2015 48th Hawaii International Conference on System Sciences

Enhanced Human-Computer Interaction for Business Applications on Mobile Devices: A Design-Oriented Development of a Usability Evaluation Questionnaire

Maximilian Wich
University of Mannheim
mwich@mail.uni-mannheim.de

Tommi Kramer
University of Mannheim
kramer@uni-mannheim.de

***Abstract*—Over the last few years mobile technology has gained enormous growth in the field of human-computer interaction as it became an essential part of our society's every-day life. Specific mobile characteristics such as availability, computational power, or high-resolution displays make these devices very useful, even in a business environment. However, usability plays a crucial role when it comes to software design for human-computer interaction with mobile devices. In order to achieve a high level of usability in user interfaces of application software, it is essential to ensure usability quality during the development process. Therefore, we have developed a specific questionnaire for evaluating the usability of mobile business apps as well as a corresponding web-based software tool for simplifying the assessment. In this paper, we follow a design science research approach and evaluate our designed artifacts in expert interviews where we demonstrate the utility and applicability of the questionnaire and tool.**

***Keywords*-Usability Engineering, Human-Computer Interaction, Mobile, Design Science**

I. Introduction

When Steve Jobs introduced Apple's first generation of iPhones in 2007, it was impossible to imagine what impact this innovation would have on daily life in society. Six years later, smartphones are every-day objects and it is hard to think of a life without them. From 2010 to 2012 the sales figures increased from 305 million devices to 722.5 million and for 2017 1,733.9 million sold smartphones are forecasted by IDC [1]. However, such an increase will not remain without consequences for desktop computers and laptops. Considering the new generation of mobile and smart devices, 87% of all connected devices (desktop computers, laptops, tablets, and smartphones) sold in 2017 will be tablets or smartphones [2].

The huge success of the new technology can be mainly ascribed to the devices' high degree of usability and user friendliness [3]. Seeing that comprehensive functionality and user-friendly human-computer interaction are no contradiction , people began to use mobile devices more and more extensively in their everyday life. Such an enormous trend had an impact on the development process of the products as well. To reach a high level of usability it is inevitable to involve future users as soon as possible in the development process. Additionally, the user-centred design approach puts user needs and goals in the focus of the engineering process and enables us to develop satisfying applications for the user [4]. An essential determinant in that process is the regular feedback of future users. In order to make this involvement efficient it is necessary to have valuable evaluation methods. For mobile applications and especially for mobile business applications, there is a lack of proven methods supporting the usability engineering process in order to improve the human-computer interaction.

A widely used method for evaluating usability of information systems is the satisfaction questionnaire, which is filled in by prospective users. Over the course of time a wide range of such questionnaires are developed, for instance the System Usability Scale [5] or the Usability Metric for User Experience by [6]. However, these methods assess the usability of a system in a generic way and do not focus on particular components of a user interface, like typography, aesthetic, layout, or terminology. Consequently, it is difficult to draw inferences from the result of the evaluation to the components of the user interface that have to be improved. In addition, they are designed for information systems in general and not specifically for mobile applications. Hence, the objective of the paper is to develop an evaluation questionnaire for assessing single components of a mobile user interface regarding their usability in order to meet the requirements of the current mobile business technology. Such a specific questionnaire helps to identify usability problems faster. Moreover, it will improve the user experience during the software development process.

This research is embedded in the design science research methodology according to [7]. Section 2 provides the theoretical background of usability evaluation questionnaires by revealing characteristics of mobile devices and mobile business applications, existing usability engineering processes, as well as state-of-the-art evaluation methods in this domain. The requirements for a user-friendly design of mobile business applications are derived from usability theories and user interface guidelines in section 3. After assembling the requirements, section 4 describes the development of the questionnaire and the corresponding software tool based on these insights. In section 5, the results of the empirical evaluation of the questionnaire and the software tool are

1530-1605/15 $31.00 © 2015 IEEE
DOI 10.1109/HICSS.2015.63

presented and discussed. Section 6 concludes the paper with a summary of the results and future work prospects. Hence, the structure of this paper addresses the rigorous development and evaluation of two artifacts (the questionnaire and the corresponding software tool) according to the principles of design science. The evaluation is conducted with user interface experts.

II. Literature Review

The following section will provide a definition of mobile devices and mobile business applications in order to get a better understanding of the unit of analysis. Additionally, the existing usability engineering process and related work are presented and discussed.

A. Mobile Technology

In the context of the paper "mobile devices are those devices that are used to connect to mobile services" [8, p. 42] like laptops, personal digital assistants (PDAs), smartphones, tablets, and mobile phones [9]. These types of devices distinguish themselves through the following characteristics: Performance [10], [11], Data Input [10], [12], Connectivity [10], [11], Mobility [10], [11], [13], Context Awareness [14], and User Interface [10].

As a consequence of the tremendous progress in the field of mobile technology in the last couple of years, the capabilities of smartphones and tablets have increased vastly. They offer more and more functionalities that were only available on desktop computers or laptops a few years ago. Furthermore, they provide several advantages of the classical mobile phone (feature phone) and an intuitive handling [15]. The new generation of mobile devices smartphones and tablets strongly differs from laptops and feature phones, although their functionalities considerably overlap. Therefore, it is crucial to consider their characteristics during the usability engineering process in order to create a satisfying user experience.

Before dealing with mobile business applications, it is necessary to define "mobile business" (m-business). Unfortunately, the literature does not provide an internationally accepted definition of this term [16]. However, Koenigstorfer was able to identify the main characteristics of mobile business by analyzing published definitions. Firstly, the services provided in the context of mobile business are available everywhere the user has access to his device. Secondly, these services contain business processes between all market participants, who could be consumer, businesses or governments [16].

After defining mobile business we are able to determine the unique features of mobile business applications. Different terms and classifications are discussed for these features in the literature. However, if you consider them as a group, they can be described as the following: Ubiquity [17], [18], Accessibility [17]–[19], Convenience [18], Personalization [17], [18], Integration [19], Context Awareness [18], Security [17], [20].

B. Usability Engineering

We already know that the success of mobile devices and mobile applications is strongly related to their high degree of usability [3]. But there is still the question how to reach this objective.

One solution is described by ISO 9241-210: "Ergonomics of human-system interaction Part 210: Human-centred design (HCD) for interactive systems" (successor of ISO 13407). The standard based on the centred design paradigm from the 1980s provides a guidance on how to manage the development process to create user interfaces with a comfortable user experience and a high degree of usability [21].

A crucial part of ISO 9241-210 is the specification of the iterative development cycle. The standard defines five main activities which are conducted sequentially to achieve the goals of HCD. Four of them form a cycle being repeated until the product fulfils the required usability objectives.

The development process starts with the planning activity. The goal of this part is to integrate the HCD processes into the overall system development process. The next activity, which is also considered the first one of the cycle, deals with understanding and specifying the context of use. Following this, the user requirements are specified. After the identification of the requirements, the insights are used to produce design solutions. Subsequently, the user-based assessment is carried out to evaluate the developed prototypes against the user and organizational objectives. The early and regular evaluation offers two essential advantages. On the one hand, changes are implemented before it becomes disproportionally expensive. On the other hand, the demands and needs of the users will be under better consideration because of the substantial and regular feedback [22]. The results of the evaluation provide a refactored basis and new insights for a further iteration of the development cycle. This procedure is repeated until the developed system "meets the user requirements" [23, p. 11].

The evaluation activity is the part of the HCD that this paper deals with. In the following subsection an overview of the different evaluation methods is provided and it is explained why a demand for a questionnaire assessing the usability of mobile devices exists.

C. Evaluation Methods for UI Design

As mentioned above, ISO 9241-210 focuses on how to manage the HCD processes, but it does not contain methods or techniques for the activities. To fill this gap and to more easily apply the HCD approach, Magurie published a paper with a comprehensive collection of methods which can be used during the development cycle [22]. Furthermore,

he suggests the following methods: ”Participatory evaluation”, ”assisted evaluation”, ”heuristic or expert evaluation”, ”controlled user testing”, ”satisfaction questionnaires”, ”assessing cognitive workload”, ”critical incidents”, and ”post-experience interviews” [22, p. 590]. According to Maguire, it is not necessary to apply all methods during development. The methods should be selected on the basis of project goals, available resources and progress of the development. Since the goal of the paper is to develop a questionnaire to assess the usability of mobile devices, we will have a closer look at the satisfaction questionnaires.

This kind of questionnaire ”captures the subjective impressions formed by users based on their experiences with a deployed system or new prototype” [22, p. 681]. The advantage of the method is that it can be carried out in an easy and inexpensive way compared to other methods. Over the course of time, many questionnaires were carefully developed for assessing the usability of information systems, like the Questionnaire for User Interface Satisfaction (QUIS) [24], Software Usability Measurement Inventory (SUMI) [25], the After-Scenario Questionnaire (ASQ), the Post-Study System Usability Questionnaire (PSSUQ), the Computer System Usability Questionnaire (CSUQ) (all three by [26]), the System Usability Scale (SUS) [5], and the Usability Metric for User Experience [6]. The common factor of these questionnaires that they are generic and therefore they assess the usability of a system on an abstract level. The advantage of a generic approach is that it makes the questionnaires applicable for almost all information systems. However, it is difficult to identify usability problems and errors in the user interface design, due to the generality of the questions. That is one reason why usability questionnaires for specific systems are developed, like Website Analysis and Measurement Inventory (WAMMI). With the arrival of mobile technology a demand for such questionnaires also emerged in this area. In 2006 Ji et al. as well as Ryu and Smith-Jackson published independently from each other a questionnaire to evaluate the usability of mobile phone user interfaces. The specification to the device type allows them to ask questions being more specific and closer to technology than the generic approaches. As a result, the success rate of identifying usability problems could be increased in both cases [27], [28]. Since both questionnaires mainly address consumer products and applications, and not business applications, [29] developed MoBiS-Q one year later. Eventhough if MoBiS-Q was designed for mobile business applications, the degree of abstraction in the questions is similar to that of the generic questionnaire. Thus, it is more difficult to draw inferences about considering the weaknesses in the usability of the user interface compared to the other two questionnaires for mobile phones.

To summarize, the benefits of specific questionnaires are obvious for the development of user interfaces, but there have been enormous strides in the field of mobile technology since 2007. Therefore, they no longer satisfy all requirements of the current generation of mobile business applications. Nevertheless, the current state of research creates a good basis for the development of such a questionnaire.

III. UI Design Requirements for Mobile Devices

The goal of this section is to determine the requirements for the user interface design of mobile business applications. They serve as a foundation for the development of the usability evaluation questionnaire, described in the next section. The requirements analysis has proceeded as follows.

Starting with an analysis and comparison of the Android and iOS design guidelines, we identify user interface components that are characteristic for the current generation of mobile applications and that are essential for usability evaluation [30], [31]. These insights are extended by additional guidelines and theories from the literature that address user interface design and development. In the second step, the requirements for the component design are formulated with the aid of the literature. To consider all aspects of usability, we ensure that the set of identified components and their requirements cover all usability principles described by [27]. Additionally, the characteristics of mobile devices and the unique features of mobile business applications are included to obtain a valid result. Some of the guidelines and theories consider older generations of mobile devices and applications. Therefore, it is necessary to prove their validity for our context. In the following subsections a short summary of the requirements is provided.

A. Navigation and Organization

The first category of design requirements is ”Navigation and Organization” dealing with how a mobile business application should be structured and how the screens of the application should be organized to create a pleasant user experience. The application has to be designed so that the user can get the required information with minimal effort [31]–[34]. To reach this, the navigation concept has to be intuitive, the differences between navigation and content elements have to be clear and the information has to be presented in a way supporting the recognition and processing of the information [30], [31], [34]. Also, the user should always be able to undo the last steps [30]–[35].

B. Terminology and Wording

Another crucial part of the design requirements is ”Terminology and Wording”, since a large proportion of the communication between user and application takes place via texts. In this way, for example, the application gives instructions to the user, informs him about its status or provides content. As a general rule, the language should be similar to the one used by the target user group. All words or texts used in a mobile context should be specific and precise, but still concise because of the small screen size and the external

distractions. Furthermore, spelling and grammar should be correct, the terminology consistent, and the abbreviations clear [30], [31]. In terms of instructions, error messages, help, and tutorial the texts should be informative and helpful but not excessive or even offensive [30], [34], [36].

C. Interaction and Feedback

In this subsection the focus is on the elements which are necessary or helpful for the interaction between the user and the mobile business application. Especially in context of mobile devices, data input is a critical success factor because the form factor aggravates the manual data input [33], [34]. Due to the technological progress, mobile application can make use of elaborate animations and transitions. However, they should only be used to support the user's interaction with the device [31], [33]. This is also the case for gestures. They have to be easy to learn and intuitive to use [34], [37]. Furthermore, the user should always be informed on what the application is doing: is it running? Is it downloading or similar? It gives the user the feeling of controlling the application. [30], [31], [35], [38], [39] Another component in this category is messages interrupting the normal workflow of the user, like notifications and modal windows. They should be well-considered before being used, otherwise, the user recognizes them with reduced attention [30], [31], [34], [39]. Summing up, all these elements serve to facilitate the interaction between user and device.

D. Visual Design and Aesthetics

According to [33], [30], and [31] an aesthetic appearance strongly improves the usability of the system. To create an aesthetic user interface, several aspects have to be considered. First of all, the design should be consistent to appear professional and to provide the user additional orientation in the application [10], [31], [33], [39]. In addition, the typography and color scheme should be adapted to the content of the application and the characteristics of mobile devices [30], [31], [40], [41]. Due to the limited screen, size images and icons are often used to communicate with user. But it has to be ensured that the images and icons are well discernible and easy to understand [30], [31], [34]. In the context of business application, branding is an important and critical topic. Too much or too intrusive branding strongly impairs the usability [31].

E. Adaptability and Integration

This category deals with the customizability of the application and how it interacts with other applications and the operating system. The usability of the initialisation process already plays an important role. If it is too elaborate, the user will close the application and is most likely not to use it again [34]. Furthermore, the user should be able to personalize the application to enhance his or her workflow and to share information easily with other users via similar applications [10], [33], [34]. Besides that, developers have to consider the behavior of the application during changes in the device orientation and the start and stop behavior of the application. Both are components influencing the usability [31], [33], [38].

F. Performance and Reliability

The last category covers the performance components considering the computational power and the connectivity. Even though if the mobile devices have performance limitations, short reaction times are still essential for a good usability [30], [31], [33], [39]. Another aspect which should be automatically handled by the application is connectivity issues. The user should be informed about changes and possible restrictions, but he or she should not have to become active to solve the problem [33], [35], [39]. In addition, an application should provide a synchronization to allow a convenient usage, if the service is used on different devices or platforms [31], [34]. Security and privacy are the last topics that have to be addressed in this category. Implemented poorly, like too complex security mechanisms, these requirements can cause a bad user experience and consequently discourage the user from using the application [10], [38].

The numerous requirements show that many aspects exist that have to be considered by the user interface designer and developer of mobile business applications. As a consequence, it is useful to develop a comprehensive and clear questionnaire covering all these aspects. With the manual evaluation of paper-based surveys being time-consuming and expensive, it is a logical consequence to provide a software tool to conduct and analyze surveys using this questionaire.

IV. Implementation Details

This section considers the development of both artifacts: the questionnaire and the corresponding software tool. In the first part the derivation of the questions from the design requirements is elaborated. Afterwards, an insight into the development of the software tool, called *mugram* (**m**obile **u**sability histo**gram**), is provided.

A. Development of the Questionnaire

The questionnaire design is heavily oriented towards the requirements described in section 3. It is structured in the same manner and the questions are directly derived from the aspects addressed by these guidelines. Table I contains six questions from the questionnaire. 100 of the 104 questions are closed-ended questions with a five point Likert scale. This means that a question contains a statement and a scale which allows the user to state how far he or she agrees or disagrees with the statement. One advantage of this question design is to make the results comparable in contrast to open questions. Furthermore, the scale offers the user more selection options than a simple yes-no question. The user

Table I
EXTRACTION OF QUESTIONS FROM CATEGORY NAVIGATION AND ORGANISATION

Question	strongly disagree				strongly agree
I have to touch too many buttons to access to the required information.	☐	☐	☐	☐	☐
It is often difficult to distinguish between navigation and content elements.	☐	☐	☐	☐	☐
The gestures differ strongly from what I am used to.	☐	☐	☐	☐	☐
The combination of colors conveys a feeling of harmony to me.	☐	☐	☐	☐	☐
The images support the recognition of the content.	☐	☐	☐	☐	☐
The latency times are too long and disturb the work flow.	☐	☐	☐	☐	☐

also has the possibility to give a neutral answer and thus the result is not distorted by false answers, if he or she is indifferent. To get a meaningful result from this type of questions it is necessary to stick to some rules. First of all, simple and unambiguous words have to be used to make the questions readily understandable for the end user [42]. Secondly, the word ”not” should be avoided in a statement because the inversion aggravates the answering of the question for the user [42]. However, the developed questionnaire contains two questions with a ”not”. But it was only used to emphasize the main statement of the question which does not include a ”not”. Furthermore, research has shown that users rather tend to agree with a statement than to disagree with it. To compensate this effect it is recommended to address an aspect with a positively and negatively formulated question [42]. Since the objective was to create a lightweight questionnaire, we avoided such phrases. Instead, the problem is handled during the processing of the results, which is explicitly explained in context of the development of the software artifact. However, it was ensured that the number of positive and negative statements is equal within the categories to obtain a valid mean value. Besides the closed-ended questions the questionnaire contains a few open questions to give the users the possibility to express their perception, impressions, and criticism.

Additionally, the ten questions from the SUS were incorporated into the questionnaire to provide a clue for the analysis of its results. These questions allow calculating a score which states how usable the entire system is. Thus, the development team does not only obtain an individual assessment for the user interface components, but also an aggregated measurement for the entire system. There are two reasons why the SUS and not another scale were selected for this purpose. It is simple and concise. Beyond that, [43] validated its reliability and validity by an empirical analysis evaluating data from nearly one decade.

B. Development of the Software Artifact

Since the conduct and analysis of paper-based surveys is expensive, the software artifact *mugram* was developed for the usability evaluation questionnaire, in order to support the evaluation process and to reduce the inhibition level of using it. Therefore, it is crucial that the tool exhibits a high degree of usability and that it is as platform independent as possible in order to facilitate its usage.

1) Technical Aspects: Due to the requirements mentioned above, *mugram* was implemented as a web application. In comparison to desktop applications or native applications for mobile devices, a web application has the advantage of the platform independence. It can be used on every device with a modern browser without developing platform specific client software. Furthermore, the centralized concept of web applications enables the conduct of parallel evaluations and supports the collaboration in the development team compared to standalone solutions.

While HTML 5 and JavaScript are used for the presentation tier of the application, the logic tier was implemented with PHP. For the persistence of the data a MySQL database server is deployed on the data tier. It should be mentioned that the separation between the logic and data is not clear. Due to performance aspects and the scalability of the system, the processing of the survey data is mainly executed by the database server.

To create a comfortable user experience on the different supported device types, the user interface of the application is based on the *Bootstrap* 3 framework developed by Twitter. This framework, written in HTML, CSS, and JavaScript, facilitates the development of responsive and touch-optimized web applications. To make use of the entire functionality of *Bootstrap*, it was necessary to incorporate the JavaScript framework *jQuery*. In addition, *Bootstrap* was extended by *bootstrap-datetimepicker*, a plugin which provides a more user-friendly method to select dates and times. To simplify the validation of form input and to display static overlay elements, the jQuery plugins *Validation* and *Clingify* were integrated. Since none of these libraries provides chart plotting functionalities, the charting library *Highcharts*, which is written in HTML5 and JavaScript, is used in version 3.0.

2) Implementation Aspects: The *mugram* tool was implemented in order to meet the requirements of an evaluation tool for the development of mobile business applications and to make the defined questionnaire practically usable.

In order to allow an intuitive operation and to manage numerous survey rounds, the backend of *mugram* is organized in projects and rounds. For every mobile application that is evaluated with *mugram* the user creates a new project.

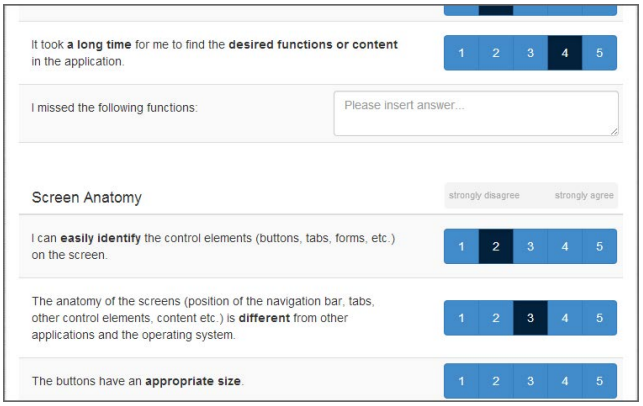

Figure 1. Frontend of mugram: questionnaire

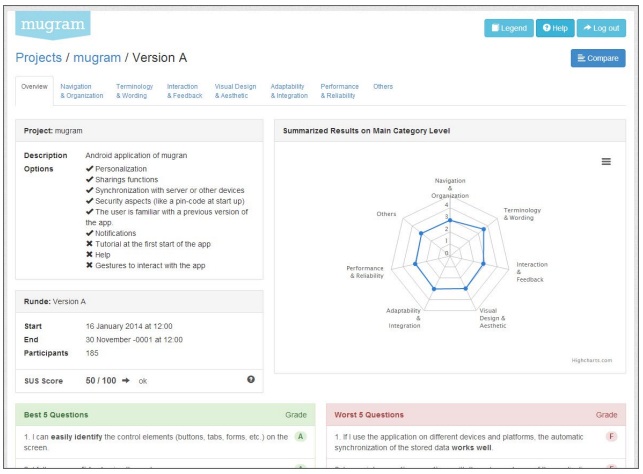

Figure 2. Backend: dashboard of mugram

During this process the user is able to adapt the scope of the questionnaire according to the functionalities of the application. If an application, for instance, provides gestures for the interaction, the user can enable this function to include the corresponding question to the questionnaire. Within a project there are rounds which represent survey rounds. It is envisaged to create an individual round for every version of the application. The reason for that becomes obvious, if the compare feature is explained. For every round there is a specific link which forwards someone to the questionnaire.

After the conduct of a survey via a form depicted in Figure 1 the backend provides the possibility to analyse the results. The first tab of the analysis provides an overview of the results in the form of a dashboard (cf. Figure 2). It informs the user about the SUS score, the five best and worst questions and a summary of the result on the main category level. The other tabs contain a detailed analysis for each question. For the closed-ended question the analysis consists of the average value, the median, the standard deviation,

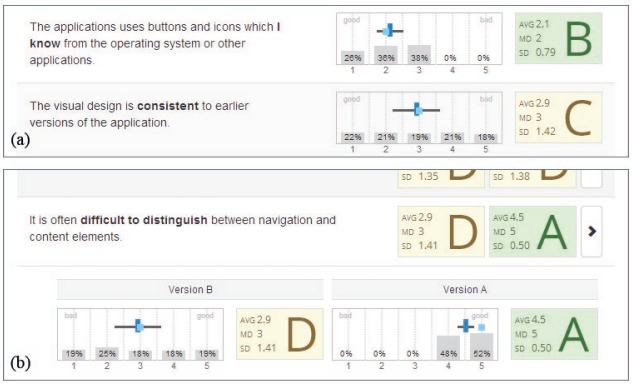

Figure 3. Backend: (a) detailed results and (b) comparison of two rounds

the distribution of the answers and a grade (cf. Figure 3a). Although the grade (between A and F) is only calculated from the average value, its usage can be ascribed to two reasons. In the first place, there are positively and negatively formulated questions, which means that sometimes 1 is the best result and sometimes 5. To facilitate the interpretation of the results, the average values are normalized by the usage of grades. In the second place, the grades allow compensation for the acquiescence effect that user rather tend to agree with a statement than to disagree with it. [42] suggests a value of 10% to describe this effect. Therefore, the interval lengths for the grades are not commensurate, but the interval covering ”strongly agree” is 10% shorter than the one covering ”strongly disagree”. The lengths of the other three intervals are also adjusted according to the harmonic sequence. However, the acquiescence effect has to be regarded only in context of the questions. For the grades of sub and main categories, the lengths of the intervals do not have to be adjusted, because the same number of positively and negatively formulated questions already compensate the effect. The answers for open questions are hidden in a collapsible container.

Since the development process of user interfaces is an iterative one, it can be interesting for the development team to see the alterations between different versions of the mobile application. Therefore, *mugram* allows the user to compare up to three rounds within a project with each other (cf. Figure 3b). The feature can be also used for A/B testing - an evaluation method that allows you to compare two versions of a design and to see which performs better [44].

C. Usage of the Questionnaire and the Software Tool

The target group of the questionnaire and the software tool *mugram* are user interface designers and developers of mobile business applications. They are supposed to use these tools to improve the usability of their applications. To achieve this, the participants of the surveys based on

the questionnaire have to be potential users of the applications. In addition, it is necessary to have a fully-functional prototype. Since the questionnaire is comprehensive and it is necessary to test an application extensively before completing the questionnaire, it is primarily suitable to use it in observed field, laboratory tests or with paid testers of the company itself. In such situations the testers are willing to invest the effort to carefully complete the questionnaire. Due to the amount of questions, it is not recommended to use the questionnaire in form of a classical online survey which is conducted without supervision.

V. Evaluation and Implications

This section illustrates the evaluation of both developed artifacts. The goal is to evaluate the extent to which the objectives for the questionnaire and the software tool are fulfilled. For this purpose the questionnaire is checked for its completeness, validity, and usefulness and the software tool for its usability and usefulness.

A. Evaluation Design

As an evaluation method, expert interviews with user interface designers and mobile application developers were chosen. It was taken into account to apply the evaluation procedure which has been used for the usability checklist of [27]. They conducted a conventional usability testing together with their developed usability checklist and compared the results in relation to the identified usability problems. In this way, the checklist could be evaluated in the context in which the tool was used later. But there were some reasons against this kind of procedure and for expert interviews. First of all, the design of the questions and the applicability of the answers should be assessed by usability experts. Secondly, the corresponding software tool should be evaluated. Since the users of the tool are user interface designers and developers, it could not be tested by conventional users of mobile business applications. For these reasons, the decision was taken in favor of the expert interviews.

The expert interview consists of two parts. The first one deals with the questionnaire, the second one with the software tool. Every part contains two sections: A quantitative section with closed-ended questions to make the interviews comparable and a qualitative one with open questions to collect the objections and the improvement suggestions of the experts. For the closed-ended question a five-point Likert scale was used (1: ”strongly disagree” to 5: ”strongly agree”). Overall, five interviews were conducted with user interface designers and mobile application developers of two start-ups and a large company of the service sector in Germany.

B. Results for the Questionnaire

The summarized results of the closed-ended questions for the first part of the evaluation are shown in the upper part of Table II.

In terms of completeness, the questionnaire has an average degree of agreement of 3.8. According to the experts, the questionnaire is very comprehensive, but there are some usability aspects which are not addressed yet. These include questions dealing with the emotional perception of the user interface, ”issues around the device itself” and the quality of translations within the application.

With regard to usefulness and validity, whether the results were helpful for the development of mobile business application, the item obtains an average degree of agreement of 4.0 and the item, whether the results could improve the usability of such an application, receive 4.2. Besides this positive feedback, some experts mentioned that the questionnaire is also an extensive item pool which can already be used as a guideline during the early stages of development.

Despite of the usefulness and the validity, the experts did not agree on the accelerating impact on the development process by the questionnaire (average degree of agreement 2.8). This can be ascribed to two reasons. Firstly, it was mentioned that the completion of the questionnaire costs too much time. Secondly, one expert explained that such a questionnaire could not be used until a fully-functional prototype is available. Therefore, it would not accelerate the actual development process. Nevertheless, the entire life cycle of a mobile business application could benefit from the results of the questionnaire. The statement, whether the experts would use the questionnaire for the development, obtains an average degree of agreement of 3.6. However, it has to be mentioned that both experts who gave a 3 for the question justified their answer with usability aspects not being addressed. The result for the question concerning the level of difficulty is not meaningful. Its average degree of agreement is 3.4 (in this case less is better). However, the standard deviation has a value of 1.34. Therefore, the result cannot be considered statistically convincing.

Beyond the aspects mentioned above, there are some further aspects which the experts especially liked about the questionnaire. Three of these are its appropriate structure, its comprehensiveness, and the question design. In addition, two experts commended the end-user oriented language of the question. Another expert considered the integration of the SUS positively, since it would increase the reliability and validity of the results. Concerning possible improvements, some experts suggested reducing the amount of questions, respectively developing a short version of the questionnaire for brief surveys. Furthermore, some of them wished for more free text answers. But this would increase the effort of completing the questionnaire and impair the comparability.

In summary, the evaluation shows that the questionnaire is a helpful tool for the development of mobile business applications and it can positively influence the usability. Moreover, it provides some interesting insights for the improvement of the questionnaire.

Table II
EVALUATION RESULTS FOR THE CLOSED-END QUESTIONS

Questionnaire (1 = strongly disagree, 5 = strongly agree)	AVG	SD
Q1: I think that the questionnaire addresses all aspects being relevant for the usability of a mobile application.	3.80	1.30
Q2: In my opinion the results of the questionnaire are helpful for the development of mobile applications.	4.00	0.71
Q3: I think that the results of the questionnaire can accelerate the development process of mobile applications.	2.80	0.84
Q4: In my opinion the results of the questionnaire can improve the usability of a mobile application.	4.20	0.45
Q5: The questionnaire is too sophisticated to be answered by the end user.	3.40	1.34
Q6: I would use the questionnaire for the development of a mobile application.	3.60	0.55
Software Tool - mugram (1 = strongly disagree, 5 = strongly agree)		
Q7: I thought the system was easy to use.	4.40	0.49
Q8: I found the system unnecessarily complex.	1.60	0.80
Q9: Mugram facilitates the conduct and analysis of a usability survey.	4.00	0
Q10: The analysis of the results is sufficient.	4.00	0.89
Q11: The comparison feature of the results is helpful.	4.60	0.80
Q12: I would use the tool for the development of a mobile application.	3.80	0.40

AVG: average; SD: standard deviation

C. Results for the Software Tool

The second part of the evaluation focuses on the software tool *mugram*. The second part of Table II shows the results for the quantitative questions.

After the adjustment of Q7s and Q8s scale, the average rating for the system's usability obtains a value of 4.4 on a scale between 1.0 and 5.0, which is a very good result. The adjustment is necessary since the item Q8 is reversed.

With regard to the range of features, *mugram* achieved similarly good results. All experts agreed (4.0) on the item that *mugram* facilitates the conduct and analysis of a usability survey. The item considering the scope of the analysis obtained an average degree of agreement of 4.2. However, it has a relatively large standard deviation of 0.89, which indicates that the experts have different requirements for the analysis of the results. A feature which was commended by four experts is the comparison feature of the rounds. The corresponding quantitative item has an average degree of agreement of 4.7. The item "I would use the tool for the development of a mobile application." obtained an average agreement of 3.8. While almost all experts gave an "agree" (4.0) for this item, one of them gave a "neither agree nor disagree" which she warranted with the missing export function for the raw data.

In addition to the compare feature, there are some more aspects or features which were commended by the experts. One of them is the platform independence which allows using *mugram* on every device with a modern web browser. Furthermore, the clean and aesthetic design of the user interface was pleasing to the experts; above all, the visual presentation of the results with the dashboard and different types of diagrams. But there were also some points of criticism, respectively for feature requests. Firstly, the radio buttons of the questionnaire should not contain numbers, since the numbers 1 to 5 are associated with grade in many countries and accordingly influence the user in his or her answer. Another improvement suggestion is to make the questionnaire more customizable. It should be possible to enable and disable entire sub categories. Overall, the questionnaire should be adaptable for every round but not for every project. The purpose behind that is to customize the rounds according to the changes in versions of an application. Beyond that, the experts required more interfaces: a REST API for the automated conduct of surveys and for the integration into management dashboards, a sharing function of the results for social networks like Google Plus or Yammer, and an export function of the raw data to conduct further analyses with SPSS, Stata or R.

In short, the software tool *mugram* received very positive feedback from the experts. The ease of use, the visual design, and the provided functionality have convinced most of the experts that they would use it for the development of their mobile business applications, if some smaller features were implemented.

VI. CONCLUSION

A. Summary

The objective of this paper is the development of an evaluation questionnaire to assess single components of a mobile user interface regarding their usability in order to meet the requirements of the current mobile business technology. The intention behind that was to create a tool which facilitates the development process, to increase the usability of mobile business applications.

According to design science research, a literature review was conducted in advance, followed by the development of two artifacts and their evaluation. The literature review exposed that there is no comparable questionnaire

for evaluating the usability of mobile business applications. Nevertheless, it seems to be a conventional technique to use questionnaires during the software development to collect feedback from future users. Based on the insights gained from other usability questionnaires and guidelines for the design of mobile user interfaces, the questionnaire and the corresponding software *mugram* were developed. To evaluate both artifacts, five usability experts were interviewed. According to the results of this evaluation, both the questionnaire and *mugram* are helpful for the development of mobile business applications and can improve the usability of these applications, which are the objectives of this paper. This statement is emphasized by the fact that most of the experts would use *mugram* for the development of their mobile applications. Overall, the experts exposed small weaknesses of the artifacts and contributed some improvement suggestions for further development.

B. Future Work

The chosen method of expert interviews for evaluating our questionnaire and tool lacks validity since it has not been based on a real application and real end-users. But this is an issue that could be addressed in future research, since it would further support the questionnaire's reliability and validity. In addition, the questionnaire should be expanded to include the missing usability aspects, which were mentioned during the evaluation, and other mobile platforms besides Android and iOS should be integrated. However, the number of questions should be reduced. One possibility is to make recourse to statistical theories, like the classical test theory or the item response theory, to identify correlating questions or questions with a conspicuously low variance. The latter aspect points to a useless question. Moreover, those methods allow determining the weight of the different main and sub categories. One specific method, for instance, could be the multivariate regression analysis. To apply such statistical methods it is necessary to collect numerous survey records. That is why *mugram* should be developed further. Offered as software as a service solution, it facilitates the collection of survey data which could provide helpful insights for the development of the questionnaire.

REFERENCES

[1] IDC. (2014) Global smartphone shipments forecast 2010-2017. [Online] Available: http://www.statista.com/statistics/263441/global-smartphone-shipments-forecast/ [Last accessed on 2014-06-01].

[2] IDC. (2013) Tablet Shipments Forecast to Top Total PC Shipments in the Fourth Quarter of 2013 and Annually by 2015, According to IDC . [Online] Available: http://www.idc.com/getdoc.jsp?containerId=prUS24314413 [Last accessed on 2014-06-01].

[3] H. B.-L. Duh, G. C. Tan, and V. H.-h. Chen, "Usability evaluation for mobile device: a comparison of laboratory and field tests," in *Proceedings of the 8th conference on Human-computer interaction with mobile devices and services*. ACM, 2006, pp. 181–186.

[4] D. Saffer, *Designing for interaction : creating innovative applications and devices.* Berkeley, Calif. : New Riders, 2010.

[5] J. Brooke, "SUS-A quick and dirty usability scale," *Usability evaluation in industry*, vol. 189, p. 194, Sep. 1996.

[6] K. Finstad, "The Usability Metric for User Experience," *Interacting with Computers*, vol. 22, no. 5, pp. 323–327, Sep. 2010.

[7] A. Hevner, S. March, J. Park, and S. Ram, "Design science in information systems research," *MIS quarterly*, vol. 28, no. 1, pp. 75–105, 2004.

[8] P. Tarasewich, "Issues in Mobile E-Commerce," *Communications of the Association for Information Systems*, 2002.

[9] L. Gorlenko and R. Merrick, "No wires attached: Usability challenges in the connected mobile world," *IBM Systems Journal*, vol. 42, no. 4, pp. 639–651, 2003.

[10] B. Adipat and D. Zhang, "Interface Design for Mobile Applications," *AMCIS*, p. 494, 2005.

[11] M. Satyanarayanan, "Fundamental challenges in mobile computing," *Proceedings of the fifteenth annual ACM symposium on Principles of distributed computing - PODC '96*, pp. 1–7, 1996.

[12] H. Heitkötter, T. Majchrzak, U. Wolffgang, and H. Kuchen, "Business Apps: Grundlagen und Status quo," 2012.

[13] P. Abrahamsson, A. Hanhineva, H. Hulkko, T. Ihme, J. Jäälinoja, M. Korkala, J. Koskela, P. Kyllönen, and O. Salo, "Mobile-D: an agile approach for mobile application development," in *Companion to the 19th annual ACM SIGPLAN conference on Object-oriented programming systems, languages, and applications*. ACM, 2004, pp. 174–175.

[14] T. Hofer, W. Schwinger, M. Pichler, G. Leonhartsberger, J. Altmann, and W. Retschitzegger, "Context-awareness on mobile devices - the hydrogen approach," in *36th Annual Hawaii International Conference on System Sciences, 2003. Proceedings of the*. IEEE, 2003, p. 10 pp.

[15] T. Goetz, N. Feldmann, and S. Schmidt, "Smarter apps motor fr geschftsmodellinnovationen," in *Smart Mobile Apps*, ser. Xpert.press, S. Verclas and C. Linnhoff-Popien, Eds. Springer Berlin Heidelberg, 2012, pp. 507–518.

[16] J. Königstorfer, *Akzeptanz von technologischen Innovationen*. Wiesbaden: Gabler, 2008.

[17] P. Alahuhta, J. Ahola, and H. Hakala, "Mobilizing business applications," *Technology Review*, vol. 167, no. 2005, 2005.

[18] D. Xiaojun, I. Junichi, and H. O. Sho, "Unique Features of Mobile Commerce," *Journal of Electronic Science and Technology of China*, vol. 2, no. 3, 2004.

[19] F. F.-H. Nah, K. Siau, and H. Sheng, "The value of mobile applications," *Communications of the ACM*, vol. 48, no. 2, pp. 85–90, Feb. 2005.

[20] F. Aloul, S. Zahidi, and W. El-Hajj, "Two factor authentication using mobile phones," in *2009 IEEE/ACS International Conference on Computer Systems and Applications*. IEEE, 2009, pp. 641–644.

[21] D. A. Norman and S. W. Draper, "User Centered System Design; New Perspectives on Human-Computer Interaction," Jan. 1986.

[22] M. Magurie, "Methods to support human-centred design," *International Journal of Human-Computer Studies*, vol. 55, no. 4, pp. 587–634, Oct. 2001.

[23] ISO 9241-210, "Ergonomics of human-system interaction – Part 210: Human-centred design for interactive systems," 2010.

[24] J. P. Chin, V. A. Diehl, and K. L. Norman, "Development of an instrument measuring user satisfaction of the human-computer interface," in *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 1988, pp. 213–218.

[25] J. Kirakowski and M. Corbett, "SUMI: the Software Usability Measurement Inventory," *British Journal of Educational Technology*, vol. 24, no. 3, pp. 210–212, Sep. 1993.

[26] J. R. Lewis, "IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use," *International Journal of Human-Computer Interaction*, vol. 7, no. 1, pp. 57–78, 1995.

[27] Y. G. Ji, J. H. Park, C. Lee, and M. H. Yun, "A Usability Checklist for the Usability Evaluation of Mobile Phone User Interface," *International Journal of Human-Computer Interaction*, vol. 20, no. 3, pp. 207–231, Jul. 2006.

[28] Y. S. Ryu and T. L. Smith-Jackson, "Reliability and Validity of the Mobile Phone Usability Questionnaire (MPUQ)," *Journal of Usability Studies*, vol. 2, no. 1, pp. 39–53, 2006.

[29] M. Markova, A. Aula, T. Vainio, H. Wigelius, and M. Kulju, "MoBiS-Q: a tool for evaluating the success of mobile business services," in *Proceedings of the 9th international conference on Human computer interaction with mobile devices and services*. ACM, 2007, pp. 238–245.

[30] Android Developers. (2013) Design Android Developers. [Online] Available: https://developer.android.com/design/index.html [Last accessed on 2014-06-01].

[31] Apple Inc. (2014) iOS Human Interface Guidelines: Designing for iOS 7. [Online] Available: https://developer.apple.com/library/ios/documentation/userexperience/conceptual/MobileHIG/ [Last accessed on 2014-06-01].

[32] D. Norman, *The design of everyday things*. Basic Books, 2002.

[33] J. Gong and P. Tarasewich, "Guidelines for handheld mobile device interface design," in *Proceedings of DSI 2004 Annual Meeting*, 2004, pp. 3751–3756.

[34] J. Nielsen and R. Budiu, *Mobile Usability*. Pearson Education, 2012.

[35] S. S. Chan, X. Fang, and J. Brzezinski, "Usability for Mobile Commerce Across Multiple Form Factors," *Journal of Electronic Commerce Research*, vol. 3, no. 3, p. 187, 2002.

[36] B. Shneiderman and C. Plaisant, *Designing the User Interface: Strategies for Effective Human-Computer Interaction*. Addison-Wesley/Pearson, 2010.

[37] D. Saffer, *Designing Gestural Interfaces*. O'Reilly Media, 2008.

[38] J. Häkkilä and J. Mäntyjärvi, "Developing design guidelines for context-aware mobile applications," in *Proceedings of the 3rd international conference on Mobile technology, applications & systems - Mobility '06*. New York, New York, USA: ACM Press, Oct. 2006, p. 24.

[39] N. Z. B. Ayob, A. R. C. Hussin, and H. M. Dahlan, "Three Layers Design Guideline for Mobile Application," *2009 International Conference on Information Management and Engineering*, pp. 427–431, 2009.

[40] J. Benson, K. Olewiler, and N. Broden, "Typography for Mobile Phone Devices: The Design of the QUALCOMM Sans Font Family," in *Proceedings of the 2005 Conference on Designing for User eXperience*, ser. DUX '05. New York, NY, USA: AIGA: American Institute of Graphic Arts, 2005, pp. 2–18.

[41] J. Tantillo, J. Di Lorenzo-Aiss, and R. E. Mathisen, "Quantifying Perceived Differences in Type Styles: An Exploratory Study." *Psychology & Marketing*, vol. 12, no. 5, pp. 447–457, 1995.

[42] J. A. Krosnick and S. Presser, "Question and questionnaire design," *Handbook of survey research*, vol. 2, pp. 263–314, 2010.

[43] A. Bangor, P. Kortum, and J. Miller, "Determining what individual SUS scores mean: Adding an adjective rating scale," *Journal of usability studies*, vol. 4, no. 3, pp. 114–123, 2009.

[44] B. Martin, B. Hanington, and B. Hanington, *Universal Methods of Design: 100 Ways to Research Complex Problems, Develop Innovative Ideas, and Design Effective Solutions*. Rockport Publishers, 2012.