Received February 10, 2020, accepted March 16, 2020, date of publication March 19, 2020, date of current version March 30, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2981892

Usability of Mobile Applications: A Systematic Literature Study

PAWEŁ WEICHBROTH
Department of Software Engineering, Faculty of Electronics, Telecommunications and Informatics, Gdańsk University of Technology, 80-233 Gdańsk, Poland
e-mail: pawel.s.weichbroth@gmail.com

ABSTRACT Since the release of the first mobile devices, the usability of on-board applications has been the concern not only of software vendors but hardware manufacturers as well. The academia community later willingly joined the discussion on usability in terms of theory and empirical measurement, having experience and knowledge in desktop settings. At first sight, such a background should guarantee a solid foundation to conduct research on software usability in a new setting. However, a preliminary study on the subject matter revealed methodological disorder in contemporary literature. As a matter of fact, a need emerged to review existing usability definitions, attributes and measures to recognize all associated aspects. In order to fill this void, we conducted a systematic literature review on usability studies indexed by the Scopus database and devoted to mobile applications. The input volume covers 790 documents from 2001 to 2018. The data analysis shows that the ISO 9241-11 usability definition has been adopted in an unchanged form and popularized as the standard by the HCI community. Secondly, in total, 75 attributes were identified and analysed. The most frequent are efficiency (70%), satisfaction (66%) and effectiveness (58%), which directly originate from the above definition. Subsequently, the less frequent are learnability (45%), memorability (23%), cognitive load (19%) and errors (17%). The last two concern simplicity (13%) and ease of use (9%). Thirdly, in the evaluation of usability, controlled observation and surveys are two major research methods applied, while eye-tracking, thinking aloud and interview are hardly used and serve as complementary to collect additional data. Moreover, usability evaluations are often confused with user experience dimensions, covering not only application quality characteristics, but also user beliefs, emotions and preferences. All these results indicate the need for further research on the usability of mobile applications, aiming to establish a consensus in the theory and practice among all interested parties.

INDEX TERMS Mobile applications, usability, attributes, measures, usability evaluation methods, systematic literature review.

I. INTRODUCTION

Amobile application is defined as ‘‘a software application developed specifically for use on small, wireless computing devices, such as smartphones and tablets, rather than desktop or laptop computers’’ [1]. A recent Statista report shows that in 2017 smartphones had a share of 77% of the global mobile device market [2], and more than 32% of the global population used a smartphone [3].

Although technological progress has been made regarding mobile devices equipped with computing power, leading to a shift from desktop computers, many limitations and challenges still remain [4]. From the many identified, usability has been the main concern, since the users of an application, and their judgment, ultimately decide on its success or failure [5]–[7]. Since the inception of the first smartphones, the subject of mobile application usability has gained attention both in academia communities and in the software vendors industry. While researchers are focused on formulating theories [8], modelling frameworks [9], and constructing methods and techniques [10], [11] for new settings, manufacturers simply desire to deliver high quality products [12].

Despite the abundance of research devoted to studies of mobile application usability on the one hand, and design patterns, prototyping tools and software frameworks on the

The associate editor coordinating the review of this manuscript and approving it for publication was Mario Luca Bernardi.

other, the term tends to be vague and loose, weakening the ability to capture its real facets and impeding the construction of measures. As a consequence, such methodological disorder violates the core assumptions and principles laying beneath the foundations of the usability notion. Therefore, considering the need for the emergence of a usability definition, its attributes and measures, along with evaluation methods, valid for mobile applications, in this paper we made an attempt to find reliable answers by conducting a systematic literature review. We expect that the obtained results can be used not only by researchers to perform further studies in this area, but also for practitioners engaged in mobile application development and quality-in-use evaluation to better understand the characteristics and measures of the notion.

The main contributions of this study include: (*i*) an evidence-based discussion of the usability definition, its attributes and measures, (*ii*) and an up-to-date map of the state of the art in usability evaluation methods (UEMs), adopted for and adapted to mobile applications, covering publications from 2001 to 2018.

The rest of the paper unfolds as follows. Section 2 provides the background on the subject addressed, and related work. Section 3 describes the research methodology. The definition and execution of the literature review are respectively presented in Sections 4 and 5. Section 6 provides an analysis of the extracted data, while the results are further discussed in Section 7, along with the future research directions. The conclusions are raised in Section 8.

II. BACKGROUND

Most people tend to use products that are easy to understand, work as expected, and eventually deliver value. In the context of the software engineering, system usability plays the crucial role in shaping perceived quality in use by its users [13], [14].

Usability is the study of the intersection of between systems and users, tasks and expectations in the context of use. Since many software products have been determined to be insufficient to meet user needs, several comprehensive studies have been conducted so far under the term usability, which move towards a better understanding and relevant measurement, aiming to cover all valid phenomena in one framework or model [15]–[17].

The results of the study, introduced by Weichbroth [18], show that over time the definition of usability has evolved. In 1991 the Organization for Standardization (ISO), in response to the emergence of the need of the software community to standardize some facets of software products, publicized the 9126 standard, which defines usability as ''a set of attributes of software which bear on the effort needed for use, and on an individual assessment of such use, by a stated or implied set of users'' [19].

Then, in 1998, ISO refashioned the usability definition in the ISO 9241-11 norm, which actually states that usability is ''the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use'' [20], [21]. While some argue that it is the most recognizable definition [18], others maintain that ''a generally accepted usability definition still does not exist, as its complex nature is hard to describe in one definition'' [22], [23].

The other usability definition can be found in ISO/IEC 25010 [24], which replaced the ISO/IEC 9126 standard from 2001 [25], and specifies usability as the ''degree to which a product or system can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use''. Here, it is worth noting that these two latest paraphrased definitions, however differently particularized, still share exactly the same three virtues which affect the user's ability to achieve specified goals.

Since the inception of the first official usability definition, one might argue that a great plethora of usability attributes have been taken into consideration regarding the ability to use particular software products, ranging from monolithic systems to lightweight web pages. Having said that, based on the literature search and analysis, in view of usability attributes that contribute to the quality in use of the desktop software, the latest study [18] shows that the most frequent are efficiency, satisfaction, learnability and effectiveness. The least frequent are understandability and operability, memorability, errors, attractiveness and accessibility.

To collect all necessary data in order to improve the quality of particular software facets, a variety of usability evaluation methods (UEMs) have been developed and empirically tested. One of the most recognized UEMs concern the family of user testing methods [26]–[28], in particular think-aloud protocol [29]–[31], question-asking protocol [32]–[34], performance measurement [35]–[37], log analysis [38]–[40], eye tracking [41]–[43], and remote testing [44]–[46]. Secondly, inspection methods, intended to be used by experts [47], refers to heuristic evaluation [48]–[50], cognitive walkthrough [51]–[53], perspective-based inspection [54]–[56], and guideline reviews [57]–[59]. Thirdly, inquiry methods, designed to gathering subjective data from users, utilize both quantitative (questionnaires [60]–[62]) and qualitative (interviews [63]–[65] and focus groups [66]–[68]) techniques. Furthermore, some authors also distinguish analytical modelling methods such as cognitive task analysis [69]–[71], task environment analysis [72]–[74] and GOMS analysis (Goals, Operators, Methods and Selection rules) [75]–[77].

Regarding the context of this study, Zhang and Adipat (2005) propose a generic framework for conducting usability tests for mobile applications through discussing existing methodologies and usability attributes [78]. As challenges, they point to the unique features of mobile devices and wireless networks which influence the usability of mobile applications, including mobile context, multimodality, connectivity, small screen size, different display resolutions, limited processing capability and power, and restrictive data entry methods. In the case of research methodologies for usability testing, they point to controlled laboratory experiments and field studies. While former limitations are ignorance of the mobile context and the preservation of reliable network conditions and other environmental factors, then later, the lack of sufficient control over participants in a study, and dealing with issues such as the selection of environmental conditions, evaluation performance, data collection and condition control. They also identify nine generic usability attributes: learnability (ease of use), efficiency, memorability, errors, user satisfaction, effectiveness, simplicity, comprehensibility (readability) and learning performance.

Hussain and Kutar (2009) introduce a hierarchical GQM (Goal Question Metric) model to evaluate mobile usability [79]. On the top level, they place three quality characteristics: effectiveness, efficiency and satisfaction. On the middle level, six guidelines are conceptualized: simplicity, accuracy, time taken, features, safety and attractiveness. Eventually, on the bottom, there is a mapping between questions and metrics, which enables the collection of quantitative data in order to evaluate usability.

Kronbauer *et al.* (2012) propose a hybrid model for the evaluation of smartphone application usability [80]. In this study, the hybrid approach blends two methods for data capture, namely, Logging and ESM (Experience Sampling Method). The first one is based on data collection related to user interaction with an application. Using sensors available in smartphones for contextual data collection, such as luminosity intensity and the device's position, allows the performance of statistical analysis regarding usability. The second one is based on the collection of users' feelings towards a specific product through questions. These two methods are respectively used to measure efficiency, effectiveness and satisfaction.

Harrison *et al.* (2013) developed the PACMAD (People At the Centre of Mobile Application Development) usability model, which identifies three major dimensions affecting the overall usability of a mobile application: the user, the task and the context of use [81]. However, the last one plays a crucial role, as an application may be used in multiple and very different contexts (e.g. environment, physical location, user's state or activity performed). The model encompasses seven attributes, which together reflect the usability of an application: effectiveness, efficiency, satisfaction, learnability, memorability, errors and cognitive load. In some studies the model has been adopted to set up testing and evaluation frameworks [82], [83]. The novelty of the model concerns cognitive load as a new usability attribute. The authors claim that it can be observed that users of mobile applications often perform additional tasks, such as walking, while using the mobile device. For this reason, these additional tasks impact the user's performance, arguing by example of a walking user who in the same time is texting a message which reduces walking speed as s/he is concentrating on typing (sending) the message. More recently, cognitive load has been acknowledged [84], or disregarded [85], as one of the usability notions.

Actual usability, located in the frames of the quality-in-use model by Lew and Olsina (2013), comprises effectiveness, efficiency, learnability in use, and communicability [86]. They also emphasize the difference between the context of mobile applications and traditional, desktop or web applications. The context does not only concern hardware limitations (e.g. size of the screen), but also other factors, such as: user activity, day/time of day, location, user profile, device and network performance.

Obviously, there are many more usability models, individually applicable to particular domains, such as mobile banking [87], or healthcare [88]; however, they were excluded from the discussion due to their specific attributes, classified as superior with respect to the others.

III. RESEARCH METHODOLOGY

A systematic literature review (SLR) in its nature differs from traditional narrative reviews by adopting a replicable, scientific and transparent process methodology. By design, it aims to reduce cognitive bias by providing an audit trail of the associated assumptions and procedures, reviewers decisions and conclusions on the one hand, and by identifying and documenting key scientific contributions to a field or question on the other.

In order to provide a body of knowledge on the usability of mobile applications, we performed a systematic literature review by adopting and adapting the approach provided by Kitchenham and Charters [89], [90], since a large majority of the reported SLRs in software engineering has been carried out in respect to their guidelines [91].

According to the research design employed, this study consists of three steps, performed in a fixed sequence. Interdependency is revealed in the one-way output/input relations.

Step 1 in the research methodology involves defining the research questions and the review protocol, which encompasses the data source and search strategy, the inclusion and exclusion criteria and the definition of the search string. The outcome of this step is described in Section 4. Step 2 in the research methodology involves executing the search string carried out on the database engine. Next, the obtained results are extracted and further processed. The outcome of this step is given in Section 5. Step 3 in the research methodology involves reviewing, analysing and reporting each data record, in order to consequently find and document answers for a defined set of the research questions. The outcome of this step is described in Section 6.

IV. SYSTEMATIC LITERATURE REVIEW DEFINITION

A. RESEARCH QUESTIONS DEFINITION

Investigating the gap in usability between desktop and mobile settings, the following three questions arose:

RQ1. How has usability for mobile applications been defined?

RQ2. What are the usability attributes for mobile applications?

RQ3. How have usability attributes for mobile applications been defined, and which measures and evaluation methods have been used?

TABLE 1. The general search query criteria.

Criterium	Search string
Article title, Abstract, Keywords	TITLE-ABS-KEY
Keywords	usability AND "mobile application"
Year	EXCLUDE (PUBYEAR, 2019)

TABLE 2. The inclusion criteria (LIMIT-TO) to the subject area (SUBJAREA).

Subject Area	Search string
Computer science	"comp"
Engineering	"engi"
Mathematics	"math"
Social Sciences	"soci"
Decision Sciences	"deci"

These three interrogative statements provide the overall framework for conducting this study, by giving direction and setting up boundaries.

B. DATA SOURCE AND SEARCH STRATEGY

In line with the research methodology, step 1 involves a systematic search of the scientific literature on the topic of mobile application usability. Performed on Scopus, the largest abstract and citation database of peer-reviewed literature, counting over 71 million records [92], the search strategy aims at identifying indexed publications. A key issue when formulating a search strategy is to define the period of time to set up time boundaries. Being in our interest to obtain reliable and concise answers to the questions, we determined the closing date in December 2018.

C. SEARCH QUERY DEFINITION

The search query was defined by the presence of ''usability'' and the string ''mobile application'' in titles, abstracts and keywords. These unique and specific terms, joined together in that order and in the extent of such meta-data, embody the authors' common declaration of their research objectives and the adopted context of their performed studies. The summary, in terms of the search query construct, is given in Table 1.

D. INCLUSION AND EXCLUSION CRITERIA

In accordance with our research objective and questions, the first applied inclusion criterion relates to the subject area, which alternatively includes: computer science, engineering, mathematics, social sciences, or decision sciences. Table 2 presents the summary of the search query construct in this scope.

In this study, usability is considered in the context of software, which is a concern of computer science and is also closely associated with the other abovementioned disciplines. In this line of thinking, we exclude irrelevant subject areas (e.g. Medicine, Health Professions, Chemistry and others). Table 3 depicts the summary of the search query construct in this scope.

The second inclusion (exclusion) criterion was the document type which alternatively encompasses: conference proceedings, journal articles or book chapters. On the other hand, we did not take into account conference reviews and other reviews, which present non-scientific contributions. Table 4 outlines the summary of the search query in this scope.

TABLE 3. The exclusion criteria (EXCLUDE) to the subject area (SUBJAREA).

Subject Area	Search string
Medicine	"medi"
Health	"heal"
Chemical Engineering	"ceng"
Environmental Science	"envi"
Physics and Astronomy	"phys"
Materials Science	"mate"
Biochemistry et al.	"bioc"
Energy	"ener"
Psychology	"psyc"
Arts and Humanities	"arts"
Earth and Planetary Sciences	"eart"
Nursing	"nurs"
Chemistry	"chem"
Neuroscience	"neur"
Economics et al.	"econ"
Agricultural and Biological Sciences	"agri"
Immunology and Microbiology	"immu"
Pharmacology et al.	"phar"

TABLE 4. The inclusion criteria (LIMIT-TO) for the document type (DOCTYPE).

Document type	Search string
Conference Paper	"cp"
Article	"ar"
Chapter	"ch"

TABLE 5. The inclusion (LIMIT-TO) and exclusion (EXCLUDE) criteria for the language.

Language	Search string
English	LIMIT-TO (LANGUAGE, "English")
Portuguese	EXCLUDE (LANGUAGE, "Portuguese")
French	EXCLUDE (LANGUAGE, "French")

Not all scientists regard conference proceedings as a reliable and valuable source of knowledge. However, from our point of view, our judgement was not solely based on the document type, but on scrupulous reading and conscientious content analysis.

The third inclusion (exclusion) criterion was the language, exclusively limited to English. Therefore, two other (Portuguese and French) were excluded. Table 5 depicts the summary of the search query construct in this regard.

English has become the modern *lingua franca* in the modern world. The major international standardization bodies publish norms and standards in English, and communication channels between experts and communities follow the same rule as well.

V. SEARCH EXECUTION

A. SEARCH AND SELECTION

In the first run, the search query (Table 1) produced 1,615 document results. To this volume, the inclusion

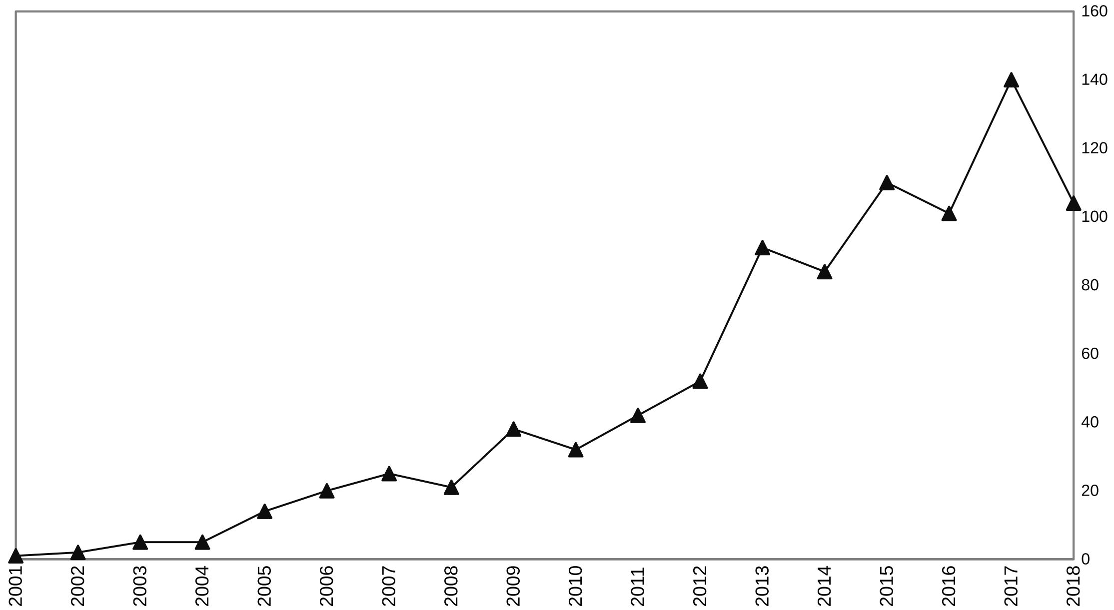

FIGURE 1. The distribution of the number of publications per year.

and exclusion criteria were applied, defined respectively in Tables 2–5. The search strings, given in all these tables, were eventually combined by the relevant Boolean operators. The final search query construct, which entirely fulfils all the requirements, is given below.

TITLE-ABS-KEY (usability AND "mobile application") AND (LIMIT-TO (SUBJAREA, "comp") OR LIMIT-TO (SUBJAREA, "engi") OR LIMIT-TO (SUBJAREA, "math") OR LIMIT-TO (SUBJAREA, "soci") OR LIMIT-TO (SUBJAREA, "deci") OR EXCLUDE (SUBJAREA, "medi") OR EXCLUDE (SUBJAREA, "heal") OR EXCLUDE (SUBJAREA, "ceng") OR EXCLUDE (SUBJAREA, "envi") OR EXCLUDE (SUBJAREA, "phys") OR EXCLUDE (SUBJAREA, "mate") OR EXCLUDE (SUBJAREA, "bioc") OR EXCLUDE (SUBJAREA, "ener") OR EXCLUDE (SUBJAREA, "psyc") OR EXCLUDE (SUBJAREA, "arts") OR EXCLUDE (SUBJAREA, "eart") OR EXCLUDE (SUBJAREA, "nurs") OR EXCLUDE (SUBJAREA, "chem") OR EXCLUDE (SUBJAREA, "neur") OR EXCLUDE (SUBJAREA, "econ") OR EXCLUDE (SUBJAREA, "agri") OR EXCLUDE (SUBJAREA, "immu") OR EXCLUDE (SUBJAREA, "phar")) AND (LIMIT-TO (DOCTYPE, "cp") OR LIMIT-TO (DOCTYPE, "ar") OR LIMIT-TO (DOCTYPE, "ch")) AND (EXCLUDE (PUBYEAR, 2019)) AND (LIMIT-TO (LANGUAGE, "English")) AND (EXCLUDE (LANGUAGE, "Portuguese") OR EXCLUDE (LANGUAGE, "French"))

The results summary was checked in order to verify whether all the criteria were successfully applied. In total, the final search query eventually produced 887 documents, published between 2001 and 2018. The details of the volume data are as follows, while the numbers in brackets indicate the total number of publications: (a) published in English (887), (b) the subject area is from: computer science (803), decision sciences (40), engineering (198), mathematics (197) and social sciences (103), and (c) the document type is: conference proceedings (666), journal articles (196) or book chapters (25). The peak year is 2017 (140), followed by the years 2015 (110), 2018 (104) and 2016 (101), with an average of 74 documents published annually between 2008-2018 (Figure 1).

The distribution of the number of publications increases in linear. However, in 2018 a fall was observed in comparison to the previous year, but still above the year 2016. The majority of documents were published by Springer in Lecture Notes in Computer Science, including sub-series Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics (148), while the largest contributor among journals is the Journal of Telecommunication, Electronic and Computer Engineering (12). The top three countries, the USA (136), Germany (81) and Malaysia (66), accounted for over 31% of the countries the authors were affiliated to.

B. DATA EXTRACTION

Having imported the reference data (authors, document title, year, and digital object identifier) to an external spreadsheet, we systematically searched for each record in full-text databases hosted by particular publishers and indicated as the source of the document. From the list of 887 records, in total 790 (89%) documents were fully available, while using a HAN system licensed account. To extract the data, three independent reviewing procedures were prepared and executed, respectively for each research question.

In the first run, each available document was screened with the aim to identify and recognize a usability definition

TABLE 6. The list of adopted usability definitions for mobile settings.

Source	Definition
ISO 9241-11	the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use [21].
ISO 25010	degree to which a product or system can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use [24].
ISO 9126-1	the capability of the software product to be understood, learned, used and attractive to the user, when used under specified conditions [25].
IEEE Glossary	the ease with which a user can learn to operate, prepare inputs for, and interpret outputs of a system or component [94].
Nielsen	ease of use and correctness of a system for a specific class of users carrying out particular tasks in a precise environment [95].
Bevan	the extent to which a user can easily interact with a product, and it measures how easy to operate, learn and (even) memorize such products are [96].

TABLE 7. The shares of adopted usability definitions for mobile settings.

Definition	Share	References
ISO 9241-11	58 (88%)	[97], [98], [99], [100], [101], [102], [103], [104], [105], [106], [107], [108], [109], [110], [111], [112], [113], [114], [115], [116], [117], [118], [119], [88], [120], [121], [122], [123], [124], [125], [126], [127], [128], [129], [130], [131], [132], [133], [134], [135], [136], [137], [138], [139], [140], [141], [142], [143], [144], [145], [146], [147], [148], [149], [150], [151], [152], [153].
ISO 25010	3 (4,5%)	[154], [155], [156].
ISO 9126-1	2 (3%)	[157], [158].
IEEE Glossary	1 (1,5%)	[159].
Nielsen	1 (1,5%)	[160].
Bevan	1 (1,5%)	[161].

referenced by the author(s). The document was classified as relevant if: (a) usability, as a term, was explicitly defined and (b) correctly referenced. If the authors provided more than one definition and did not indicate a particular one as valid, then the first one given was assumed to be adopted. Eventually, 66 (8%) documents were classified as relevant as the input for analysis, with the aim of formulating an answer to the first research question.

In the second run, each document was screened again to determine the overall quality and its relevance. A document was classified as relevant if: (a) the subject of the research was addressed to the usability of mobile applications, and (b) was not biased by a context of the research, such as: (*i*) application type or (*ii*) user-specific properties, such as: age, occupation, sex or (*iii*) system-specific support features, like visually impaired or disability. The review of the list produced 53 (7%) documents as relevant as the input for analysis with the aim of formulating an answer to the second research question.

In the third run, the above list was reviewed and examined again with the aim of extracting attribute definitions, measures and UEMs. The document was classified as relevant if: (a) usability attributes being the subject of the study were explicitly defined, whereas a measure was valid if it captures the quantitative data which accurately describes one particular usability attribute. Ultimately, 39 (5%) documents were classified as relevant as the input for analysis with the aim of formulating an answer to the third research question.

VI. DATA ANALYSIS

This section addresses the analysis of the data extracted from the studies in accordance with the three defined research questions. We used a qualitative content analysis, which focuses on the characteristics of language as a communication channel, with attention to the specific subjects, narrowed and directed by particular research questions.

RQ1. How has usability for mobile applications been defined?

To this day, none of the authors have introduced any formal definition of usability associated with an application (system) running on a mobile device. Therefore, all identified and recognized definitions have been adopted from the existing general norms, standards and definitions.

The great majority of authors (88%) have defined usability solely in terms of the ISO 9241-11 norm, while others have also made reference to ISO 25010 (4,5%) and ISO 9126 (3%) norms, as well as to the IEEE Glossary (1,5%), the Nielsen (1,5%) and Bevan (1,5%) definitions. Table 6 includes the full text of these six definitions, whereas Table 7 depicts findings of the shares of adopted usability definitions for mobile settings.

RQ2. What are the usability attributes for mobile applications?

In total, 75 usability attributes were identified and analysed. Among them, the most frequent are efficiency (70%), satisfaction (66%) and effectiveness (58%). Less frequent are learnability (45%), memorability (23%), cognitive load (19%) and errors (17%). The last two concern simplicity (13%) and ease of use (9%). The remaining attributes occurred four times or less. Table 8 outlines the details in this regard (the attributes which occurred only once are not included).

TABLE 8. The list of adopted usability attributes for mobile settings.

Attribute	Share	References
efficiency	37 (70%)	[97], [154], [102], [104], [162], [105], [163], [120], [164], [84], [165], [166], [99], [115], [167], [168], [116], [169], [82], [88], [155], [170], [124], [157], [129], [171], [83], [172], [173], [86], [174], [175], [176], [177], [178], [78], [153].
satisfaction	35 (66%)	[97], [104],[179], [162], [105], [163], [120], [164], [84], [165], [166], [180], [115],[168], [116], [169], [126], [118], [82], [88], [155], [181], [124], [129], [171], [182], [83], [172], [173], [174], [175], [183], [177], [178], [78].
effectiveness	31 (58%)	[97], [154], [104], [162], [163], [120], [164], [84], [165], [99], [115], [167], [168], [116], [169], [126], [82], [88], [170], [124], [157], [129], [171], [83], [172],[173], [86], [174], [175], [177], [78].
learnability	24 (45%)	[179], [184], [103], [162], [105], [84], [166], [180], [115], [168], [185], [169], [126], [82], [88], [155], [156], [129], [171], [83], [86], [175], [178], [78].
memorability	12 (23%)	[105], [84], [186], [166], [168], [169], [82], [88], [155], [171], [83], [78].
cognitive load	10 (19%)	[184], [84], [115], [168], [185], [169], [82], [88], [83], [153].
errors	9 (17%)	[105], [84], [166], [168], [169], [82], [171], [83], [78].
simplicity	7 (13%)	[186], [168], [118], [88], [171], [175], [78].
ease of use	5 (9%)	[154], [179], [155], [157], [178].
navigation	4 (8%)	[103], [118], [155], [153].
operability	4 (8%)	[162], [169], [156], [187].
usefulness	4 (8%)	[179], [88], [157],[175].
attractiveness	3 (6%)	[162], [186], [169].
comprehensibility	3 (6%)	[171], [175], [78].
aesthetics	3 (6%)	[119], [156], [187].
accessibility	2 (4%)	[156], [187].
accuracy	2 (4%)	[186], [169].
adaptability	2 (4%)	[184], [175].
consistency	2 (4%)	[103], [185].
interaction	2 (4%)	[118], [155].
learning performance	2 (4%)	[171], [78].
training	2 (4%)	[162], [188].
understandability	2 (4%)	[162], [169].
user error protection	2 (4%)	[156], [187].

RQ3. How have usability attributes for mobile applications been defined, and which measures and evaluation methods have been used?

The foremost attribute, *efficiency* is the ability of a user to complete a task with speed and accuracy. Efficiency is measured in a number of ways, such as the duration spent on each screen, the duration to complete a given task (a set of tasks), and the user's error rate. Two evaluation methods are used: controlled observation and survey.

Satisfaction is a user's perceived level of comfort and pleasure, or a user's perceived level of fulfilment of his expectations and needs. Satisfaction is measured only by using survey, with predefined statements with the Likert-scale rating system, which is typically used to capture a user's intangible attitude towards an application.

Effectiveness is the ability of a user to complete a task in a given context. It is measured by the number of successfully completed tasks, the number of steps required to complete a task, the number of double taps unrelated to the operation of an application, and the number of times that a back button is used by the mobile device (not the application).

Learnability is defined twofold. *First-time* learnability refers to the degree of ease with which a user can interact with a newly-encountered system without getting guidance or referring to documentation. It is measured by the number of attempts to solve a task, the number of assists during performing a task, and the number of errors performed by a user. Learnability *over time*, on the contrary, is the capacity of a user to achieve proficiency with an application. Typically, a user's performance during a series of tasks is observed to measure how long it takes these participants to reach a pre-specified level of proficiency. Similarly to effectiveness, two evaluation methods are used: controlled observation and survey.

Memorability is the degree of ease with which a user can remember how to use an application effectively. It is measured by asking users to perform a series of tasks after having become proficient with the use of the application, and afterwards asking them to perform similar tasks after a period time of inactivity. To determine how memorable the application was, a comparison is made between the two sets of results. In this case, the eye-tracking technique is also used as the method to collect gaze data which is further used to evaluate usability.

Cognitive load refers to the amount of mental activity imposed on a user's working memory during application usage. Cognitive load theory differentiates cognitive load into three types: extraneous, intrinsic and germane. Firstly, *extraneous* cognitive load refers to instructional and presentation schemas, caused by the mental activities and elements that do not directly support application usage. Secondly, *intrinsic* cognitive load refers to the task complexity, caused by

TABLE 9. The top most frequent usability attributes, their measures and associated usability evaluation methods (UEMs) for mobile settings.

Attribute	Measures	UEMs & References
efficiency	– duration spent on each screen, – duration to complete task, – user's error rate,	controlled observation [102], [189], [105], [165], [168], [116], [82], [177]; survey [162], [164], [84], [155], [178].
satisfaction	– predefined statements with Likert-scale rating,	survey [189], [162], [105], [164], [84], [165], [180], [168], [116], [118], [82], [88], [155], [174], [177], [178].
effectiveness	– number of successfully completed tasks, – number of steps required to complete a task, – number of taps related to an app usage, – number of taps unrelated to an app usage, – number of times that a back button was used;	controlled observation [189], [165], [168], [116], [82], [177]; survey [162], [164], [84].
learnability	first-time: – number of attempts to solve a task, – number of assists during performing a task, – number of errors performed by a user; over time: – repeating similar pairs of tasks in each session, – duration to reach a pre-specified proficiency,	controlled observation [105], [180], [168], [82], [155]; survey [162], [84], [155], [178].
memorability	– effects of response time, – duration of pauses, – predefined statements with Likert-scale rating;	controlled observation [105], [168], [82]; eye tracking [106], survey [84], [186], [155].
cognitive load	– fixations, gaze points and heat maps, – text summarization, – predefined statements with Likert-scale rating;	controlled observation [168], eye tracking [143], thinking aloud [82], survey [84].
errors	– amount and type of errors occurred, – an application ability to recover from errors;	controlled observation [105], [168], [116], [178]; survey [84].
simplicity	– the number of menu levels, – the number of performed gestures to reach a destination object – the duration of searching a button to perform a specific function, – predefined statements with Likert-scale rating;	survey [186], [168], [118], [88]; controlled observation [168], [78]; interview [175].
ease of use	– predefined statements with Likert-scale rating,	survey [155], [178].

the number of elements in a task and the degree to which these elements are related to each other. Thirdly, *germane* cognitive load refers to the amount of mental effort used to form schemas and actively integrate new information with prior knowledge during application usage. In the practice of cognitive load measurement, instruments such as a subjective rating scale, a thinking aloud dual task protocol or eye tracking are in common use.

Errors refers to the amount and type of errors which occur during task performance by a user. On the other hand, it is the ability of an application to recover from occurred errors. Both these definitions also respectively reflect the measures of attribute.

Simplicity is the degree of being easy to understand or being uncomplicated in form or design, described by such characteristics as the number of menu levels, the number of performed gestures to reach a destination object, and the duration of searching a button to perform a specific function. On the other hand, simplicity is the level of comfort with which a user is able to complete a task, measured by predefined statements with the Likert-scale rating.

Ease of use is the perceived level of the user's effort related to usage of the application. The survey instrument is used to collect data from users on perceptions concerning their experienced interaction with the application.

Table 9 presents a summary in which each attribute is associated with the valid measures, along with the usability evaluation methods used to collect the necessary data to improve particular software artefacts.

From the variety of available methods, the most frequent is survey, based on the questionnaire instrument, which has been used to collect data from a sample of the participants, as a representation of the population of interest. Controlled observation of the user while interacting with an application is the second most frequent method applied to usability evaluation. The remaining three, namely eye-tracking, thinking aloud and interview, are hardly used and serve as complementary to collect additional data. Table 10 presents the details showing the number of occurrences of all identified UEMs applied for particular attributes.

VII. DISCUSSION

Based on the obtained results, we argue that the ISO 9241-11 definition has been widely accepted in a non-changed form, and since the inception of research on mobile application usability, has been, de facto, popularized as the standard by the HCI community. Having said that, it is worth noting that other definitions are not contrary to each other. Moreover, they have in common the software capability to interact with a user, yet emphasize different aspects of his/her proficiency. On the other hand, usability is always associated with the product, except for the IEEE Glossary and Bevan definitions, which focus first and foremost on the user.

TABLE 10. The number of occurrences of usability evaluation methods (UEMs) applied to particular usability attributes.

Attribute	Survey	Controlled observation	Eye tracking	Thinking aloud	Interview
efficiency	5	8	0	0	0
satisfaction	15	0	0	0	0
effectiveness	3	6	0	0	0
learnability	4	5	0	0	0
memorability	3	3	1	1	0
cognitive load	1	1	1	1	0
errors	1	4	0	0	0
simplicity	4	2	0	0	1
ease of use	2	0	0	0	0

TABLE 11. The percentage of studies concerning usability attributes of mobile applications.

Attribute	Coursaris (2007)	Harrison (2013)	This study (2020)
effectiveness	62%	51%	58%
efficiency	33%	55%	70%
satisfaction	20%	58%	66%
errors	0%	33%	17%
learnability	11%	21%	46%
cognitive load	0%	23%	19%
accessibility	7%	0%	4%
memorability	2%	2%	23%
operability	4%	0%	8%
flexibility	2%	0%	0%
acceptability	2%	0%	0%

The most frequent attributes originate from the usability definition adapted to mobile applications. In such a case, the main usability characteristics are device-agnostic. In other words, efficiency, satisfaction and effectiveness are valid for studying the usability of both desktop and mobile applications. In a similar manner, however with minor extensions, the remaining attributes have been assimilated as well.

By design, cognitive load is related to the mental effort required by the user to perform tasks using a mobile device. While it is neither novel, nor high-ranking in usability research, it has now gained a larger audience due to the fact that a user's attention is usually divided among other simultaneously performed tasks.

If one breaks down usability into two parts, one gets two nouns: "use" and "ability". According to this line of thinking, the ability to use an application, in particular, means the ability to learn, memorize, navigate and operate. Besides this, ease of use seems related to the sense of presence of these abilities, facilitated by the errorless behaviour of the application.

If one takes into account the research methods applied to the problem of usability evaluation, the attributes of the studied phenomena can be divided into two groups: quantitative and qualitative. However, based on such criteria, every attempt to formulate distinct groups will have its pros and cons, because each attribute has been measured depending on observation and survey. Nevertheless, the existing measures can be unambiguously classified if one still makes a clear distinction between facts and opinions. In other words, quantitative-oriented attributes have the advantage of being clearly definable and objectively measurable, using measures that are not influenced by the user's personal judgement.

On the other hand, one can point to user-oriented measures, and on the contrary, to application-oriented measures. Last but not least, it appears that existing measures intertwine user and application performance in one integrated artefact. It seems obvious that observational data are required to discover an application's bottlenecks and general areas for improvement, thereby optimizing its operational capabilities by reducing the time and effort involved in its usage.

To collect quantitative and qualitative data, questionnaires and controlled observation, respectively, have been typically applied, occasionally supported by eye-tracking, thinking aloud and interview techniques (see Table 10). In order to obtain numerical measures, a retrospective audio/video analysis is performed, while in some studies, third-party tools were installed which log all user interaction with an application with the date and time of the event, including the buttons that they chose, the gestures that were made and the functions that were recalled. After completing the task scenario, a user was asked to rank their agreement (disagreement) with predefined statements on a Likert scale or other rating scale.

In comparison with the results obtained from studies with similar objectives, conducted by Coursaris and Kim [93] and Harrison *et al.* [81], our findings are consistent in the extent of the top three attributes, which concern efficiency, satisfaction and effectiveness (see Table 11). An increased interest in learnability and memorability can also be noticed, while errors and cognitive load are less appreciated. While simplicity and ease of use have not been indicated before, and being complementary, are neither novel nor visionary. The rest of the attributes seem to be extensions of existing ones, however, unbigoted by usability, they correspond to both explicit and implicit application properties. On the other hand, if one explores the user's preferences instead of his/her ability to use an application, the results refer more to the user experience domain, yet less to usability.

A. LIMITATIONS

Although this study contributes to the field of human-computer interaction, certain limitations exist within the research design.

Firstly, one of the major limitations is that only one data source was involved. However, the indexation process covers varied sources of scientific content, ranging from conference proceedings to journal articles, which are reviewed each year to ensure quality standards are maintained.

Secondly, inclusion and exclusion criteria, as they permit documents only published in English, may be a subject of critique. In this manner, our intention was not to disregard other foreign languages, but was determined by the global status of the English language in modern science.

Thirdly, regarding the search query construct, including only the terms "usability" and "mobile applications" might have excluded potentially relevant documents concerning other related studies (i.e. user experience or design thinking) from the scope of the search results, and later, from the analysis, though one should bear in mind that evidence is defined as the "synthesis of the best quality scientific studies on a specific topic (...)" [90]. Nevertheless, by design, the goal of this study was to provide an evidence-based contribution on the usability of mobile applications, thus this limitation is simply the result of the application of SLR methodology principals.

Ultimately, the applied reviewing procedures might be seen as too strict or hard to follow. However, we assumed to identify only such attributes and measures which can be replicated in any extent, and arbitrarily extended if necessary.

B. FUTURE RESEARCH

The obtained results uncover the trend in time of producing "new" attributes which unnecessarily contribute to the usability of mobile applications. And yet, one might try to assume that there are still some vulnerable properties laying beneath their quality of use. Nevertheless, one of the issues which unfolds definitely concerns how to consolidate the existing attributes into one compact model which reflects all identified and relevant usability facets.

Moreover, in addressing the topic of usability evaluation, there is still little known about the simulation methods which might admittedly replace both experts and users in application evaluation in view of selecting its properties and behaviour. On the one hand, software vendors will benefit by reducing engaged time and effort, while on the other, the users will take advantage of the better application in daily usage. Therefore, a second suggestion is practical in nature, and relates to the matter of developing a tool in which implemented methods to automate and simulate the users' interaction with the application may reduce the participation of both experts and users.

Currently, we are developing a usability inspection method which aims to fully automate application testing due to evaluating its compliance with efficiency and effectiveness requirements and to detecting bugs and errors. The latest application version enables usability engineers to perform video analysis annotation, which aims at measuring the duration of actions on a time scale, embedded on a ribbon within a visual diagram editor. Moreover, it also allows the tasks on the layers to be graphically decomposed into smaller units (subtasks). The first results are promising, showing that if we isolate user activities from application responses, then it allows us to analyse and evaluate both the user and the application separately, which adequately produces a reliable outcome for interface designers and developers.

VIII. CONCLUSION

The results of the systematic literature review show that the ISO 9241-11 definition has been adapted by the majority as the baseline in the studies of mobile application usability. In total, 75 attributes were distinguished in the body of 790 documents, indexed by the Scopus database. The most frequent are efficiency (70%), satisfaction (66%) and effectiveness (58%), which originate from the above definition. Afterwards, the less frequent are learnability (45%), memorability (23%), cognitive load (19%) and errors (17%). The last two concern simplicity (13%) and easy of use (9%). The remaining attributes occurred four times or less.

We observed that 91% of documents lack a usability definition. While not providing a formal and legitimate definition even seems to be acceptable in some circumstances, measuring and explaining the facets of the phenomena exclusively on the grounds of common sense might be questionable. As a matter of fact, over 90% of documents did not meet the inclusion criteria for analysis, although some report valuable results. On the other hand, a small number of the remaining documents zealously "produced" new attributes with associated "unique" measures, which usually concerned unobservable properties, measured by a set of explicit statements. Many of these constructs lack theoretical foundations and empirical evidence to expose their worthiness.

To complicate the matter even further, most of the introduced attributes have focused on user beliefs, emotions, preferences, perceptions, physical and psychological responses, behaviours and accomplishments that occur before, during and after application use, which concern, in particular or as a whole, dimensions of user experience [21]. Such a combination of objective and subjective assessments eventually produces an outcome which refers to neither application usability nor user experience. This reflects an ignorance to methodological rigour, which negatively affects the validity of the results.

REFERENCES

[1] M. Rouse and I. Wigmore. (2013). *Definition. Mobile App*. [Online]. Available: https://whatis.techtarget.com/definition/mobile-app

[2] Statista. (2019). *Global Mobile Device Market Share in 2017 and 2018, by Device Type*. [Online]. Available: https://www.statista.com/statistics/183530/worldwide-market-share-mobile-device/

[3] Statista. (2019). *Smartphone User Penetration as Percentage of Total Global Population From 2014 to 2021*. [Online]. Available: https://www.statista.com/statistics/203734/global-smartphone-penetration-per-capita-since-2005/

[4] H. M. Salman, W. F. Wan Ahmad, and S. Sulaiman, ''Usability evaluation of the smartphone user interface in supporting elderly users from experts' perspective,'' *IEEE Access*, vol. 6, pp. 22578–22591, 2018.

[5] M. Sikorski, ''Beyond product usability: User satisfaction and quality management,'' in *Proc. CHI Extended Abstr. Hum. Factors Comput. Syst. (CHI)*. New York, NY, USA: ACM, 2000, pp. 61–62.

[6] H. Feng, T. Hoegler, and W. Stucky, ''Exploring the critical success factors for mobile commerce,'' in *Proc. Int. Conf. Mobile Bus.*, Jun. 2006, p. 40.

[7] E. Garcia-Lopez, A. Garcia-Cabot, C. Manresa-Yee, L. de-Marcos, and C. Pages-Arevalo, ''Validation of navigation guidelines for improving usability in the mobile Web,'' *Comput. Standards Interfaces*, vol. 52, pp. 51–62, May 2017.

[8] N. Tractinsky, ''The usability construct: A dead end?'' *Hum.-Comput. Interact.*, vol. 33, no. 2, pp. 131–177, Mar. 2018.

[9] B. Shackel, ''Usability—Context, framework, definition, design and evaluation,'' *Interacting Comput.*, vol. 21, nos. 5–6, pp. 339–346, Dec. 2009.

[10] J. Kjeldskov and J. Stage, ''New techniques for usability evaluation of mobile systems,'' *Int. J. Hum.-Comput. Stud.*, vol. 60, nos. 5–6, pp. 599–620, May 2004.

[11] K. Gulzar, J. Sang, M. Ramzan, and M. Kashif, ''Fuzzy approach to prioritize usability requirements conflicts: An experimental evaluation,'' *IEEE Access*, vol. 5, pp. 13570–13577, 2017.

[12] P. Weichbroth, ''Delivering usability in IT products: Empirical lessons from the field,'' *Int. J. Softw. Eng. Knowl. Eng.*, vol. 28, no. 7, pp. 1027–1045, Jul. 2018.

[13] M. S. Bashir and A. Farooq, ''EUHSA: Extending usability heuristics for smartphone application,'' *IEEE Access*, vol. 7, pp. 100838–100859, 2019.

[14] R. Parente Da Costa, E. D. Canedo, R. T. De Sousa, R. De Oliveira Albuquerque, and L. J. Garcia Villalba, ''Set of usability heuristics for quality assessment of mobile applications on smartphones,'' *IEEE Access*, vol. 7, pp. 116145–116161, 2019.

[15] A. Seffah, M. Donyaee, R. B. Kline, and H. K. Padda, ''Usability measurement and metrics: A consolidated model,'' *Softw. Qual. J.*, vol. 14, no. 2, pp. 159–178, Jun. 2006.

[16] S. Winter, S. Wagner, and F. Deissenboeck, ''A comprehensive model of usability,'' in *Proc. IFIP Int. Conf. Eng. Hum.-Comput. Interact.* Berlin, Germany: Springer, 2007, pp. 106–122.

[17] K. Majrashi and M. Hamilton, ''A cross-platform usability measurement model,'' *Lect. Notes Softw. Eng.*, vol. 3, no. 2, p. 132, 2015.

[18] P. Weichbroth, ''Usability attributes revisited: A time-framed knowledge map,'' in *Proc. Federated Conf. Comput. Sci. Inf. Syst.*, Sep. 2018, pp. 1005–1008.

[19] *Software Enginnering—Product Quality*, Standard ISO/IEC 9126:1991, ISO/IEC, Geneva, Switzerland, 1991.

[20] *Ergonomic Requirements for Office Work With Visual Display Terminals (VDTs)—Part 11: Guidance on Usability*, Standard ISO 9241-11:1998(en), ISO, Geneva, Switzerland, 1998. [Online]. Available: https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-1:v1:en

[21] *Ergonomics of Human-System Interaction—Part 11: Usability: Definitions and Concepts*, ISO 9241-11:2018(en), ISO, Geneva, Switzerland, 2018. [Online]. Available: https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-2:v1:en

[22] J. R. Lewis, ''Usability: Lessons learned... and yet to be learned,'' *Int. J. Hum.-Comput. Interact.*, vol. 30, no. 9, pp. 663–684, Sep. 2014.

[23] D. Quiñones and C. Rusu, ''How to develop usability heuristics: A systematic literature review,'' *Comput. Standards Interfaces*, vol. 53, pp. 89–122, Aug. 2017.

[24] *Systems and Software Engineering—Systems and Software Quality Requirements and Evaluation (SQuaRE)—System and Software Quality Models*, Standard ISO/IEC 25010:2011, ISO, Geneva, Switzerland, 2011.

[25] *Software Engineering—Product Quality—Part 1: Quality Model*, Standard ISO/IEC 9126-1:2001, ISO, Geneva, Switzerland, 2001. [Online]. Available: https://www.iso.org/standard/22749.html

[26] J. Nielsen, ''Iterative user-interface design,'' *Computer*, vol. 26, no. 11, pp. 32–41, Nov. 1993.

[27] N. Bevan, ''Usability is quality of use,'' in *Advances in Human Factors/Ergonomics*, vol. 20. Amsterdam, The Netherlands: Elsevier, 1995, pp. 349–354.

[28] J. M. C. Bastien, ''Usability testing: A review of some methodological and technical aspects of the method,'' *Int. J. Med. Inform.*, vol. 79, no. 4, pp. e18–e23, Apr. 2010.

[29] T. Boren and J. Ramey, ''Thinking aloud: Reconciling theory and practice,'' *IEEE Trans. Prof. Commun.*, vol. 43, no. 3, pp. 261–278, 3rd Quart., 2000.

[30] S. Richardson, R. Mishuris, A. O'Connell, D. Feldstein, R. Hess, P. Smith, L. McCullagh, T. McGinn, and D. Mann, '''Think aloud' and 'Near live' usability testing of two complex clinical decision support tools,'' *Int. J. Med. Inform.*, vol. 106, pp. 1–8, Oct. 2017.

[31] G. Deniz and P. O. Durdu, ''A comparison of mobile form controls for different tasks,'' *Comput. Standards Interfaces*, vol. 61, pp. 97–106, Jan. 2019.

[32] T. Kato, ''What 'question-asking protocols' can say about the user interface,'' *Int. J. Man-Mach. Stud.*, vol. 25, no. 6, pp. 659–673, Dec. 1986.

[33] T. Obata, T. Daimon, and H. Kawashima, ''A cognitive study of in-vehicle navigation systems: Applying verbal protocol analysis to usability evaluation,'' in *Proc. Vehicle Navigat. Inf. Syst. Conf. (VNIS)*, Oct. 1993, pp. 232–237.

[34] R. Nagpal, D. Mehrotra, and P. K. Bhatia, ''The state of art in Website usability evaluation methods,'' in *Design Solutions for User-Centric Information Systems*. Harrisburg, PA, USA: IGI Global, 2017, pp. 275–296.

[35] et. al. R. Nagpal, ''FAHP approach to rank educational Websites on usability,'' *Int. J. Comput. Digit. Syst.*, vol. 4, no. 4, pp. 251–260, Oct. 2015.

[36] A. Hinderks, M. Schrepp, F. J. Domínguez Mayo, M. J. Escalona, and J. Thomaschewski, ''Developing a UX KPI based on the user experience questionnaire,'' *Comput. Standards Interfaces*, vol. 65, pp. 38–44, Jul. 2019.

[37] J.-P. Jain and W.-C. Shen, ''A study of the optimization of app design with regards to usability and user experience—a case study of SunlineApp,'' in *Proc. Eng. Innov. Design, 7th Int. Conf. Innov., Commun. Eng. (ICICE)*, Hangzhou, China. CRC Press, May 2019, p. 175.

[38] A. Holzinger, ''Usability engineering methods for software developers,'' *Commun. ACM*, vol. 48, no. 1, pp. 71–74, Jan. 2005.

[39] L. F. Gonçalves, L. G. Vasconcelos, E. V. Munson, and L. A. Baldochi, ''Supporting adaptation of Web applications to the mobile environment with automated usability evaluation,'' in *Proc. 31st Annu. ACM Symp. Appl. Comput. (SAC)*. New York, NY, USA: ACM, 2016, pp. 787–794.

[40] J. Grigera, A. Garrido, J. M. Rivero, and G. Rossi, ''Automatic detection of usability smells in Web applications,'' *Int. J. Hum.-Comput. Stud.*, vol. 97, pp. 129–148, Jan. 2017.

[41] J. Falkowska, J. Sobecki, and M. Pietrzak, ''Eye tracking usability testing enhanced with EEG analysis,'' in *Proc. Int. Conf. Design, User Exper., Usability*. Cham, Switzerland: Springer, 2016, pp. 399–411.

[42] P. Weichbroth, K. Redlarski, and I. Garnik, ''Eye-tracking Web usability research,'' in *Proc. Federated Conf. Comput. Sci. Inf. Syst.*, Oct. 2016, pp. 1681–1684.

[43] P. Realpe-Muñoz, C. A. Collazos, J. Hurtado, T. Granollers, J. Muñoz-Arteaga, and J. Velasco-Medina, ''Eye tracking-based behavioral study of users using e-voting systems,'' *Comput. Standards Interfaces*, vol. 55, pp. 182–195, Jan. 2018.

[44] T. Tullis, S. Fleischman, M. McNulty, C. Cianchette, and M. Bergel, ''An empirical comparison of lab and remote usability testing of Web sites,'' in *Proc. Usability Professionals Assoc. Conf.*, 2002, pp. 1–8.

[45] J. Dong, S. L. Martin, J. M. Mullaly, and A. R. Tannenbaum, ''Method, system and program for performing remote usability testing,'' U.S. Patent 6 526 526, Feb. 25, 2003.

[46] J. Sauer, A. Sonderegger, K. Heyden, J. Biller, J. Klotz, and A. Uebelbacher, ''Extra-laboratorial usability tests: An empirical comparison of remote and classical field testing with lab testing,'' *Appl. Ergonom.*, vol. 74, pp. 85–96, Jan. 2019.

[47] M. J. Scott, F. Spyridonis, and G. Ghinea, ''Designing for designers: Towards the development of accessible ICT products and services using the VERITAS framework,'' *Comput. Standards Interfaces*, vol. 42, pp. 113–124, Nov. 2015.

[48] J. Díaz, C. Rusu, and C. A. Collazos, ''Experimental validation of a set of cultural-oriented usability heuristics: E-commerce Websites evaluation,'' *Comput. Standards Interfaces*, vol. 50, pp. 160–178, Feb. 2017.

[49] T. C. Lacerda and C. G. von Wangenheim, "Systematic literature review of usability capability/maturity models," *Comput. Standards Interfaces*, vol. 55, pp. 95–105, Jan. 2018.

[50] D. Quiñones, C. Rusu, and V. Rusu, "A methodology to develop usability/user experience heuristics," *Comput. Standards Interfaces*, vol. 59, pp. 109–129, Aug. 2018.

[51] I. Connell, A. Blandford, and T. Green, "CASSM and cognitive walkthrough: Usability issues with ticket vending machines," *Behav. Inf. Technol.*, vol. 23, no. 5, pp. 307–320, Sep. 2004.

[52] T. Hollingsed and D. G. Novick, "Usability inspection methods after 15 years of research and practice," in *Proc. 25th Annu. ACM Int. Conf. Design Commun. (SIGDOC)*. New York, NY, USA: ACM, 2007, pp. 249–255.

[53] L. Alonso-Virgós, J. P. Espada, and R. G. Crespo, "Analyzing compliance and application of usability guidelines and recommendations by Web developers," *Comput. Standards Interfaces*, vol. 64, pp. 117–132, May 2019.

[54] Z. Zhang, V. Basili, and B. Shneiderman, "Perspective-based usability inspection: An empirical validation of efficacy," *Empirical Softw. Eng.*, vol. 4, no. 1, pp. 43–69, 1999.

[55] M. N. Mahrin, P. Strooper, and D. Carrington, "Selecting usability evaluation methods for software process descriptions," in *Proc. 16th Asia–Pacific Softw. Eng. Conf.*, Dec. 2009, pp. 523–529.

[56] C. Wilson, *User Interface Inspection Methods: A User-Centered Design Method*. London, U.K.: Newnes, 2013.

[57] J. Vanderdonckt and A. Beirekdar, "Automated Web evaluation by guideline review," *J. Web Eng.*, vol. 4, no. 2, pp. 102–117, 2005.

[58] M. Speicher, A. Both, and M. Gaedke, "Towards metric-based usability evaluation of online Web interfaces," in *Proc. Mensch Comput. Workshopband*, 2013, pp. 277–281.

[59] T. S. D. Silva, M. S. Silveira, and F. Maurer, "Usability evaluation practices within agile development," in *Proc. 48th Hawaii Int. Conf. Syst. Sci.*, Jan. 2015, pp. 5133–5142.

[60] G. Büyüközkan, "Determining the mobile commerce user requirements using an analytic approach," *Comput. Standards Interfaces*, vol. 31, no. 1, pp. 144–152, Jan. 2009.

[61] K. Curcio, A. Malucelli, S. Reinehr, and M. A. Paludo, "An analysis of the factors determining software product quality: A comparative study," *Comput. Standards Interfaces*, vol. 48, pp. 10–18, Nov. 2016.

[62] I. Arroyo, F. Giné, C. Roig, and A. Granollers, "Performance and usability tradeoff in a cluster display wall," *Comput. Standards Interfaces*, vol. 62, pp. 53–63, Feb. 2019.

[63] A. Darejeh and D. Singh, "An investigation on ribbon interface design guidelines for people with less computer literacy," *Comput. Standards Interfaces*, vol. 36, no. 5, pp. 808–820, Sep. 2014.

[64] A. Shrestha, A. Cater-Steel, M. Toleman, and T. Rout, "Benefits and relevance of international standards in a design science research project for process assessments," *Comput. Standards Interfaces*, vol. 60, pp. 48–56, Nov. 2018.

[65] P. Efe and O. Demirors, "A change management model and its application in software development projects," *Comput. Standards Interfaces*, vol. 66, Oct. 2019, Art. no. 103353.

[66] S. Caplan, "Using focus group methodology for ergonomic design," *Ergonomics*, vol. 33, no. 5, pp. 527–533, May 1990.

[67] J. Nielsen, "The use and misuse of focus groups," *IEEE Softw.*, vol. 14, no. 1, pp. 94–95, 1997.

[68] A. Stefani and M. Xenos, "Weight-modeling of B2C system quality," *Comput. Standards Interfaces*, vol. 33, no. 4, pp. 411–421, Jun. 2011.

[69] A. Sutcliffe, "Task analysis, systems analysis and design: Symbiosis or synthesis?" *Interacting Comput.*, vol. 1, no. 1, pp. 6–12, Apr. 1989.

[70] D. Benyon, "Adaptive systems: A solution to usability problems," *User Model. User-Adapted Interact.*, vol. 3, no. 1, pp. 65–87, 1993.

[71] A. Crystal and B. Ellington, "Task analysis and human-computer interaction: Approaches, techniques, and levels of analysis," in *Proc. 10th Americas Conf. Inf. Syst. (AMCIS)*, 2004, p. 391.

[72] G. de Haan, G. C. van der Veer, and J. C. van Vliet, "Formal modelling techniques in human-computer interaction," *Acta Psychol.*, vol. 78, nos. 1–3, pp. 27–67, Dec. 1991.

[73] A. Badre, *Shaping Web Usability: Interaction Design in Context*. Reading, MA, USA: Addison-Wesley, 2002.

[74] S. R. Humayoun, Y. Dubinsky, and T. Catarci, "A three-fold integration framework to incorporate user–centered design into agile software development," in *Proc. Int. Conf. Hum. Centered Design*. Berlin, Germany: Springer, 2011, pp. 55–64.

[75] S. K. Card, A. Newell, and T. P. Moran, *The Psychology of Human-Computer Interaction*. Trenton, NJ, USA: Lawrence Erlbaum Associates Inc., 1983.

[76] S. A. Oyewole and J. M. Haight, "Determination of optimal paths to task goals using expert system based on GOMS model," *Comput. Hum. Behav.*, vol. 27, no. 2, pp. 823–833, Mar. 2011.

[77] D. Fernandez-Lanvin, J. D. Andres-Suarez, M. Gonzalez-Rodriguez, and B. Pariente-Martinez, "The dimension of age and gender as user model demographic factors for automatic personalization in e-commerce sites," *Comput. Standards Interfaces*, vol. 59, pp. 1–9, Aug. 2018.

[78] D. Zhang and B. Adipat, "Challenges, methodologies, and issues in the usability testing of mobile applications," *Int. J. Hum.-Comput. Interact.*, vol. 18, no. 3, pp. 293–308, Jul. 2005.

[79] A. Hussain and M. Kutar, "Usability metric framework for mobile phone application," in *Proc. 10th Annu. PostGraduate Symp. Converg. Telecommun., Netw. Broadcast.*, 2009, pp. 978–982.

[80] A. H. Kronbauer, C. A. S. Santos, and V. Vieira, "Smartphone applications usability evaluation: A hybrid model and its implementation," in *Proc. Int. Conf. Hum.-Centred Softw. Eng.* Berlin, Germany: Springer, 2012, pp. 146–163.

[81] R. Harrison, D. Flood, and D. Duce, "Usability of mobile applications: Literature review and rationale for a new usability model," *J. Interact. Sci.*, vol. 1, no. 1, p. 1, 2013.

[82] I. do Nascimento Mendes and A. C. Dias-Neto, "A process-based approach to test usability of multi-platform mobile applications," in *Proc. Int. Conf. Design, User Exper., Usability*. Cham, Switzerland: Springer, 2016, pp. 456–468.

[83] A. Saleh, R. B. Isamil, and N. B. Fabil, "Extension of pacmad model for usability evaluation metrics using goal question metrics (GQM) approach," *J. Theor. Appl. Inf. Technol.*, vol. 79, no. 1, pp. 90–100, 2015.

[84] N. Parsazadeh, R. Ali, and M. Rezaei, "A framework for cooperative and interactive mobile learning to improve online information evaluation skills," *Comput. Edu.*, vol. 120, pp. 75–89, May 2018.

[85] E. A. Khan and M. K. Y. Shambour, "An analytical study of mobile applications for Hajj and Umrah services," *Appl. Comput. Informat.*, vol. 14, no. 1, pp. 37–47, Jan. 2018.

[86] P. Lew and L. Olsina, "Relating user experience with MobileApp quality evaluation and design," in *Proc. Int. Conf. Web Eng.* Cham, Switzerland: Springer, 2013, pp. 253–268.

[87] H. I. Abubakar, N. Hashim, and A. Hussain, "Verification process of usability evaluation model for m-banking application," in *Proc. Recent Adv. Comput. Sci. 14th Int. Conf. Appl. Comput. Appl. Comput. Sci. (ACACOS)*, 2015, pp. 23–25.

[88] E. Kaur and P. D. Haghighi, "A context-aware usability model for mobile health applications," in *Proc. 14th Int. Conf. Adv. Mobile Comput. Multi Media (MoMM)*. New York, NY, USA: ACM, 2016, pp. 181–189.

[89] B. Kitchenham, "Procedures for performing systematic reviews," *Keele, UK, Keele Univ.*, vol. 33, no. 2004, pp. 1–26, 2004.

[90] B. Kitchenham and S. Charters, "Guidelines for performing systematic literature reviews in software engineering (version 2.3)," Keele Univ. Durham Univ., Keele, U.K., & Durham, U.K., Tech. Rep., EBSE-2007-01, 2007.

[91] H. Zhang and M. Ali Babar, "Systematic reviews in software engineering: An empirical investigation," *Inf. Softw. Technol.*, vol. 55, no. 7, pp. 1341–1354, Jul. 2013.

[92] *Scopus. An Eye on Global Research: 5,000 Publishers. Over 71 M Records and 23,700 Titles*, Amsterdam, The Netherlands, Elsevier, 2018. [Online]. Available: https://www.elsevier.com/__data/assets/pdf_file/0017/114533/SC_FS_overview_WEB.pdf

[93] C. K. Coursaris and D. Kim, "A research agenda for mobile usability," in *Proc. CHI Extended Abstr. Hum. Factors Comput. Syst. (CHI)*. New York, NY, USA: ACM, 2007, pp. 2345–2350.

[94] *IEEE Standard Glossary of Software Engineering Terminology*, IEEE Standard 610.12-1990, 1990, pp. 1–84.

[95] J. Nielsen, "Durability of usability guidelines," Alertbox, Bengaluru, India, Tech. Rep., Jan. 2005.

[96] N. Bevan, "International standards for HCI and usability," *Int. J. Hum.-Comput. Stud.*, vol. 55, no. 4, pp. 533–552, Oct. 2001.

[97] L. Barnett, C. Harvey, and C. Gatzidis, "First time user experiences in mobile games: An evaluation of usability," *Entertainment Comput.*, vol. 27, pp. 82–88, Aug. 2018.

[98] N. Bessghaier and M. Souii, "Towards usability evaluation of hybrid mobile user interfaces," in *Proc. IEEE/ACS 14th Int. Conf. Comput. Syst. Appl. (AICCSA)*, Oct. 2017, pp. 895–900.

[99] R. Duran-Saez, X. Ferre, H. Zhu, and Q. Liu, "Task analysis-based user event logging for mobile applications," in *Proc. IEEE 25th Int. Requirements Eng. Conf. Workshops (REW)*, Sep. 2017, pp. 152–155.

[100] S. Caro-Alvaro, E. Garcia-Lopez, A. Garcia-Cabot, L. de-Marcos, and J.-J. Martinez-Herraiz, "Identifying usability issues in instant messaging apps on iOS and Android platforms," *Mobile Inf. Syst.*, vol. 2018, pp. 1–19, Oct. 2018.

[101] J. Espinoza, P. Loarte, C. Espinoza, F. Paz, and J. Arenas, "A new software development model: Innovation through mobile application with UCD," in *Proc. Int. Conf. Design, Exper., Usability*. Cham, Switzerland: Springer, 2018, pp. 673–692.

[102] F. Z. Ghazizadeh and S. Vafadar, "A quantitative evaluation of usability in mobile applications: An empirical study," in *Proc. Int. Symp. Comput. Sci. Softw. Eng. Conf. (CSSE)*, Oct. 2017, pp. 1–6.

[103] Z. Huang and Z. Tian, "Analysis and design for mobile applications: A user experience approach," in *Proc. Int. Conf. Design, User Exper., Usability*. Cham, Switzerland: Springer, 2018, pp. 91–100.

[104] A. Hussain, E. O. Mkpojiogu, S. Mortada, and W. S. Yue, "Mobile experience evaluation of an e-reader app," *J. Telecommun., Electron. Comput. Eng. (JTEC)*, vol. 10, nos. 1–10, pp. 11–15, 2018.

[105] J. Krzewińska, A. Indyka-Piasecka, M. Kopel, E. Kukla, Z. Telec, and B. Trawiński, "Usability testing of a responsive Web system for a school for disabled children," in *Proc. Asian Conf. Intell. Inf. Database Syst.* Cham, Switzerland: Springer, 2018, pp. 705–716.

[106] K. M. Malan, J. H. P. Eloff, and J. A. De Bruin, "Semi-automated usability analysis through eye tracking," *South Afr. Comput. J.*, vol. 30, no. 1, pp. 66–84, 2018.

[107] S. S. Nathan, A. Hussain, and N. L. Hashim, "Usability evaluation of deaf mobile application interface: A systematic review," *J. Eng. Appl. Sci.*, vol. 13, no. 2, pp. 291–297, 2018.

[108] R. Pakomera, D. van Greunen, and A. Veldsman, "Usability and usefulness in designing patient-centric systems in afro-centric setting," in *Proc. 6th Int. Conf. Rel., Infocom Technol. Optim. (Trends Future Directions) (ICRITO)*, Sep. 2017, pp. 576–579.

[109] S. Shafiq and T. A. Khan, "Role & value of usability in educational learning via game based apps," *Int. J. Sci. Technol. Res.*, vol. 7, pp. 70–77, Nov. 2018.

[110] V. N. Adama, I. S. Shehu, S. A. Adepoju, and R. G. Jimoh, "Towards designing mobile banking user interfaces for novice users," in *Proc. Int. Conf. Design, User Exper., Usability*. Cham, Switzerland: Springer, 2017, pp. 181–197.

[111] M. Aguilar and C. Zapata, "Integrating UCD and an agile methodology in the development of a mobile catalog of plants," in *Advances in Ergonomics Modeling, Usability & Special Populations*. Cham, Switzerland: Springer, 2017, pp. 75–87.

[112] S. Caro-Alvaro, E. Garcia-Lopez, A. Garcia-Cabot, L. de-Marcos, and J.-M. Gutierrez-Martinez, "A systematic evaluation of mobile applications for instant messaging on iOS devices," *Mobile Inf. Syst.*, vol. 2017, pp. 1–17, Oct. 2017.

[113] Z. M. C. Cruz, J. J. R. Alpay, J. D. D. Depeno, M. J. C. Altabirano, and R. Bringula, "Usability of 'Fatchum': A mobile application recipe recommender system," in *Proc. 6th Annu. Conf. Res. Inf. Technol.* New York, NY, USA: ACM, 2017, pp. 11–16.

[114] S. Ibtasam, H. Mehmood, L. Razaq, J. Webster, S. Yu, and R. Anderson, "An exploration of smartphone based mobile money applications in pakistan," in *Proc. 9th Int. Conf. Inf. Commun. Technol. Develop. (ICTD)*. New York, NY, USA: ACM, 2017, p. 1.

[115] S. Prezenski, D. Bruechner, and N. Russwinkel, "Predictive cognitive modelling of applications," in *Proc. 12th Int. Joint Conf. Comput. Vis., Imag. Comput. Graph. Theory Appl.*, 2017, pp. 165–171.

[116] Y. Wang and R. Brennan, "Building high usability consumer-oriented linked data mobile apps," in *Proc. Irish Conf. Artif. Intell. Cogn. Sci. (AICS)*, Dublin, Republic of Ireland, Dec. 2017, pp. 1–12.

[117] Q. Ye, S. A. Boren, U. Khan, and M. S. Kim, "Evaluation of functionality and usability on diabetes mobile applications: A systematic literature review," in *Proc. Int. Conf. Digit. Hum. Modeling Appl. Health, Saf., Ergonom. Risk Manage.* Cham, Switzerland: Springer, 2017, pp. 108–116.

[118] V. V. S. M. Chintapalli, W. Tao, Z. Meng, K. Zhang, J. Kong, and Y. Ge, "A comparative study of spreadsheet applications on mobile devices," *Mobile Inf. Syst.*, vol. 2016, pp. 1–10, Feb. 2016.

[119] H. Hoehle, R. Aljafari, and V. Venkatesh, "Leveraging Microsoft's mobile usability guidelines: Conceptualizing and developing scales for mobile application usability," *Int. J. Hum.-Comput. Stud.*, vol. 89, pp. 35–53, May 2016.

[120] R. Liu, J. Cao, K. Zhang, W. Gao, J. Liang, and L. Yang, "When privacy meets usability: Unobtrusive privacy permission recommendation system for mobile apps based on crowdsourcing," *IEEE Trans. Services Comput.*, vol. 11, no. 5, pp. 864–878, Sep. /Oct. 2018.

[121] I. Nascimento, W. Silva, B. Gadelha, and T. Conte, "Userbility: A technique for the evaluation of user experience and usability on mobile applications," in *Proc. Int. Conf. Hum.-Comput. Interact.* Cham, Switzerland: Springer, 2016, pp. 372–383.

[122] P. Paškevicius and R. Damaševicius, "Design of usable interface for a mobile e-commerce system," in *Proc. Int. Conf. Young Researchers Inform., Math. Eng.*, 2016, pp. 31–35.

[123] M. Tenemaza, E. Mena-Maldonado, F. Ledesma, M. Llumiquinga, and S. Luján-Mora, "Augmented reality in limited environment, analysis of usability," in *Proc. 9th Int. Conf. ICT, Soc. Hum. Beings (ICT)*, 2016, pp. 277–279.

[124] A. S. A. Al-Aidaroos and A. Abdul Mutalib, "Design of the usability measurement tool for multimodal mobile applications," *J. Teknologi*, vol. 77, no. 29, pp. 41–47, 2015.

[125] M. Al-khomsan, A. Al-Arjan, A. Al-Amro, and K. Al-Nafjan, "Usability evaluation of Twitter on Android platform for elderly arab users using morae evaluation tool," in *Proc. 10th Int. Conf. Internet Technol. Secured Trans. (ICITST)*, Dec. 2015, pp. 446–447.

[126] H. M. Az-Zahra, A. Pinandito, and H. Tolle, "Usability evaluation of mobile application in culinary recommendation system," in *Proc. IEEE Asia–Pacific Conf. Wireless Mobile (APWiMob)*, Aug. 2015, pp. 89–94.

[127] Z. Elkheir and A. Abdul Mutalib, "Mobile learning applications designing concepts and challenges: Survey," *Res. J. Appl. Sci., Eng. Technol.*, vol. 10, no. 4, pp. 438–442, 2015.

[128] H. Hoehle and V. Venkatesh, "Mobile application usability: Conceptualization and instrument development," *MIS Quart.*, vol. 39, no. 2, pp. 435–472, Feb. 2015.

[129] A. Hussain, H. I. Abubakar, and N. B. Hashim, "Evaluating mobile banking application: Usability dimensions and measurements," in *Proc. 6th Int. Conf. Inf. Technol. Multimedia*, Nov. 2014, pp. 136–140.

[130] M. S. Kim, A. A. Ben Ramadan, M. A. Clarke, M. K. Markey, K. J. Lage, M. R. Aro, K. L. Ingalls, and V. Sindhwani, "Usability of mobile applications supporting training in diagnostic decision-making by radiologists," in *Int. Conf. Digit. Hum. Modeling Appl. Health, Saf., Ergonom. Risk Manage.* Cham, Switzerland: Springer, 2015, pp. 448–454.

[131] P. Kortum and M. Sorber, "Measuring the usability of mobile applications for phones and tablets," *Int. J. Hum.-Comput. Interact.*, vol. 31, no. 8, pp. 518–529, Aug. 2015.

[132] S. J. Levulis and D. J. Harris, "Are all tests equal? A comparison of emulator and device testing for mobile usability evaluation," in *Proc. Hum. Factors Ergonom. Soc. Annu. Meeting*, vol. 59, no. 1. Los Angeles, CA, USA: Sage, 2015, pp. 976–980.

[133] M. Plachkinova, G. Faddoul, and S. Chatterjee, "Designing a mobile application for complementary and alternative medicine: A usability approach," in *Proc. Int. Conf. Hum.-Comput. Interact.* Cham, Switzerland: Springer, 2015, pp. 345–349.

[134] F. Botella, J. P. Moreno, and A. Peñalver, "How efficient can be a user with a tablet versus a smartphone?" in *Proc. 15th Int. Conf. Hum. Comput. Interact.* New York, NY, USA: ACM, 2014, p. 64.

[135] W. Y. Chua, M. P.-H. Wan, K. Chang, and W. Yi, "Improving mobile applications usage experience of novice users through user-acclimatized interaction: A case study," in *Proc. Americas Conf. Inf. Syst.*, 2014, pp. 1–10.

[136] A. Hussain, N. L. Hashim, and N. Nordin, "MGQM: Evaluation metric for mobile and human interaction," in *Proc. Int. Conf. Hum.-Comput. Interact.* Cham, Switzerland: Springer, 2014, pp. 42–47.

[137] K. Lapin, "Deriving usability goals for mobile applications," in *Proc. Mulitmedia, Interact., Design Innov. Int. Conf. Multimedia, Interact., Design Innov. (MIDI)*. New York, NY, USA: ACM, 2014, pp. 1–6.

[138] J.-M. López-Gil, M. Urretavizcaya, B. Losada, and I. Fernández-Castro, "Integrating field studies in agile development to evaluate usability on context dependant mobile applications," in *Proc. 15th Int. Conf. Hum. Comput. Interact.* New York, NY, USA: ACM, 2014, pp. 24:1–24:8.

[139] A. S. Tsiaousis and G. M. Giaglis, "Mobile websites: Usability evaluation and design," *Int. J. Mobile Commun.*, vol. 12, no. 1, pp. 29–55, 2014.
[140] L. Wein, "Visual recognition in museum guide apps: Do visitors want it?" in *Proc. 32nd Annu. ACM Conf. Hum. Factors Comput. Syst. (CHI)*. New York, NY, USA: ACM, 2014, pp. 635–638.
[141] B. Aryana and T. Clemmensen, "Mobile usability: Experiences from Iran and Turkey," *Int. J. Hum.-Comput. Interact.*, vol. 29, no. 4, pp. 220–242, Mar. 2013.
[142] M. N. Islam, "A systematic literature review of semiotics perception in user interfaces," *J. Syst. Inf. Technol.*, vol. 15, no. 1, pp. 45–77, 2013.
[143] P. Chynał, J. M. Szymański, and J. Sobecki, "Using eyetracking in a mobile applications usability testing," in *Proc. Asian Conf. Intell. Inf. Database Syst.* Berlin, Germany: Springer, 2012, pp. 178–186.
[144] D. Flood, R. Harrison, C. Iacob, and D. Duce, "Evaluating mobile applications: A spreadsheet case study," *Int. J. Mobile Hum. Comput. Interact.*, vol. 4, no. 4, pp. 37–65, Oct. 2012.
[145] F. Nayebi, J.-M. Desharnais, and A. Abran, "The state of the art of mobile application usability evaluation," in *Proc. 25th IEEE Can. Conf. Electr. Comput. Eng. (CCECE)*, Apr. 2012, pp. 1–4.
[146] M. Nourbakhsh, R. Mohamad Zin, J. Irizarry, S. Zolfagharian, and M. Gheisari, "Mobile application prototype for on-site information management in construction industry," *Eng., Construct. Architectural Manage.*, vol. 19, no. 5, pp. 474–494, Aug. 2012.
[147] R. Hegarty and J. Wusteman, "Evaluating EBSCOhost mobile," *Library Hi Tech*, vol. 29, no. 2, pp. 320–333, 2011.
[148] B. Biel, T. Grill, and V. Gruhn, "Exploring the benefits of the combination of a software architecture analysis and a usability evaluation of a mobile application," *J. Syst. Softw.*, vol. 83, no. 11, pp. 2031–2044, Nov. 2010.
[149] K. Chen, S. Wang, and Y. Pei, "The adoption and use of mobile commerce: An information processing perspective," in *Proc. Int. Conf. Manage. Service Sci.*, Sep. 2009, pp. 1–4.
[150] B. Falchuk, "Visual and interaction design themes in mobile healthcare," in *Proc. 6th Annu. Int. Conf. Mobile Ubiquitous Syst., Comput., Netw. Services*, Jul. 2009, pp. 1–10.
[151] D. Zhang, B. Adipat, and Y. Mowafi, "User-centered context-aware mobile applications—The next generation of personal mobile computing," *Commun. Assoc. Inf. Syst.*, vol. 24, no. 1, p. 3, 2009.
[152] K. Liukkunen, M. Eteläperä, M. Oivo, J.-P. Soininen, and M. Pellikka, "Virtual prototypes in developing mobile software applications and devices," in *Proc. Int. Conf. Product Focused Softw. Process Improvement*. Berlin, Germany: Springer, 2008, pp. 174–188.
[153] K. H. Moe, B. Dwolatzky, and R. V. Olst, "Designing a usable mobile application for field data collection," in *Proc. IEEE Africon. 7th Africon Conf. Afr.*, vol. 2, Sep. 2004, pp. 1187–1192.
[154] C. D. de Oliveira, R. P. de Mattos Fortes, and E. F. Barbosa, "An analysis of crossword learning: A mobile application for the elderly," in *Proc. Int. Conf. Hum. Aspects IT Aged Population*. Cham, Switzerland: Springer, 2018, pp. 501–515.
[155] O. Korhan and M. Ersoy, "Usability and functionality factors of the social network site application users from the perspective of uses and gratification theory," *Qual. Quantity*, vol. 50, no. 4, pp. 1799–1816, Jul. 2016.
[156] W.-M. Han, C.-H. Hsu, and C.-Y. Yeh, "Using dematel to analyze the quality characteristics of mobile applications," in *Proc. Int. Conf. Future Inf. Eng. Manuf. Sci. (FIEMS)*, vol. 2. Boca Raton, FL, USA: CRC Press, 2015, p. 131.
[157] S. Geiger-Prat, B. Marin, S. Espana, and G. Giachetti, "A GUI modeling language for mobile applications," in *Proc. IEEE 9th Int. Conf. Res. Challenges Inf. Sci. (RCIS)*, May 2015, pp. 76–87.
[158] R. Capilla, L. Carvajal, and H. Lin, "Addressing usability requirements in mobile software development," in *Relating System Quality and Software Architecture*. Amsterdam, The Netherlands: Elsevier, 2014, pp. 303–324.
[159] S. P. Osorio, L. M. Aristizabal, and C. A. Zuluaga, "Development of a command interface based on handheld devices for remotely operated vehicles," in *Proc. IEEE Colombian Conf. Robot. Autom. (CCRA)*, Sep. 2016, pp. 1–5.
[160] H. I. Abubakar, N. L. Hashim, and A. Hussain, "Usability evaluation model for mobile banking applications interface: Model evaluation process using experts' panel," *J. Telecommun., Electron. Comput. Eng.*, vol. 8, no. 10, pp. 53–57, 2016.
[161] R. Inostroza, C. Rusu, S. Roncagliolo, V. Rusu, and C. A. Collazos, "Developing SMASH: A set of smartphone's usability heuristics," *Comput. Standards Interfaces*, vol. 43, pp. 40–52, Jan. 2016.
[162] M. A. Kabir, O. A. M. Salem, and M. U. Rehman, "Discovering knowledge from mobile application users for usability improvement: A fuzzy association rule mining approach," in *Proc. 8th IEEE Int. Conf. Softw. Eng. Service Sci. (ICSESS)*, Nov. 2017, pp. 126–129.
[163] A. Kulpy and G. Bekaroo, "Fruitify: Nutritionally augmenting fruits through markerless-based augmented reality," in *Proc. IEEE 4th Int. Conf. Soft Comput. Mach. Intell. (ISCMI)*, Nov. 2017, pp. 149–153.
[164] Y. Man, M. Lützhöft, N. A. Costa, M. Lundh, and S. N. MacKinnon, "Gaps between users and designers: A usability study about a tablet-based application used on ship bridges," in *Proc. Int. Conf. Appl. Hum. Factors Ergonom.* Cham, Switzerland: Springer, 2017, pp. 213–224.
[165] A. Widyanti and S. A. Q. Ainizzamani, "Usability evaluation of online transportation' user interface," in *Proc. Int. Conf. Inf. Technol. Syst. Innov. (ICITSI)*, Oct. 2017, pp. 82–86.
[166] A. Hussain, H. A. Razak, and E. O. Mkpojiogu, "The perceived usability of automated testing tools for mobile applications," *J. Eng., Sci. Technol.*, vol. 12, no. 4, pp. 89–97, 2017.
[167] W. Roh and S.-W. Lee, "An ontological approach to predict trade-offs between security and usability for mobile application requirements engineering," in *Proc. IEEE 25th Int. Requirements Eng. Conf. Workshops (REW)*, Sep. 2017, pp. 69–75.
[168] A. Saleh, R. Ismail, and N. Fabil, "Evaluating usability for mobile application: A MAUEM approach," in *Proc. Int. Conf. Softw. e-Bus. (ICSEB)*. New York, NY, USA: ACM, 2017, pp. 71–77.
[169] Z. Zali, "An initial theoretical usability evaluation model for assessing defence mobile e-based application system," in *Proc. Int. Conf. Inf. Commun. Technol. (ICICTM)*, May 2016, pp. 198–202.
[170] I. Costa, W. Silva, A. Lopes, L. Rivero, B. Gadelha, E. Oliveira, and T. Conte, "An empirical study to evaluate the feasibility of a UX and usability inspection technique for mobile applications," in *Proc. 28th Int. Conf. Softw. Eng. Knowl. Eng.*, Jul. 2016, pp. 1–5.
[171] B. A. Kumar and P. Mohite, "Usability of mobile learning applications: A systematic literature review," *J. Comput. Edu.*, vol. 5, no. 1, pp. 1–17, Mar. 2018.
[172] D. F. O. de Paula, B. H. X. M. Menezes, and C. C. Araújo, "Building a quality mobile application: A user-centered study focusing on design thinking, user experience and usability," in *Proc. Int. Conf. Design, User Exper., Usability*. Cham, Switzerland: Springer, 2014, pp. 313–322.
[173] A. Hussain, N. L. Hashim, N. Nordin, and H. M. Tahir, "A metric-based evaluation model for applications on mobile phone," *J. Inf. Commun. Technol.*, vol. 12, pp. 55–71, Dec. 2013.
[174] J. Eustáquio Rangel de Queiroz and D. de Sousa Ferreira, "A multidimensional approach for the evaluation of mobile application user interfaces," in *Proc. Int. Conf. Hum.-Comput. Interact.* Berlin, Germany: Springer, 2009, pp. 242–251.
[175] M. Kenteris, D. Gavalas, and D. Economou, "An innovative mobile electronic tourist guide application," *Pers. Ubiquitous Comput.*, vol. 13, no. 2, pp. 103–118, Feb. 2009.
[176] A. Hussain and E. Ferneley, "Usability metric for mobile application: A goal question metric (GQM) approach," in *Proc. 10th Int. Conf. Inf. Integr. Web-Based Appl. Services (iiWAS)*. New York, NY, USA: ACM, 2008, pp. 567–570.
[177] J. Häkkilä, P. Korpipää, S. Ronkainen, and U. Tuomela, "Interaction and end-user programming with a context-aware mobile application," in *Proc. IFIP Conf. Hum.-Comput. Interact.* Berlin, Germany: Springer, 2005, pp. 927–937.
[178] C. Ryan and A. Gonsalves, "The effect of context and application type on mobile usability: An empirical study," in *Proc. 28th Australas. Conf. Comput. Sci.*, vol. 38. Darlinghurst, NSW, Australia: Australian Computer Society, 2005, pp. 115–124.
[179] P. Jarupunphol, W. Buathong, T. Chansaeng, and N. Laosen, "A descriptive design for a smart kitchen management application (SKM)," in *Proc. Int. Conf. Inf. Comput. Technol. (ICICT)*, Mar. 2018, pp. 61–66.
[180] K. McDonald, W. Abell, C. Smith, and S. Gibbs, "Lessons learned from evaluating a mobile app out in the field," in *Proc. 4th Int. Conf. User Sci. Eng. (i-USEr)*, Aug. 2016, pp. 5–10.
[181] E. O. Mkpojiogu, N. L. Hashim, and R. Adamu, "Observed demographic differentials in user perceived satisfaction on the usability of mobile banking applications," in *Proc. 8th Knowl. Manage. Int. Conf. (KMICe)*, Chiang Mai, Thailand, 2016, pp. 29–30.
[182] P. Pocatilu, I. Ivan, A. Zamfiroiu, and C. Boja, "An aggregate indicator for mobile application quality assessment," *TURKISH J. Electr. Eng. Comput. Sci.*, vol. 23, pp. 956–973, 2015.

[183] M. Fetaji, Z. Dika, and B. Fetaji, ''Usability testing and evaluation of a mobile software solution: A case study,'' in *Proc. ITI-30th Int. Conf. Inf. Technol. Interfaces*, Jun. 2008, pp. 501–506.

[184] N. K. Chuan, A. Sivaji, F. A. Loo, W. F. W. Ahmad, and S. S. Nathan, ''Evaluating 'gesture interaction' requirements of mobile applications for deaf users: Discovering the needs of the hearing-impiared in using touch-screen gestures,'' in *Proc. IEEE Conf. Open Syst. (ICOS)*, Nov. 2017, pp. 90–95.

[185] S. A. A. Shukri, H. Arshad, and R. Z. Abidin, ''Mobile augmented reality system design guidelines based on tourist's emotional state,'' *J. Telecommun., Electron. Comput. Eng.*, vol. 9, nos. 2–12, pp. 75–79, 2017.

[186] T. Aljaber and N. Gordon, ''A guidance and evaluation approach for mHealth education applications,'' in *Proc. Int. Conf. Learn. Collaboration Technol.* Universiti Teknikal Malaysia Melaka: Springer, 2017, pp. 330–340.

[187] M. A. Redondo, A. I. Molina, and C. X. Navarro, ''Extending CIAM methodology to support mobile application design and evaluation: A case study in m-learning,'' in *Proc. Int. Conf. Cooperat. Design, Visualizat. Eng.* Cham, Switzerland: Springer, 2015, pp. 11–18.

[188] S. Valdivia, R. Blanco, A. Uribe-Quevedo, L. Penuela, D. Rojas, and B. Kapralos, ''Development and evaluation of two posture-tracking user interfaces for occupational health care,'' *Adv. Mech. Eng.*, vol. 10, no. 6, Jun. 2018, Art. no. 168781401876948.

[189] A. Hussain and R. Fitria, ''Mobile flight and hotel booking application: A heuristic and UX test,'' *J. Telecommun., Electron. Comput. Eng.*, vol. 10, nos. 1–11, pp. 93–101, 2018.

PAWEŁ WEICHBROTH received the M.A. degree in statistics from the University of Gdańsk, Poland, in 2003, and the Ph.D. degree in artificial intelligence from the Katowice University of Economics, Poland, in 2014.

He is currently an Assistant Professor with the Department of Software Engineering, Gdańsk University of Technology. In this regard, he has authored more than 40 research articles as journal articles, conference papers, and book chapters. His main research areas concern software quality, machine learning, and knowledge management. He has been a member of the Scientific Community of Business Informatics, a member of several international conference program committees, and currently acting as a Reviewer in journals with an impact factor. Moreover, for more than 20 years, he has worked as a Business Consultant and an IT Lecturer. Since 2018, he has been acting as an Expert of the Ministry of Digital Affairs in a project for the development of public digital services.