field /lab	Usability issues	Solutions	Ref.
Overcoming challenges in developing more usable pedestrian navigation systems	Ioannis Delikostidis, Corné P.J.M. van Elzakker, Menno-Jan Kraak	2016	Requirement analysis	Questionnaire, thinking aloud, audio-visual observation, synchronous screen logging/multicamera recording and semi-structured interview	iGo My Way v8.0 and Google Mobile Maps	Real app		-	effectiveness	Task completion	Users	8	-	24-47	8	Field	North-up map	Rotating map	1
			comparative	questionnaires, synchronized video/audio recordings and interview audio recordings, thinking aloud	Google maps and LandNavin	Prototype	Samsung Galaxy S		efficiency	Task completion times, number of stops	Users	24		18-60	16	Field	Continues zooming	Reverse overview + detail	
Comparative Usability Evaluation of Mobile Map Applications	Patarada Thanachan, Arisara Jiamsanguanwong	2016	comparative	Video recording, questionnaire	NOSTRA map and Google maps	Real app	iPhon 5	Mean and Standard deviation, paired-samples t-test, Pareto histogram	Learnability	The time duration to work successfully for the first time	-	5	2	23-35	-	Lab	Words used on interface were misinterpreted by users, users cannot find category and need to search, Icon sub-category was not easily noticeable, cannot save the favorite places, function finding route was complicated, cannot open the list of favorite places, unable to show detail result page, didn't see current location button, cannot chose to Hybrid Map, get lost into Measurement Tools function, the overall problems founded were related to the design of the icons and their location in both apps which inappropriately presented, when participants lost during any tasks in NOSTRA map app, they often selected more menu instead of main menu	redesign of icons and change their location on screen, words use in apps should be minimized and according to individual user, comprehension should be confirmed, the more menu should be integrated into main menu to reduce user confusions	2
									Efficiency	The time duration to work successfully for experienced user									
									Effectiveness	The task success ratio (TSR) by completion ratio multiplies with accuracy ratio of completion ratio (number of application pages/expected number of number of application pages) and accuracy ratio (actual number of click/expected number of click)									
									Memorability	The time duration to work successfully after avoid using system for 5 days									
									satisfaction	Post-Study System Usability Questionnaire (PSSUQ)									

Title	Author	Year	Type of study	Usability evaluation method	Evaluated application	The stage of the evaluated app	Tested device	Method used in analyzing the results	Usability metrics	measurable criteria	Test Persons (TPs) or Subjects	Number of TPs	Male	Age	TPs number with relevant knowledge	Field /lab	Usability issues	Solutions	Ref.
Bridging the Gap Between Field- and Lab-Based User Studies for Location-Based Services	Ioannis Delikostidis, Holger Fritze, Thore Fechner and Christian Kray	2015	Comparative	pre-selection questionnaire, think-Aloud, multi-camera recording system, audio recording	Google Maps	Real app	LG Optimus P990	-	usefulness, ease of use and satisfaction	USE questionnaire	-	18	-	-	-	Field and lab	Stacking in the previous position in Google maps when user has moved to a new position, visualizing a big landmark in reality, but not in Google Maps	-	3
									effectiveness and efficiency	measurement of the performances (video recordings)									
Usability Evaluation of Mobile Passenger Information Systems	Shirley Beul-Leusmann, Christian Samsel, Maximilian Wiederhold	2014	Comparative	think aloud, screen record software, questionnaire, System Usability Scale (SUS), interview	DB Navigator and a prototype	Prototype	iPhone 4s	-	-	Time-on-task	Users	20	10	20-32	17	Lab	Lack of auto-completing in the text fields and lack of overall view, lack of automatically selection of surrounding bus stops, color code, small screen, problem in deleting the texts in the text fields, technical problems of LBS	-	4
												20	10	-	-	field			
Pedestrian navigation with augmented reality, voice and digital map: final results from an in situ field study assessing performance and user experience	Karl Rehrl, Elisabeth Häusler, Sven Leitinger & Daniel Bell	2014	comparative	Santa Barbara Sense-of-Direction Scale (SBSDS) for measure the scale of environmental spatial ability of the participants, system usability scale (SUS) for measure satisfaction, efficiency and effectiveness, voice recording, post-questionnaire to gather further qualitative feedback	Self-implemented mobile app with three interfaces: map interface, voiced-based interface and AR interface	Prototype	Apple's iPhone 4	ANOVA	effectiveness	number of stops (and reasons for stops), GPS accuracy	Users	48	24	21-73	SBSDS test	field	The streets names were partly unreadable because of their wrong adjustment on the map (upside down) that used standard OSM tiles and these map tiles are aligned to the north	for the second iteration, different map tiles for the four cardinal directions were rendered	5
									efficiency	walking time, task completion time, duration of stops									
									satisfaction	NASA TLX							Sensor inaccuracies in AR	Image recognition	
Factors influencing users' employment of mobile map services	Eunil Park, Jay Ohm	2013	-	In-depth interview, pre-survey, questionnaire, online survey	-	Real app	-	Mean, structural equation modeling (SEM), LISREL 8.0	Perceived locational accuracy, Satisfaction, Service & display quality, Perceived mobility, Perceived usefulness, Attitude, Flow state, Intention to use	In-depth interview, TAM, SDQ, PLA	Experts and Users	1109	648	18->60	-	-	Location accuracy, processing speed, display and service quality	-	6

Title	Author	Year	Type of study	Usability evaluation method	Evaluated application	The stage of the evaluated app	Tested device	Method used in analyzing the results	Usability metrics	measurable criteria	Test Persons (TPs) or Subjects	Number of TPs	Male	Age	TPs number with relevant knowledge	Field /lab	Usability issues	Solutions	Ref.
Introducing Usability Heuristics for Mobile Map Applications	Liisa Kuparinen, Johanna Silvennoinen, Hannakaisa Isomäki	2013	Comparative	Post-test questionnaire	NavFree	Real app	iPhone 3GS and Android phone Samsung Galaxy Nexus	-	They measured the suitability of the HE for MMA in comparison to Nielsen's Heuristic	-	experts	4	-	-	-	-	More usability problem was found with the proposed HE for MMA	-	7
A Study of User Perception, Interface Performance, and Actual Usage of Mobile Pedestrian Navigation Aides	James Wen, William S. Helton, and Mark Billinghurst	2013	Comparative	Pre-test and post-test questionnaires, NASA TLX survey, actual usage time, average traversal speed	5 different interfaces	Real app and prototype	iPhone	Post hoc Bonferroni analyses, repeated-measure ANOVA, Greenhouse-Geisser Correction, Bonferroni correction	Ease of use, usefulness, intuitiveness of the interface, goal is obvious in the interface	Perceived usability questionnaire	Users	30	19	19-42	-	field	Problem in directional orientation with simple north-up map	Forward-up map, which shows the direction of the device during the navigation	8
Exploring the use of handheld AR for outdoor navigation	Andreas Dunser, MarkBillinghurst, JamesWen, VilleLehtinen, AnttiNurminen	2012	Comparative	Pre-test questionnaire, NASA TLX, interview, Post-task questionnaire, video recording, user comments, experimenter observations	3 difference interfaces	Prototype	HTC Desire	Friedman test	-	-	Users	22	11	19-47	-	-	GPS accuracy and compass input, shakiness, screen brightness, recognition of dead ends routs in AR interface (lack of overview)	-	9
Lost in Navigation: Evaluating a Mobile Map App for a Fair	Anders Bouwer, Frank Nack, Abdallah El Ali	2012	Evaluation	Pre-questionnaire, video recording, post-test questionnaire	Indoor mobile map app	Real app	HTC Desire HD	Median and inter-quartile ranges	Perceived usefulness of functionality, acceptance, usefulness and helpfulness and ease of use	Post-study questionnaire	Users	14	9	20-54	-	field	Map orientation	Lost in Navigation: Evaluating a Mobile Map App for a Fair	10
A mobile 3D-GIS hybrid recommender system for tourism	José M. Noguera, Manuel J. Barranco, Rafael J. Segura, Luis Martínez	2012	Comparative	Usability questionnaire (7-point Likert scale), subjective rating usability test	-	Prototype	iPhone	-	Easiness and efficiency	questionnaire	Users	27	19	24-48	-	Field and lab	Possibility to switch from the 3D to the 2D interface and vice versa, Integration with social networks	-	11
Collaborative Map Exploration Using Multitouch Surfaces	Pedro G. Villanueva, Ricardo Tesoriero, and María D. Lozano	2012	Comparative	Questionnaire	Collaborative Map Explorer	Prototype	-	-	Effectiveness Productivity user satisfaction	effectiveness, task completion, and error frequency task time and task efficiency questionnaire	-	10	7	10-20	-	-	-	-	12

Title	Author	Year	Type of study	Usability evaluation method	Evaluated application	The stage of the evaluated app	Tested device	Method used in analyzing the results	Usability metrics	measurable criteria	Test Persons (TPs) or Subjects	Number of TPs	Male	Age	TPs number with relevant knowledge	Field /lab	Usability issues	Solutions	Ref.
Usability Evaluation of a Map-Based Multi-Publishing Service	Hanna-Marika Flink, Juha Oksanen, Ulla Pyysalo, Mikko Rönneberg and L. Tiina Sarjakoski	2011	Qualitative	Pre-test and post-test questionnaires, think aloud, SUS	MenoMaps	Prototype	iPhone	affinity diagram	Satisfaction	Post questionnaire with SUS	Users	6	3	32-58	-	Lab	Interpreting some background maps were difficult Doesn't have search function	Legend can be suggested Search field or choose from a list	13
A Comparison of Communicative Modes for Map-Based Tasking	Eli R. Hooten, Sean T. Hayes, and Julie A. Adams	2011	Comparative	Scenario completion times, subjective ratings using Likert scale questionnaires and NASA TLX subjective workload, subjective comparation	MMBTI	Real app	Dell XT2 tablet laptop running Windows 7	Stepdown Bonferroni tests, T-test, Wilcoxon signed-rank test	Effectiveness Overall preference, task comprehension and ease of use	Scenario completion time Subjective ranking questionnaire	Users	8	6	18-26	-	Lab	Missing the overall understanding of the area because of small screen size	Paper maps	14
Visualizing references to off-screen content on mobile devices: A comparison of Arrows, Wedge, and Overview + Detail	Stefano Burigat, Luca Chittaro	2011	comparative	demographic questionnaire, automatically logged task completion times and error rates, subjective preference	Three visualization techniques for off-screen objects	Real app	Asus P535 Windows Mobile phone	Shapiro–Wilk test and ANOVA, Tukey's post-hoc test, Shapiro-Wilk test and ANOVA-Type Statistic (ATS), Dunn's post-hoc test, non-parametric ATS	effectiveness Ease of use	Marking off-screen object on a printed version of the visualization Low error rates	Users	24	9	20-59	9	lab	When users zoom to a specific area, at the same time lose the overview of the space information	Three visualization techniques for off-screen objects	15
Tilt and go: exploring multimodal mobile maps in the field	Andrew Ramsay, Marilyn McGee-Lennon, Graham A, Wilson Steven J. Gray, Philip Gray, François De Turenne	2010	Comparative	Pre-test questionnaire, audio and video recording, post-hoc interview	-	-	Nokia N95-2	-	-	-	Users	18	16	17-48		Field	Significant delay during scrolling while new map tiles were downloaded from the remote server	Catching the tiles covering the area	16
Semi-Automatic Zooming for Mobile Map Navigation	Sven Kratz, Ivo Brodien, Michael Rohs	2010	Comparative	NASA TLX, task completion time and USE	slippy map	Prototype	iPhone 3GS	-	Satisfaction, learnability, usefulness and ease of use	USE questionnaire	-	13	8	24-28	-	-	unwanted loss of control in SDAZ zooming technique, for simple zoom interface: occlusion, slowness, sticky fingers problem	One hand interaction, quick zooming	17

Title	Author	Year	Type of study	Usability evaluation method	Evaluated application	The stage of the evaluated app	Tested device	Method used in analyzing the results	Usability metrics	measurable criteria	Test Persons (TPs) or Subjects	Number of TPs	Male	Age	TPs number with relevant knowledge	Field /lab	Usability issues	Solutions	Ref.
COMPARISON OF EVALUATION METHODS FOR FIELD-BASED USABILITY STUDIES OF MOBILE MAP APPLICATIONS	BURGHARDT D., WIRTH K.	2010	Evaluate the Garmin GPS (not mobile phone)	Video recording, Likert scale and open questions	Garmin receiver	-	-	t-test	-	-	-	18	-	12-17 18-60 60<	-	-	Some errors happened according to age and applied evaluation method	-	18
Is Tilt Interaction Better Than Keypad Interaction for Mobile Map-based Applications?	Bradley van Tonder, Janet Wesson	2010	Comparation of tilt and keypad interaction	Post-task questionnaire (based on NASA-TLX), post-test questionnaire	MapExplorer	Prototype	Nokia N97	-	Satisfaction, Perceived Efficiency, Effectiveness, controllability and ease of use Perceived workload, Preferred interaction and overall impressions of the two interaction techniques in terms of perceived effectiveness, efficiency, controllability and ease of use	Post-task questionnaire Post-test overall rating-seven point semantic differential scale (Tilt = 1 and Keypad = 7)	Users	32	23	18-29	20	lab	In touch-screen interaction the display can be obscured by user's hand	Tilt interaction	19
USABILITY TESTING OF A PROTOTYPE MOBILE NAVIGATION INTERFACE FOR PEDESTRIANS	Ioannis Delikostidis, Corné P.J.M. van Elzakker	2010	comparative	pre-test questionnaire, thinking aloud (for user satisfaction), synchronized video and audio recording and audio recording of the post-session interviews and users' task performance	Google maps and LandNavin (a prototype)	Prototype	Samsung Galaxy S	Atlas.ti (qualitative research software) for verbatim transcription	overall efficiency and effectiveness and the satisfaction	Task completion times and the success and error rates	Users	24	10	18-40	16	field	lack of automatic rotation of the North-up map in Google Maps, GPS and compass inaccuracies and icon overlapping in particular zoom levels	compass-based heading-up (rotated) map, landmark pop-up information, multi-perceptive photos and landmark symbology, vertical scale bar with the combination of distance and time needed, landmark filtering and dual map	20

Title	Author	Year	Type of study	Usability evaluation method	Evaluated application	The stage of the evaluated app	Tested device	Method used in analyzing the results	Usability metrics	measurable criteria	Test Persons (TPs) or Subjects	Number of TPs	Male	Age	TPs number with relevant knowledge	Field /lab	Usability issues	Solutions	Ref.
Use, User, and Usability Research for Mobile Geo-Applications for Personal Orientation and Navigation	Corné P.J.M. van Elzakker, Ioannis Delikostidis	2010	Requirement analysis	A combination of questionnaires, observation, thinking aloud, audio / video recording, screen logging and semi-structured interviews	Different apps used by user in the requirement analysis stage	Real apps	Android-base	-	Effectiveness and efficiency	-	Users	18	-	-	-	field	-	Continues and accurate automatic rotation of mobile map, simplicity of the map that should not be overloaded with many symbols or 3D landmarks, landmark photos that pop-up when clicking them are more preferable, by presenting landmarks in successive scales, avoiding frequent zooming and panning, the spatial information on the map should be represented in a way that users spend more time to observe surrounding to develop mental maps than looking at the mobile map, more choices for pedestrians that means they should be left free to select any possible route (flow channels) that lead to the destination, automatic landmark recognition with using integrated digital camera with GPS position and heading information with landmark visibility map data on image recognition algorithm, a technique opposite of overview+detail that represent the overview on the full screen and detail view on the thumbnail	21
Navigation techniques for small-screen devices: an evaluation on maps	Stefano Burigat, Luca Chittaro and Silvia Gabrielli	2008	Comparative	automatically logged task completion times, user interface actions and accuracy, SpatialMemory Task, subjective	-	Real app	624Mhz Pocket PC with a 3.5" display	Kolmogorov-Smirnov test of normality, ANOVA, Tukey post-hoc test, Friedman's	Performance and satisfaction	-	Users	20	12	21-39	-	lab	Occupying the screen with hand or stylus	DoubleScrollBar Technique	22

and web pages				preference, user workload				test, Dunn's Multiple Comparison post-hoc test, T-test, Friedman's test									Users' orientation	ZEN (Overview and Detail) technique	
Title	**Author**	**Year**	**Type of study**	**Usability evaluation method**	**Evaluated application**	**The stage of the evaluated app**	**Tested device**	**Method used in analyzing the results**	**Usability metrics**	**measurable criteria**	**Test Persons (TPs) or Subjects**	**Number of TPs**	**Male**	**Age**	**TPs number with relevant knowledge**	**Field /lab**	**Usability issues**	**Solutions**	**Ref.**
Map, Diagram, and Web Page Navigation on Mobile Devices: the Effectiveness of Zoomable User Interfaces with Overviews	Stefano Burigat, Luca Chittaro, Edoardo Parlato	2008	comparative	Interview, logged code	Web pages, Diagrams and Maps	Prototype	Windows mobile 5.0 PDA phone	Kolmogorov-Smirnov test of normality, ANOVA, Tukey post-hoc test and Friedman's test	Effectiveness	task completion times and user interface actions	Users	24	13	16-37	10	lab	Users lose the overview when use zooming during navigation	Overview&detail interfaces	23
									User preference	Preference analysis									
Field-Based Usability Evaluation Methodology for Mobile Geo-Applications	Corné P.J.M. van Elzakker, Ioannis Delikostidis, Peter J. M. van Oosterom	2008	Propose a methodology for usability evaluation with Comparation of different usability testing methods	observation, thinking aloud, video/audio recording (screen logging) and semi-structured interview	iGO My way 2006	Real app	HP iPAQ hx4700 PDA	-	Efficiency	Time consumed for each task	Users	18	12	25-40	-	field	Absence of properly placed landmarks, inability of the mobile map to be oriented toward the actual view point of the user, initial misunderstanding of users' location	-	24
									Effectiveness	The percentage of the successful completion of each task									
									User satisfaction	post-survey semi-structured interview									
Framy – visualising geographic data on mobile interfaces	Luca Paolino, Monica Sebillo, Genoveffa Tortora & Giuliana Vitiello	2008	comparative	Think aloud, questionnaire,	MapGIS	Prototype	-	SYSTAT (ver.12), t-test	Quantitatively measure efficacy and efficiency	Time requested to complete the task (TRC), The number of steps to complete a task (NSC)	Users	20	-	-	-	lab	Multipole zooming and panning	Proposing a new method for off-screen visualization (Framy)	25
									satisfaction	Post-questionnaire with Likert scales									

Appendix C

Extracted data for analysing the second round of iteration of the review

Title	Author	Year	Usability evaluation method(s)	Evaluated application(s)	The stage of the evaluated app	Apparatus	Usability metrics	measurable criteria	Test Persons (TPs)	Number of TPs	Male	Age	Number of TPs with relevant knowledge	Field /lab	Usability issues	Solution(s)	Ref.
CaMap: Camera-based Map Manipulation on Mobile	Liang Chen, Dongyi Chen	2018	Pre and post-test questionnaire	CaMap	Prototype	4.3 inch	Easy to use, efficiency and effective	Subjective questionnaire	Users	8	4	23.4	8	Lab	the button's position made it less convenient to trigger the zoom mode, the button's position should be designed more properly, less usable results in panning operation	No solution	1
															Automatic zooming without user's intention	Using valume-up button in zoom operation and valume-down for click operation	
A USABILITY EVALUATION OF A 3D MAP DISPLAY FOR PEDESTRIAN NAVIGATION	Trias Aditya, Dany Laksono, Heri Sutanta, Nur Izzahudin, Febrian Susanta	2018	Questionnaire, video recording with logged app	A 3D map and Google Maps	-	-	5Es (effective, efficient, engaging, error tolerant, and easy to Learn)	Subjective questionnaire	Users	16	10	-	-	Field	Availability and quality of pedestrian navigation lines available in 3D map, accuracy and correctness, they have difficulties to differentiate individual buildings in 3D map	-	2
Maps, vibration or gaze? Comparison of novel navigation assistance in indoor and outdoor environments	Charalampos Gkonos, Ioannis Giannopoulos and Martin Raubal	2017	Santa Barbara sense of direction scale (SBSODS), pre-questionnaire for self-reporting of their spatial abilities user experience questionnaire (UEQ)	-	Prototype	Samsung Galaxy Nexus	Navigation performance	Completion time and errors	Users	10	7	29	-	Field	Too much engagement of user with map interface and distracting from the environment	Navigation with assistant of vibration and gaze	3
							Cognitive workload	NASA TLX questionnaire									
							(attractiveness, perspicuity, efficiency, dependability , stimulation, novelty)	User Experience Questionnaire (UEQ)									
							Cognitive workload	raw NASA TLX questionnaire									
Follow the Signs—Countering Disengagement from the Real World During City Exploration	Markus Konkol, Christian Kray and Morin Ostkamp	2017	Pre-test questionnaire, Semi-structured interviews, think aloud	-	Prototype	Samsung Galaxy S4 Mini (4.3")	-	-	Users and experts	9	1	34	7	Field	Mobile maps: High cognitive load, disengagement from the surrounding environment. Prototype: GPS inaccuracy, late appearance of the signage pop up message, the wrong direction provided by the compass and arrow	When user reach a territory of a signage or landmark in his route a pop up containing a photograph of sign and a message appear to guide him, using arrow point directly to the target	4

Title	Author	Year	Usability evaluation method(s)	Evaluated application(s)	The stage of the evaluated app	Apparatus	Usability metrics	measurable criteria	Test Persons (TPs)	Number of TPs	Male	Age	Number of TPs with relevant knowledge	Field /lab	Usability issues	Solution(s)	Ref.
															Google maps: need to enter addresses and zooming in and out Overall issue: localization and orientation inaccuracies		
Towards interfaces of mobile pedestrian navigation systems adapted to the user's orientation skills	Christina Ohm, Stefan Bienk, Markus Kattenbeck, Bernd Ludwig, Manuel Müller	2016	Self-report SoD-questionnaires, questionnaire	-	Prototype	-	-	-	Users	112	58	23.4	-	Field	-	The navigation system should be adaptable to users' sense of direction, abstract interface with highlighted salient objects	5
Mobile User Experience in Augmented Reality vs. Maps Interfaces: A Case Study in Public Transportation	Manousos Kamilakis, Damianos Gavalas, Christos Zaroliagis	2016	Questionnaire, logged data	-	-	-	Ease of use	Questionnaire	Users	22	10	31.86	-	Field	Interacting with (tapping on) the markers in AR was more difficult than 2D map	-	6
MapCube: A Mobile Focus and Context Information Visualization Technique for Geographic Maps	Bjorn Werkmann, Matthias Hemmje	2016	Experiment observations	-	Prototype	-	effectiveness efficiency	Error rate Time-on-task	Users	10	-	-	-	Lab	The problem of off-screen object view	Proposing an information visualization technique for showing simultaneously focus and context information on the map	7
EYE TRACKING TO EXPLORE THE IMPACTS OF PHOTOREALISTIC 3D REPRESENTATIONS IN PEDSTRIAN NAVIGATION PERFORMANCE	Weihua Dong, Hua Liao	2016	Pre-test questionnaire, Combination of eye tracking (video and audio of TPs faces were recorded too) and think aloud	Google Street View and Google Maps	Real app	4 inch	Effectiveness, efficiency and cognitive workload	Time-on-task	Users	20	6	21	-	Lab	3D map representation needs more attention of users with higher workload than 2D map representation	Combine 2D and 3D methods for map representation, in 3D representations only the most important information should show to decrease the information density, in 2D maps, important landmarks should be included to help users locate and orient themselves	8
The utility of Magic Lens interfaces on handheld devices for touristic map navigation	Jens Grubert, Michel Pahud, Raphael Grasset, Dieter Schmalstieg, Hartmut Seichter	2015	semi-structured interviews, questionnaire, experiment observations	-	-	-	Effectiveness, efficiency	Task completion time, Error rate	Users	18	12	48.76	-		Zooming issue in small screens that cause losing overview	Simultaneously zooming and panning, support one handed spatial navigation	9
GazeNav: Gaze-Based Pedestrian Navigation	Ioannis Giannopoulos, Peter Kiefer, Martin Raubal	2015	Wizard of Oz, pre-test questionnaire (spatial ability), Questionnaire	-	-	-	Efficiency Effectiveness	Task completion time Reaching the target successfully (interruptions)	Users	32	19	31.97	-	Lab	Map interface requires visual attention (switching the users' attention several times to the navigation device) and high spatial abilities	Independent direct mobile interaction (only interact with the environment that lead users to acquire significantly better local spatial knowledge)	10
Using split screens to combine maps and images for pedestrian navigation	Dirk Wenig, Stefan Brending, Nina Runge and Rainer Malaka	2014	think aloud, SUS questionnaire (UTAUT), interview, logged data	OpenStreetMap data	Prototype	-	-	-	Users	16	6	17-54	-	Field	Street names were not always readable because of the rotating map, the choice of the route must be questioned	One hand interaction, reduce the interaction with split the screen with map and image that can be seen simultaneously	11
The influence of gaze history visualization on map interaction sequences and cognitive maps	Ioannis Giannopoulos, Peter Kiefer, Martin Raubal	2013	Santa Barbara Sense of Direction Scale, experiment observations	OpenStreetMaps	-	4.6 inch	Effectiveness, efficiency	Task completion time	Users	40	-	27.9	-	Lab	The problem of continues zooming and panning and poor spatial knowledge of users because of the effect of current map-based interactions	Reducing panning interaction significantly by GeoGazemarks concept	12
Relationships between Methods for Presenting Information on Navigation Tools and Users' Wayfinding Behavior	Toru Ishikawa, Kazunori Takahashi	2013	Santa Barbara Sense of Direction Scale, Questionnaire, experiment observations	Using Google maps API Using Google maps API	Prototype Prototype	3.5 inch 3.5 inch	Memory of scenes Time spent looking at the tool	Scene recognition task Self-reported in questionnaire	Users Users	24 24	12 12	22.3 21.8	- -	Field Field	Distracting users' attention from the environment to the tool, poor remembering of	Using paper map	13

Title	Author	Year	Usability evaluation method(s)	Evaluated application(s)	The stage of the evaluated app	Apparatus	Usability metrics	measurable criteria	Test Persons (TPs)	Number of TPs	Male	Age	Number of TPs with relevant knowledge	Field /lab	Usability issues	Solution(s)	Ref.
															surrounding scenes, the tools require users to follow directions without knowledge where they are heading or having a mental picture of the whole route, lack of positioning users' current location on the map		
Moving Beyond the Map: Automated Landmark Based Pedestrian Guidance Using Street Level Panoramas	Jason Wither, Carmen E. Au, Ray Rischpater, Radek Grzeszczuk	2013	Questionnaire, interview	City Scene	Real app	Nokia N9	efficiency	Number of errors and task completion time	Users	8	4	20-50	-	Field	Visual complexity in panorama interfaces Requires more user attention	-	14
															discrete routes in panorama interfaces opposite to map interfaces that show the entire route	Switch between two interface modes Combining both modes	
City Scene: Field Trial of a Mobile Street-imagery-based Navigation Service	Tuomas Vaittinen, Miikka Salminen, Thomas Olsson	2013	Questionnaire, interview, logged data	City Scene	Real app	Nokia N9	-	-	Users	10	6	29.2	-	Field	The images in interface were not up to date, determining the right direction when starting the navigation was reported that induce users to walk and look where the GPS pointer is moving and tapping the buttons for moving between the waypoints while walking was also reported to be cumbersome, the long time of downloading the panoramas	When GPS avatar on the map mode didn't move as expected, the panorama view might be helpful, using panorama view in recognizing the destination Simplicity of the interface is very important The system's warning should be simpler to let the users pay more attention to the environment and enjoy their visit The compass- and map-based view should be prioritized while images are downloading Seamless switching between simple and visually reach views	15

Title	Author	Year	Usability evaluation method(s)	Evaluated application(s)	The stage of the evaluated app	Apparatus	Usability metrics	measurable criteria	Test Persons (TPs)	Number of TPs	Male	Age	Number of TPs with relevant knowledge	Field /lab	Usability issues	Solution(s)	Ref.
User-Center Design of Mobile Geo-Applications	Corné P.J.M. van Elzakker, Ioannis Delikostidis	2012	Questionnaire, observation, thinking aloud, audio/video record, interview	iGo My way, Google Maps	Real app	-	-	-	Users	26	-	-	-	Field	Maps are overcrowded with too much information	The map should be simple (following color coding and should not be overloaded with many symbols or 3D buildings), the street sizes and patterns should properly reflect these parameters of reality, landmark photos that pop up when clicking them are more preferable than 3D models, preserving landmark visibility in successive scales, using an algorithm to calculate landmark visibilities in any point of the users' possible route on the map, presenting a North-up map at the beginning of any navigation task that may be transformed automatically by animation into heading-up maps during navigation whilst maintaining the connection between the two views, providing the user with a tool for automatic recognition-identification of landmarks through augmented reality	16
															Overview and detail	Paper maps	
User Experience of Photorealistic Urban Pedestrian Navigation	Timo Partala, Miikka Salminen	2012	Questionnaire, Latin squares method (for counterbalancing the experiment), AttrakDiff Questionnaire (pragmatic quality, hedonic – identification, hedonic stimulation and attractiveness), NASA TLX questionnaire for measuring task load, subjective preference questionnaire	Fonecta maps, Google Maps	Real and prototype	Nokia N97	Location and direction knowledge scales questionnaire	Measure how well the used map visualization supported the participants' awareness of their current position	Users	9	4	24.1	-	Field	head-down interaction	-	17
							Current walking direction	-									
							Identification of buildings scale	how easy it was for the participants to make associations between the buildings and landmarks on the map and the corresponding real-world objects									
							Time spent looking at the map	-									
Improving the controllability of tilt interaction for mobile map-based applications	Bradley Paul van Tonder, Janet Louise Wesson	2012	Logged data, NASA TLX questionnaire	MapExplorer	Prototype	Samsung Google Nexus S	Efficiency	Task completion time	Users	30	21	20-32	30	Lab	The problems of keypad and touch-screen interaction techniques such as less or no control over panning speed mostly for long distance panning, both hands are engaged in interaction, occlusion of the display, controllability relating to zooming and panning operations while walking	Tilt zooming (use of tilting gesture to control zooming speed), adapting the sensitivity of tilt interaction according to the users' current context (seated or walking)	18
							Perceived workload	NASA TLX questionnaire (mental demand, physical demand, temporal demand, performance, effort and frustration)									
							User satisfaction, perceived effectiveness, efficiency, ease of use and controllability	Questionnaire									
		2011	Logged data, questionnaire	MapExplorer	Prototype		Efficiency	Task times	Users	17	11	24	-	Lab			19

Title	Author	Year	Usability evaluation method(s)	Evaluated application(s)	The stage of the evaluated app	Apparatus	Usability metrics	measurable criteria	Test Persons (TPs)	Number of TPs	Male	Age	Number of TPs with relevant knowledge	Field /lab	Usability issues	Solution(s)	Ref.
The Impact of Sensor Fusion on Tilt Interaction in a Mobile Map-Based Application	Bradley van Tonder, Janet Wesson					Samsung Google Nexus S	controllability	For location tasks, the number of times the cursor entered and left the target POI region and the time between POI region entry and selection, for navigation tasks, planned routes and the number of waypoints missed							Position of the zooming button was difficult to reach for one of the TPs	Zoom independent of panning operation with tilting, change the zoom level more than one level at a time, providing selection operation with a cursor which with using vibration noticing user and with standard Android back button returning to the map display again, smoother interaction	
							Perceived controllability	questionnaire									
IntelliTilt: An Enhanced Tilt Interaction Technique for Mobile Map-Based Applications	Bradley van Tonder, Janet Wesson	2011	Post-test questionnaire	MapExplorer	Prototype	Nokia N97	Perceived workload	Post-task questionnaire	Users	16	11	20-29	-	Lab	In touch-screen interaction the display can be obscured by user's hand, controllability, zooming operation sometimes accidentally is going out of user's control, mental demand, sensitivity, practicality	Allow users browse maps in a wide geographic area at a wide range of zoom levels	20
							User satisfaction, perceived efficiency, effectiveness, ease of use, controllability	Standard After-Scenario Questionnaire (ASQ)									
							performance	Task times									
Enhancing Handheld Navigation Systems with Augmented Reality	Alessandro Mulloni, Hartmut Seichter, Dieter Schmalstieg	2011	Video recording, semi-structured interview (subjective feedback), logged data	-	-	-	-	-	Users	9	-	28.1	0	Field	Users need to exploit multiple interfaces, photos do not match the appearance of environment due to its variability and they are rarely taken from the exact users' position	Multimodal navigation system with using AR and audio and vibration instructions, tracking accuracy must be communicated by the visualization	21
A Location-based Content Search System Considering Situations of Mobile Users	Mayu IWATA, Takahiro HARA, Kentaro SHIMATANI, Tomohiro MASHITA, Kiyoshi KIYOKAWA	2011	Questionnaire, logged data	-	Prototype	iPhone	-	-	Users	11	8	-	-	Field	Difficult map interaction	Easy with few operations, Easy to grasp information, Adapting to users' situations	22
Integration of Cognition-based Content Zooming and Progressive Visualization for Mobile-based Navigation	Yik Kong Cheung, Zhilin Li and Wu Chen	2009	Questionnaire	-	-	-	-	-	Users	20	-	20-40	-	Lab	Too many map levels	Map matching address	23
															Refresh of screen	Progressive map visualization	
															Map tiles load problems	better to load the centred map tile first before the other for better visualization	
															Reducing the number of zoom operation	Content zooming (analogous to textual address)	
Geo-Identifi cation and Pedestrian Navigation with Geo-Mobile Applications: How Do Users Proceed?	Ioannis Delikostidis and Corné P. J. M. van Elzakker	2009	Questionnaire, thinking aloud with audio/visual observation with synchronous screen logging, semi-structure interview	iGo My way v. 8.0, Google Maps	Real app	PDA HP iPAQ 4700hx, PDA-smartphone i-mate Ultimate 9502	-	-	Users	8	5	24-47	8	Field	Google Maps: missing important landmarks on the map displays iGo: screen was overloaded with too many 3D buildings that made interaction slow and difficult, applying same color for different buildings on map made distinguishing difficult, the software disability of fast	The landmarks should be made more distinguishable (colur, shape, size) or additional information should be provided (photoes, text), using pop up photoes instead of 3D repreentation for landmarks, properly representation of accurate streets' size and pattern visualization according to reality, representing pedestrian paths as an important landmark on the map, the users' orientation shouln't rely merely on dirrection of position arrow on the map since according to the speed of walking is not accurate (it should take advantage of landmarks), the map	24

Title	Author	Year	Usability evaluation method(s)	Evaluated application(s)	The stage of the evaluated app	Apparatus	Usability metrics	measurable criteria	Test Persons (TPs)	Number of TPs	Male	Age	Number of TPs with relevant knowledge	Field /lab	Usability issues	Solution(s)	Ref.
															processing of geographic data in a common mobile device in order to achieve graphically smooth changes during zoom operation	should rotate toward the direction of user even when he/she is not moving,	

Appendix D

Extracted data for analysing the third round of iteration of the review

Title	Author	Year	Usability evaluation method(s)	Evaluated application(s)	The stage of the evaluated app	Test Persons (TPs)	Number of TPs	Male	Age	Number of TPs with relevant knowledge	Field /lab	Usability issues	Solution(s)	Ref.
Pedestrian Navigation and GPS Deteriorations: User Behavior and Adaptation Strategies	Champika Ranasinghe, Sven Heitmann, Albert Hamzin, Max Pfeiffer and Christian Kray	2018	Questionnaire, santa barbara sense of direction scale, semi-structured interview, video and audio recording, logged data	-	prototype	Users	21	12	22-38	-	Field	Weak offline features of Google maps	Map layers and photographs would be downloaded for offline usage	1
												GPS deteriorations	Using prominent landmarks, combining geometrical and basic thematic information with photographs of real world entities (could be downloaded for offline usage), users could be able to chose the granularity of information according to their situation, system should notify user about the prombel, different visualizations for diffetent types of deteriorations (e.g. red dot for no signal)	
Block Party: Synchronized Planning and Navigation Views for Neighbourhood Expeditions	Huiyuan Zhou , Aisha Edrah , Bonnie MacKay , Derek Reilly	2017	Santa Barbara Sense of Direction Questionnaire (SBSOD), video recording, logged data, observation, interview	Black party	prototype	Users	10	6	18-35 1>50	-	Field	the ability to divert from the recommended route, and to personalize routes, unable to save or search a POI during navigation task without going back to the map view, destinations didn't match expectations (e.g., a museum looked like an ordinary house), in Google Maps same data (e.g. address) is often available in one view (e.g. map) but not in another (e.g. navigation), GPS updates and accuracy	easy switching among distinct views that support diverse tasks, synchronized data/operations among views, a directory-based list of POIs, explicit support of planning, creation, and modification of multi-point itineraries, Map, List and Immersive views toggled with the bottom menu to facilitate easy switch among different ways to view the data, selecting a category to show the POIs of that type show on the map view, saving POIs during the route without abandon of the route	2
							10	5	18-38	1	Field			
Personalized Compass: A Compact Visualization for Direction and Location	Daniel Miau, Steven Feiner	2016	Error rates and task completion time	-	prototype	Users	26	13	20-39	25	Lab	Off-screen objects Weaknesses of 'Wedge' technique Losing overview Too much zooming and panning operations Screen occlusion of information	Replace the compass with Personalized compass (P-Compass) A combination of Wedge with P-compass (using P-compass for distant off-screen POIs and using Wedge for nearby off-screen POIs) since they are complementary Automatic switching between Wedge and P-Compass according to the distance of POIs and with the control of users	3

Title	Author	Year	Usability evaluation method(s)	Evaluated application(s)	The stage of the evaluated app	Test Persons (TPs)	Number of TPs	Male	Age	Number of TPs with relevant knowledge	Field /lab	Usability issues	Solution(s)	Ref.
Technology Literacy in Poor Infrastructure Environments: Characterizing Wayfinding Strategies in Lebanon	Abdallah El Ali, Khaled Bachour, Wilko Heuten and Susanne Boll	2016	Interview and web survey	Google Maps	Real app	Users	12	7	18-35	2	Lab	outdated mapping of locations on digital maps, GPS inaccuracy, incorrect route plans, smartphone battery life, poor accuracy of using Google Maps, the technology aid in question did not keep up with the naming conventions used by people, the language used to describe directions was highly imprecise among people, some street names were not referred to by their official names, direction giving strategy is irrelevant to technologically literacy of user, problems of navigation in rural area, inaccurate position marker on the map, incorrect or missing place listings on the map, poor network connectivity, incorrect bus route plans, maps only showing partial information (e.g. main roads)	-	4
							85	56	17-74	-	-			
Overview "vs" Detail on mobile devices: a struggle for screen space	Concalves, Tiago Paula Afonso, Ana Biatriz Carmo, Maria Rombinho, Paula	2012	Interview, logged data	-	-	Users	30	20	18-53	27	Lab	For overview & detail technique: greater physical and mental effort, reducing the available space and some information of the detail might be hidden behind the overview, the small size of the overview	Resizable overview thumbnail for Overview&Detail technique	5
Influence of Anchor Management on Anchored Navigation in Mobile Maps	Rodrigo de A. Maués, Eduardo F. Nakamura and Simone D. J. Barbosa	2012	Questionnaire, logged data, think aloud and interview	-	-	Users	36	24	21	-	Lab	The general issue of switch between zooming and panning	Facilitate the acquisition of off-screen POI by reducing mode-switching (switching between panning and zooming operations)	6
User expectations of the design of a map-based mobile guide system for public arts	Ting-sheng Lin, Hubert Gee, Shelley S. C. Young	2010	Questionnaire	OhMyArt	prototype	Users	87	47	-	36	Lab	-	A simple map representation with GPS functionality based on realistic landmark rather than abstract form of representation (iconic symbols or words)	7