Guerilla testing of mobile medical applications in clinical settings

Experiences from field testing in four Ugandan health institutions

Lina Linnea Aspen
Department of Product Design
Norwegian University of Science and Technology

ABSTRACT

Human factors considerations and usability evaluations are included in international guidelines and regulations for development of healthcare technology to maximize the likelihood that new medical devices will be safe and effective for the intended users, uses and use environments. The established usability testing methods are however resource intensive and intimidate developers who are already obliged to undergo substantial clinical testing. The contextual complexity of clinical settings is difficult to simulate realistically in a lab, while demands for objectiveness, for instance, makes field testing a demanding activity. Guerilla testing is a cheap and manageable alternative method for usability testing in the field. It is a suitable method for testing in clinical settings because of its non-invasive and flexible setup, while simultaneously enabling for testing in several situs. Risks are related to participant recruitment and unpredictability of the use situation. This paper will discuss guerilla testing as a method to test mobile medical apps by reviewing literature and reflecting upon the author's experiences from using the method to test a mobile medical application in Ugandan health institutions.

KEYWORDS: Guerilla testing, mobile medical application

1. INTRODUCTION

With usability engineering and human factor principles entering the regulations and guidelines [1,2,3,4] for developing medical informatics, there is a need for reevaluating the conventional methods of usability testing and explore which techniques are more practically manageable now that smartphone technologies open up for smaller less, established companies to enter the market. Researchers have long elaborated the importance of field testing [5 pp. 110], while in practice it can be so demanding that some argue the added value is not worth the efforts [6, 7]. The conventional understanding of usability testing attach importance to obtaining objective results to undergo the evaluation and consequently lab testing has gained momentum. The bottleneck of successful lab tests is the construction of realistic scenarios, which is especially challenging for clinical settings as they are characterized by unpredictability and complex social structures [7].

The main reasons companies do not conduct usability testing are due to the intimidating costs and complexity related to established techniques [8]. To counteract this there has evolved alternative discount approaches [8, 9] to lower the threshold of conducting usability testing. One low cost method for field testing is guerilla testing which emphasizes the value of qualitative

testing with the intended users, instead of striving for statistical results [10].

This paper will present how guerilla testing can be suitable for testing of mobile medical tools in clinical settings, and discuss the related risks and benefits. This will be seen in conjunction with existing literature on both field and lab testing, and medical tools for professional use. The article also contains reflections on personal experiences from a case using the method to test a diagnostic smartphone application in Ugandan health institutions.

1.1 Usability is context dependent

The International Organization for Standardization defines usability as "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" [1]. This definition lets usability be measured by three targets – effectiveness, efficiency and satisfaction, in relation to three aspects of the use situation – the users, their goals and the context of use [11]. This states how the usability of a product is dependent on how well it fits, amongst other, the use situation and hence is a context dependent factor [11].

As field testing is demanding to conduct, especially if comparable results are required, alternative methods have evolved. One accepted alternative is to set up a simulation in a lab. This way the test can be performed in controlled environments facilitated for collecting comparable data. By presenting scenarios and recreating environments within the lab, this approach to usability testing has acquired acceptance by professionals and research shows it is a good substitute for naturalistic studies [12]. Some even argue more usability problems are detected by lab testing compared to in-situ; because of the unpredictable aspect of field tests which makes it demanding when striving for comparable results [6].

With the emergence of mobile technology, lab testing lost some of its dominant because it fails to support mobile user scenarios. Mobility is not well supported within a lab and researches again raised the voice for field testing [13, 14 pp. 49]. In addition, a realistic recreation of a use context within a lab is challenging to achieve. In relation to clinical settings it has proven to be difficult to stage a realistic environment due the complexity of this use situation [11]. Clinical settings are diverse and often include a social aspect involving several users, as well as an unpredictable environment which is hard to replicate in a lab [11].

1.2 Why field testing is demanding

Field testing of clinical products is in nature challenging to conduct due to safety and privacy restrictions to preserve safe operation and patient interaction. Obtaining access to health facilities is in itself a barrier to perform field testing. Due to the variations in national and institutional regulations, this topic is not discussed in this paper.

Another demanding factor of field testing that is reinforced when addressing clinical settings, is the requirements of objective measurements for the usability evaluation [15]. Attaining this is hard when the use context in nature is unpredictable and dynamic [15]. This uncontrolled environment makes it difficult to carry out identical tests that can provide comparable results for an objective evaluation. The requirement of objective results has been long debated and Jakob Nielsen has expressed concerns that it overrules the actual aim of usability testing – gathering insights and understanding user experiences [8].

Objectiveness also requires careful data collection for post analysis. New methods and equipment to record field tests are being explored, but it is challenging to capture contextual issues in an uncontrolled environment without invading the situation and affect the results [15]. Observers, record equipment and test moderators are all intruders of the use

context and should be minimized to avoid implications of the Hawthorne effect.

1.2 The Hawthorne Effect

The Hawthorne effect is generally defined as when participants in field experiments are aware of their participation and hence modifies their behavior [16]. It implies that subjects don't act naturally when they are under observation [14 pp. 29]. The awareness of being part of a research experiment can result in the participants recognizing the motive of the experiment and, consciously or unconsciously, act differently than they would under normal conditions. In regard to usability testing this can for example play out with the participant recognizing the investments in the project and trying to please the test moderator by downplaying issues that arise. It can also occur if the participant receive an allowance for their participation, and therefore feel their contribution should correspond to this and hence adjust their behavior to provide useful insights to make the test moderators perceive the sessions as worth the costs. This paper will argue that guerilla testing to a lesser extent causes Hawthorne effects since the method is less obtrusive, the participants less prepared, and the relationship between the moderator and the participant is less established.

1.3 Mobile medical apps

The U.S. Food and Drug Administration (FDA) offer the following definition on mobile medical apps: "Mobile apps are software programs that run on smartphones and other mobile communication devices. They can also be accessories that attach to a smartphone or other mobile communication devices, or a combination of accessories and software. Mobile medical apps are medical devices that are mobile apps, meet the definition of a medical device and are an accessory to a regulated medical device or transform a mobile platform into a regulated medical device." [17].

This paper addresses this specific product group as it differs from other medical tools with its properties of mobility and as part of a multi-functional device. In addition this article will limit its scope to only address medical apps for professional use by healthcare personnel.

The case study used in this paper is in line with FDA's definition. The product in the case consists of a software application running on a smartphone with an accessory in form of an accompanying color calibration card. It is targeting healthcare workers in health institutions or in the field as a tool to diagnose jaundice in newborns.

2. GUERILLA TESTING

The barrier to undergo usability testing in the field is apparent when developing mobile medical applications, but definitely desired so as to minimize risks and ensure safe use. As described earlier the clinical setting is hard to simulate in the lab due to its complex and unpredictable nature. The following chapter will reflect upon the benefits and risks related to guerilla testing as a manageable method to conduct field testing of mobile medical applications.

2.1 The Method

Guerilla testing is a low cost method to quickly test prototypes with end users in the intended context of use [10]. The method can be characterized as a quick and low-cost usability method manageable for non-professionals [10]. The sessions are short and often informal in nature, and well suited for iterative processes by enabling for quick execution and analysis to feed into the next cycle of development [10]. It is a qualitative method giving rich insights instead of statistical results, with the goal of improving the design – not proving it [18 pp. 12]. By not aiming for objective results, the setup gains an aspect of flexibility which makes it suitable the uncontrollability related to field tests. Guerilla testing also differs from other methods by involving participants to test ad-hoc. This means

approaching potential users and ask if they are willing to take part in a test there and then. This approach puts limitations on the test setup, but also offers opportunities that other methods don't. Without previously agreeing on participation, the test setup must be of a nature that attracts people to be willing to take part. This also puts constraints on the length and content of the test to perceive people to attend. The setup leaves the participant with little time to get familiar with the topic at hand, nor conduct extensive scenarios. Topics to explore during the test should therefore not go too deep or be too complex as that would require more time to get acquainted with the concept being tested.

The guerilla method is more manageable in terms of both time and money spent organizing and executing it than conventional field testing or lab testing. In addition, expertise needed to conduct guerilla testing is lower than for testing striving for objective results. The UK Government Service Design Manual says that "Anyone on the service team can conduct 'guerrilla testing' on their site or service" [10]. This is one of the main advantages of the method and shows how valuable it can be for developing teams by contributing to overcome the issue that Jakob Nielsen describes as: "the perception that anybody touching usability will come down with a bad case of budget overruns" [8].

The affordability and feasibility of guerilla testing have to a large degree been its promoted attributes, as a way to lower the threshold of conducting usability tests. This is also the main reason for why the author made use of the method. Still, through the experiences made from guerilla testing of a mobile medical application in Uganda, this paper wants to advocate for a larger acceptance of guerilla testing as a method which also has disciplinary benefits and especially is well-suited for testing mobile medical apps in clinical settings.

3. EXPERIENCES FROM A CASE

The author conducted field testing of a mobile medical tool in four different health institutions in Uganda using the guerilla testing method. Underneath follows experiences from using the method, and reflection on its pros and cons regarding testing of mobile medical applications for clinical use.

3.1 The case

The product being tested was a smartphone application for diagnosing jaundice in newborns, and the focus was to gather insights and feedback on how this tool could be made relevant for healthcare workers in low resource settings by, amongst other, evaluating the current design. The research in Uganda consisted of expert interviews with scholars in newborn care and rural healthcare, as well as interviews and testing of the prototype with nurses, doctors and midwives in four different health institutions. An iOS prototype initiated the usability testing, while also functioned as a way to materialize the concept and demonstrate the technology when talking with experts and health workers. The field trip was conducted by the author alone, due to financial limitations, who adopted the role of test moderator and interviewer.

The vision initiating the product was to bring diagnostic equipment to areas in the world with little access to the existing tools of today due to lack of infrastructure or financial limitations. As the health institutions in the company's home state Norway differed too much with the intended context of use, it was necessary to conduct the testing in a more relevant country to make sure the product would meet the needs of healthcare workers in low resource settings. Examples of aspects of the use scenario that characterizes many low resource health systems are shortage in staff and expertise, shortage in medical equipment, shortage in power supply, lack of infrastructure and scarce rollout of health services to rural areas [25]. These are all aspects that define limitations and opportunities that the design of the diagnostic tool must meet to ensure a good user experience.

3.2 How the testing was conducted

The institutions visited were selected based on the wish to address the whole specter of healthcare services offered in low income countries. There is a large variation in healthcare quality, equipment availability and staff expertise within the Ugandan health system [19], and we wanted to explore if the tool could fit the different levels or if there was a need for segmentation.

The sites visited ranged from national hospitals to rural clinics. The author would meet with a contact person who served as a guide to the premises. On each site the author would visit the maternity ward, observe and discuss the current work conditions, interact with the staff, and inquire doctors, nurses and midwifes to participate in a test of the prototype. Those who had the opportunity to participate were given a brief introduction to the project and how the test would play out. They were informed on test procedure principles, including the common phrase – "it is not you we are testing it is the prototype", and encouraged to talk out loud about how they were experiencing the tool while using it. It was emphasized how the prototype was not working properly, so they should not rely on the diagnose it provided. Next they got some information on how the tool worked – that you had to put the calibration card on the baby's chest and use the app to take pictures of it. Finally they were handed the calibration card and the prototype phone and asked to test a newborn for jaundice.

Figure 1: The setup: A doctor testing the diagnosing app on a newborn by photographing the color calibration card on its chest.

During the test the test moderator would observe, take pictures and notes. If the participant inquired assistance, the moderator would kindly refrain and urge them to try to figure it out by themselves. When the task was completed and the participant had received a test result; the moderator would ask questions on how they understood the test result and follow up on issues observed during the testing. The were variations on how much time the participants had at hand, but the test and the prototype often served as a base for further exchange of experiences from their work regarding neonatal jaundice and newborn health. When participants had additional time to spare, they were asked to conduct the same task again, on another baby; as babies behave differently. During a second, or even third, test, the environment could have changed – revealing new issues regarding use scenarios. These occasions would provide insights on the learning curve of the application.

The number of participants attending test sessions at each institution would vary. There was also a variation in the composition of disciplines, the level of expertise and the conditions at the locations.

The aspects described above show how this case can be categorized as guerilla testing. It is in line with the method in how participants were approached ad hoc, the sessions were short and informal, the focus was on gathering rich insights - not statistical results, and it addressed the end users and the intended use context.

3.1 Why guerilla testing suited this case

The guerilla method was the preferred approach due to several reasons. First of all there was a strong wish to operate in the field so as to gather more insights on the use context, but arranging conventional field tests was too time consuming and expensive. The limited time frame spent in Uganda also pushed in favor of Guerilla testing as it doesn't require preparations on site or from the participants [10]. The concerns when choosing this method was the little control of the execution of the sessions and the insecurity regarding participant recruitment [10]. In the end we managed to engage enough participants and were also satisfied with the composition of disciplines and ages.

The variations within the Ugandan health system called for testing in additional situs to ensure covering the whole specter of potential users and scenarios. Arranging several tests in several locations or constructing lab tests with different setups; require a lot of resources and time, which was unachievable for the start-up organizing the field trip. Guerilla testing, on the other hand, fits well for testing on multiple sites due to its non-invasive and temporary setup [20].

When deciding to go for in-situ testing, the hectic and unpredictable environment of clinical setting calls for a flexible and unobtrusive setup. Guerilla testing is a method that can be conducted on the

terms of the participants and the situation [21], making it suitable for the unpredictable environment of clinical settings. The field testing in Uganda was based on a modular setup that allowed for interruptions and resumption when the situations that required it occurred. Tasks where manageable to get acquainted with and conduct within reasonable time. The test consisted of one main task and related questions, and repetition of this task with variations if participants had more time to spare. This modular setup ensured useful results regardless of how much time the participants had at disposal. The different institutions that were visited in Uganda belonged to different levels of the health system which led to a variation in atmosphere and opportunities for testing. It proved valuable to have a flexible setup that suited the busy surroundings of the maternity ward of large hospitals as well as the slower and modest surroundings of rural clinics. Enabling for testing in several situs proved valuable as many of the usability issues discovered were context related, and hence varied in occurrence from site to site. For example the rural clinics showed a need for a slower step by step process as the staff generally had a lower level of education and less experience with smartphone use. Health workers in large hospitals on the other hand, called for a more streamlined process of using the tool due to the constant time pressure on the ward. Other insights could be more context specific – for example the poor lighting conditions in the rural clinics which challenges the performance of the technology.

4. DISCUSSION

The uncontrolled environment and the loosely structured approach of guerilla testing makes it difficult to achieve statistical and objective results. The method provides qualitative data and can be a gateway to gather context related insights. The lack of objectiveness has been criticized and caused the low recognition of guerilla testing as professional method. A discussion has taken place, with amongst other Jakob Nielsen wishing to shift the focus from

objectiveness and scientific results to user experiences and insights to improve the design [8]. Nielsen describes the problem of how published work of usability research normally describe cases where extra resources were used on deriving results qualifying for publication, while the needs for developing a better design can be met in much simpler ways [8]. He argues that the emphasis on objectivity in usability research is overrated, and that practitioners shouldn't strive for statistical results, but rather results that can assist in making the best design decisions [8]. Regarding testing of mobile medical applications, an additional argument goes that it is more valuable to obtain a variation in users and scenarios to, than achieving comparable results based on a limited set of use settings.

The demand for objectiveness requires comprehensive data collection of the test sessions to undergo post analysis [8]. This creates another a barrier for practitioners as it can be hard to achieve appropriate quality in the field without being intrusive, as well as the difficulties regarding privacy issues at health facilities. All together this represents costs not necessarily resulting in gathering more usability insights [8].

Another challenge derives from the participant recruitment approach. Since guerilla testing requires less from the participants in terms of time investment and the spontaneous aspect of participation, participants can be left indifferent to the completion of the test. This can result in for example hurrying through the setup or investing little thought in their responds [20]. This can also be a result of the little established relationship with the test moderator. To avoid this, preparation and rehearsal by the test moderator is helpful to quickly establish mutual trust and engagement [21].

Guerilla testing in clinical settings must consider the related safety risks to control these. A risk can appear in the prototype being experienced as too realistic, leading to participants taking actions based on this, which is a safety risk if it is not fully developed. In the case described here, this was

amplified by the language barrier and played out by some participants expressing a desire to act on the diagnose provided by the prototype. The prototype used for testing was a high level prototype simulating the process of testing for neonatal jaundice. It did not provide reliable diagnoses, representing a safety risk if participants made decisions based on these. The reason for using a high level prototype was to explore if the health workers could trust a mobile app as a medical tool in areas with little experience in mobile medical apps. This safety risk suggests using a low level prototypes when this is considered to provide sufficient feedback.

In addition there is a risk concerning the uncontrollability of guerilla testing which leaves you uninsured of obtaining insight on desired topics [20]. Doing guerilla testing you have to accept that the execution is done on the terms of the participants [21]. When guerilla testing is done in clinical environments this can imply waiting for the opportunity of spare time to arise. In Uganda some preparations were done on beforehand, but to different degrees in the different institutions. One place the author would meet with workers who were off duty but could still access the premises, which lead to time at disposal and opportunity for testing in situ. Other places this was not possible, and a lot of waiting was involved to wait for some spare time from health workers at duty. This also worked out well, and had a bonus of providing time for observations of the use context.

Being only one person was useful to ensure the unobtrusive quality of guerilla testing, but challenging when gathering data. Solutions to this can be explored in different camera set-ups to enable for easy sharing with the rest of the team when doing post analysis.

5. FINDINGS

5.1 Benefits and risks with guerilla testing of mobile medical apps

From this review and reflection of experiences, guerilla testing proves to have both benefits and risks when applied to testing mobile medical apps in clinical settings. The method suits this domain of products and related contexts because of its non-invasive setup that makes it possible to enter the intended context of use without much disruption and avoid emergence of the Hawthorne effect. Guerilla testing allows for a flexible setup which is valuable when addressing testing in the unpredictable and hectic environment that many clinical operations inhabit. In addition, the setup for guerilla testing enables to easily test in several situs, which is a bonus for products addressing a large specter of health institutions and health workers.

Guerilla testing also opens up for financially and organizational affordable solutions, which is vital for companies skeptical to perform field testing, as well as start-ups with limited funds – a group whom are making an entrance into the field of medical applications these days [22]. In the case described, the guerilla testing method enabled the start-up to incorporate vital field testing in the development process which otherwise would have been omitted.

The outcome of guerilla testing are discovery of usability issues, as well as new insights related to contexts and users [23 pp. 312]. Based on this case there cannot be concluded in whether these usability issues would have been discovered during off-site testing or not. Still there is no question a realistic lab simulation would have been challenging to construct. One consequence is not succeeding in simulating the complexity of the situation, which would lead to the participant to not experience the same cognitive load as if doing the same test in situ, which would result in leaving some usability issues undiscovered. The case described earlier chose guerilla testing to simultaneously gather insights on the area of application, and the method proved suitable for combination with other research methods. For example the test sessions worked well as a conversation starter for more in-depth interviews, and the prototype was helpful as a

way to conceptualize and discuss the project in cases where the language barrier appeared.

On the other hand, if you are striving for statistical result, the flexibility of the setup, the unpredictability of the use setting and recruitment of participants, can be a downside as it challenges the attainment of comparable results as described in the previous chapter. Risks related to guerilla testing are linked to the dependence on spontaneous and willing participants, as well as the unpredictable environment leaving you with little control of the execution [21]. Another risk concerning the participants is that they don't have any incentives to invest thought in their execution of the test and responses to questions. Because they are attending spontaneously and the relationship to the test moderator is little established, the participants might share poorly conceived responses which can result in misleading results [20]. Concerning the non-invasive setup there is a risk of the sessions being too realistic resulting in the participants taking unsafe actions based on the prototype.

5.2 Future perspectives in research and design practice and Relation to usability testing

The market for mobile medical applications [24] is growing and lean startup methodologists are gaining momentum. With new actors and products entering the domain, there has evolved guidelines and regulations to ensure safe use and minimize risks [22]. However, the required activities and authorization processes are demanding and resource intensive, challenging the smaller operators who try enter the market. This calls for a recognition of guerilla testing as a professional method, so as to lower the threshold to undergoing usability evaluations for producers already pressured by extensive regulations. A more accessible and manageable method will enable more testing during the development phase, resulting in lowering costs and being upfront on regulations on usability.

Guidelines on how to plan and conduct guerilla testing in clinical settings, and gain the most effect out of scarce resources, should be developed and shared amongst medical informatics start-ups.

6. CONCLUSIONS

This paper contributes with a new take on the guerilla testing method as a much needed tool for usability testing of mobile medical tools, by discussing its benefits and risks when applied in clinical settings. Benefits are its non-invasive and flexible setup, making it fit for the unpredictable and dynamic environment of clinical settings, as well as enabling for testing in several situs. It has limitations in terms of scope and complexity, and must be adjusted to for the concept at hand in terms of preparations, setup and execution. Risks are related to the uncontrolled environment and participant recruitment process. The method is manageable for non-professionals, but there is a need for further guidelines so as to ensure that practitioners gain the most value out of the tool. To further lower the alienation of guerilla testing as a method to be applied for usability testing of mobile medical apps, more publications on experiences should be made to raise its acknowledgment as a professional usability method.

REFERENCES

[1] International Organization for Standardization. *35.240.80: IT applications in health care technology*. Retrieved 19 April, 2016, from http://www.iso.org/iso/products/standards/catalogue_ics_browse.htm?ICS1=35&ICS2=240&ICS3=80

[2] U.S Food and Drug Administration. *Human Factors and Medical Devices.* Retrieved 19 April, 2016, from http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/HumanFactors/

[3] National Institute of Standards and Technology (NIST). *Health IT Usability*. Retrieved 19 April, 2016, from http://www.nist.gov/healthcare/usability/index.cfm

[4] Agency for Healthcare Research and Quality (AHRQ). *Designing Consumer Health IT: A Guide for Developers and Systems Designers.* Retrieved 19 April, 2016, from https://healthit.ahrq.gov/health-it-tools-and-resources/designing-consumer-health-it-guide-developers-and-systems-designers

[5] Nielsen, J. (1994). *Usability Engineering*. [Online]. Retrieved 19 April, 2016, from http://proquestcombo.safaribooksonline.com/book/web-development/usability/9780125184069/front-cover/cover?uicode=ntnu

[6] Kjeldskov, J. (2004). Is It Worth the Hassle? Exploring the Added Value of Evaluating the Usability of Context-Aware Mobile Systems in the Field. In *Mobile Human-Computer Interaction - Mobile HCI 2004*, 3160, 61-73. Retrieved 3 May, 2016, from http://link.springer.com/chapter/10.1007%2F978-3-540-28637-0_6

[7] Nielsen, C. M., Overgaard, M., Pedersen, M. B., Stage, J., Stenild, S. (2006). It's Worth the Hassle! The Added Value of Evaluating the Usability of Mobile Systems in the Field. In *Proceedings of the 4th Nordic conference on Human-computer interaction: changing roles*, 272-280. Retrieved 19 April, 2016, from http://dx.doi.org/10.1145/1182475.1182504

[8] Nielsen, J. (1994, 1 January). Guerrilla HCI: Using Discount Usability Engineering to Penetrate the Intimidation Barrier. Retrieved 19 April, 2016, from https://www.nngroup.com/articles/guerrilla-hci/

[9] Abras, C., Maloney-Krichmar, D., Preece, J. (2004). User-Centered Design. *Encyclopedia of Human-Computer Interaction.* Thousand Oaks: Sage Publications.

[10] UK Government. Government Service Design Manual. Guerilla testing. Retrieved 19 April, 2016, from https://www.gov.uk/service-manual/user-centred-design/user-research/guerrilla-testing.html

[11] Svanæs, D., Das, A., Alsos, O. A. (2008*).* The contextual nature of usability and its relevance to medical informatics. *Studies in Health Technology and Informatics*, 136, 541-546.

[12] Kushniruk, A. W., Borycki, E. M., Kuwata, S. Kannry, J. (2011*).* Emerging Approaches to Usability Evaluation of Health Information Systems: Towards In-Situ Analysis of Complex Healthcare Systems and Environments. *Studies*

in Health Technology and Informatics, 169, 915 - 919. Retrieved 3 May, 2016, from http://dx.doi.org/10.3233/978-1-60750-806-9-915

[13] Duh, H. B.-L., Tan, G. C. B., Chen, V. H. (2006), 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*, 181-186. Retrieved 19 April, 2016, from http://dx.doi.org/10.1145/1152215.1152254

[14] Hall, E. (2013). *Just Enough Research*. New York: Jeffrey Zeldman

[15] Svanæs, D., Alsos, O. A., Dahl, Y. (2010), Usability testing of mobile ICT for clinical settings: Methodological and practical challenges. *International Journal of Medical Informatics*, 79(4), e24–e34. Retrieved 3 May, 2016, from http://dx.doi.org/10.1016/j.ijmedinf.2008.06.014.

[16] Adair, J.G. (1984). The Hawthorne effect: A reconsideration of the methodological artifact. *Journal of Applied Psychology*, 69(2), 334-345. Retrieved 3 May, 2016, from http://dx.doi.org/10.1037/0021-9010.69.2.334

[17] U.S Food and Drug Administration. *Mobile Medical Application*. Retrieved 19 April, 2016, from http://www.fda.gov/MedicalDevices/DigitalHealth/MobileMedicalApplications/default.htm

[18] Krug, S. (2009) *Rocket Surgery Made Easy: The Do-It-Yourself Guide to Finding and Fixing Usability Problems*. [Online]. Retrieved 19 April, 2016, from http://proquestcombo.safaribooksonline.com/ book/web-development/usability/9780321702821

[19] Ministry of Health, Health Systems 20/20, and Makerere University School of Public Health. (2012). *Uganda Health System Assessment 2011*. [Online]. Retrieved 19 April, 2016, from http://health.go.ug/docs/hsa.pdf

[20] White, E. (2014) *Usability and UX for libraries*. Virginia Commonwealth University, Retrieved 19 April, 2016, from http://scholarscompass.vcu.edu/cgi/viewcontent.cgi?article=1005&context=libraries_present

[21] Kellingley, N. 10 Hints for Carrying Out Better Guerrilla Usability Testing. Retrieved 19 April2016, from https://www.interaction-design.org/literature/article/10-hints-for-carrying-out-better-guerrilla-usability-testing

[22] MIT Technology Review. (2014, 18 November). Mobile Medical Apps: A Market on the Move. Retrieved 19 April, 2016, from https://www.technologyreview.com/s/532661/mobile-medical-apps-a-market-on-the-move/

[23] Rubin, J. Chisnell, D. (2008). *Handbook of Usability Testing*. Indianapolis: Wiley Publishing Inc.

[24] Transparency Market Research. (2016, 16 February). Mobile Medical Apps Market to Grow due to Easy Integration of Mobile Medical Apps in Clinical Workflows. Retrieved 19 April, 2016, from http://www.medgadget.com/2016/02/mobile-medical-apps-market-to-grow-due-to-easy-integration-of-mobile-medical-apps-in-clinical-workflows.html

[25] Health equity: challenges in low income countries. (2009). *African Health Sciences*, *9*(Suppl 2), S49–S51.