This paper was presented at The XXVIII ISPIM Innovation Conference – Composing the Innovation Symphony, Austria, Vienna on 18-21 June 2017. The publication is available to ISPIM members at www.ispim.org.

Assessing User Experience of Context-Aware Interfaces in a Retail Store

Stephanie Van Hove*

imec-MICT-Ghent University, Miriam Makebaplein 1, 9000 Ghent, Belgium.
E-mail: stephanie.vanhove@ugent.be

Aron-Levi Herregodts

imec-MICT-Ghent University, Miriam Makebaplein 1, 9000 Ghent, Belgium.
E-mail: aronlevi.herregodts@ugent.be

Dimitri Schuurman

imec-MICT-Ghent University, Miriam Makebaplein 1, 9000 Ghent, Belgium.
E-mail: dimitri.schuurman@ugent.be

Lieven De Marez

imec-MICT-Ghent University, Miriam Makebaplein 1, 9000 Ghent, Belgium.
E-mail: lieven.demarez@ugent.be

* Corresponding author

Abstract: Context-awareness is becoming an essential functionality of mobile applications. However, it remains challenging to capture the contextual experience in innovation research, since early-stage technologies have not reached maturity to be implemented in a real-life context. Moreover, users have difficulty in evaluating implicit interactions with context-aware interfaces since imagination of users is limited. Assuming that context impacts user experience, virtual reality (VR) provides an untapped potential for the domain of innovation research. The aim of this study (in progress) is to investigate the potential of user tests in virtual reality (here virtual retail store) for human-computer interaction to better match the needs of users and designers. Initially, the mock-up has been implemented in a retail store with its context-awareness being simulated using the Wizard of Oz methodology (N = 18). This approach is found to be time-consuming and not sufficient for evaluating radical context-aware innovations.

Keywords: user experience; context-awareness; implicit HCI; disruptive innovation; Wizard of Oz; virtual reality; retail; user innovation

1 Introduction

77% of mobile device users enable location services when asked by an app, which indicates the increasing proliferation of tracking sensors (e.g., gyroscope, accelerometer, magnetometer) in mobile applications (In The Pocket, 2016). In literature, these applications are coined as "context-aware and self-adaptive technologies" that need to adjust to the user's situations, habits, and intentions, without interrupting and distracting the user (Evers *et al.*, 2014). Dey and Abowd (1999) defined a context-aware system as a technology that "uses context to provide relevant information and/or services to the user, where relevancy depends on the user's task." This implies that context is either explicitly or implicitly indicated by the user, and always related to a user's goals. In addition to explicit HCI, e.g. telling the computer what it is expected to do, these mobile applications are mostly used while doing something else, which is implicit HCI. Implicit HCI is "an action performed by the user that is not primarily aimed to interact with a computerized system but which such a system understands as input" (Schmidt, 2000, p. 192). From a user-centric point of view, this conceptualization is also valuable for investigating contextual user experience.

The shift in these interfaces, and thus new human-computer interactions (HCI), points towards the need for new methodologies to explore the still unmet user needs and wants, but also to validate these new user interactions. Testing context-aware and self-adaptive technologies in innovation research remains challenging, because of some reasons. First, it cannot be tested in a context-agnostic lab setting with little ecological validity. Multiple key requirements rely on the context of the interactions and thus are mostly used while doing something else (Schmidt, 2000; Neill *et al.*, 2007), such as the navigation app Waze notifying you to leave for your meeting, based on your calendar, current location, and traffic. Second, innovations in an early new-product development (NPD) stage, have not reached maturity to be implemented in a real-life context. As a result, user evaluation is usually only executed at the end of the innovation process resulting in only incremental changes. Third, usability guidelines and principles are lacking or not yet applicable for self-adaptive technologies (Evers *et al.*, 2014). While self-adaptive technologies aim to deliver the most appropriate service depending on the user's situation, usability is restricted to ease of use and understanding of a particular functionality (Evers *et al.*, 2014). These challenges in combination with increasing attention devoted to user experience (Rebelo *et al.*, 2012), mitigates new methodologies to evaluate user experience with context-aware mobile applications that goes further than "usability and task-oriented instrumental values" (Hassenzahl, 2005).

Although evaluating self-adaptive technologies in a real context, such as retail, is promising. Still, some unanswered questions remain. What are the user experiences of these disruptive innovations, and how is this user experience ideally evaluated since it enables users to do what they never did before? Assuming that users lack knowledge and motivation to assess disruptive technologies (Lettl, 2007), and also context determines user experience (Neill *et al.*, 2007; Rebelo *et al.*, 2011), Wizard of Oz (WOz) tools and VR simulations of product usage contexts provide untapped potential for the domain of innovation research to explore this contextual user experience. WOz tools let designers evaluate design assumptions in realistic situations without being constraint by technical requirements, and investing time and money to build a complete prototype (Dow, MacIntyre, and Lee, 2005; Li, Hong, and Landay, 2007). Moreover they support

researchers to explore users' preferred interactions (Dow, MacIntyre and Lee, 2005). Taking product testing to the next level, VR simulations of product usage contexts can automate wizard tasks, which relieves the wizard from routine tasks (Li, Hong, and Landay, 2007). In contrast to experiments that are often criticized for their lack of ecological validity, prototype testing in VR benefits from a more ecological evaluation (Rebelo *et al.*, 2011, 2012).

Research question 1: What are the strengths and weaknesses of testing a context-aware shopping assistant using the Wizard of Oz methodology?

Research question 2: What would be the added value of VR in this process?

The structure of this paper is as follows. First, the current research context is delineated. Second, insights of applying the Wizard of Oz methodology are presented. The discussion section is devoted to a reflection on the strengths and weaknesses of the WOz methodology, and opportunities of virtual reality as an empirical research tool.

2 Methodology

Research design

The LUNAR research project aims to develop an accurate and scalable solution for the real-time tracking of customers' shopping carts in retail markets using ultra-wideband technology. To meet the needs of the customer, a user-centric and context-aware shopping assistant will be developed that can be accessed through the tablet that is attached to the shopping cart. Three co-creation workshops with customers guided the user-centric innovation development process; starting from the exploration of needs and frustrations, towards a long-list of potential features and requirement, and the concept evaluation. Stakeholder workshops revealed additional (technical) requirements and opportunities. Based on feedback on the clickable mock-up during the last co-creation workshop, the concept has been improved.

The contextual evaluation of the clickable mock-up proceeds in two stages. First, the shopping assistant is evaluated by the end-user using the Wizard of Oz methodology in a real supermarket. Eighteen participants (72.2% men, M_{age} = 37.29 years, min_{age} = 24, max_{age} = 59) were invited to test the clickable mock-up, each user test lasting about one hour. The user test started with introducing the study, taking a pre-test survey, and preparing the eye tracking glasses (SMI ETG) setup, i.e. fit and calibration. Once the participant was equipped, s/he was guided to the shopping cart on which a GoPro was mounted, as well as the shopping assistant tablet. During the test the participant was shadowed by an experimenter (the "wizard") and an observer. The experiment ended with a post-test survey and evaluative interview.

In a second stage, the shopping assistant will be implemented in a virtual retail store. The computer-generated retail store has already been developed in VR. However, the clickable mock-up should yet be integrated. This study will immerse the user in a fully immersive and interactive VR through HTC VIVE's head-mounted display and remote controllers (Gutiérrez, Vexo and Thalmann, 2008). The high level of immersion allows full agency

(e.g. guiding a shopping cart) and perception through vision, sound and touch feedback (e.g. controller vibrates if you drop a loaf of bread) (Rebelo *et al.*, 2012; Olsson *et al.*, 2013). The level of realism of the virtual retail store is high, because of its high level of detail.

Figure 1 Exemplary view of the real (left) and virtual store (right).

Procedure

The participants were randomly assigned to one of the two conditions: with or without a shopping list. Both conditions consisted of buying six products, which were mostly the same. The conditions differed in the availability of a digital shopping list in the application, as well as the tested functionalities. The users in the "shopping list"-condition had to, amongst other things, compare nutritional product information and watch an ad at the checkout. The users in the "no shopping list"-condition were provided with a recipe inspiration and needed to notify the isle manager using the application because the fresh chive was out of stock.

Afterward, the participants were instructed to fill out the post-test survey, and a post-session interview took place drawing upon the issues that raised during the testing.

3 Preliminary results

On the one hand, this study provides insights to improve the shopping assistant concept and its interactions with the user that enhance the in-store experience. On the other, this study evaluates the Wizard of Oz methodology to improve the internal and ecological validity of future user tests.

Regarding the design process of the mock-up, we found that participants' eye gaze was directed more than half of the time towards the shopping assistant tablet. Accordingly, ad-hoc obstacles, which were not mentioned in the mock-up, were not noticed, such as other customers and loading carts. Furthermore, the high number of saccades (i.e., the rapid movement of the eyes to reorient gaze) shows that there is room for improvement regarding the usability, as well as the user experience of the shopping assistant.

> "The application shuts you from the outside world. When I am doing the groceries in the supermarket I am used to go, I want to talk with people."
>
> (man, 48 years old, shops multiple times a week)

Concerning the implementation of the Wizard of Oz methodology it was observed that the tasks were too rigid, as people were forced to change their shopping habits (e.g., "*When you see the pasta pesto recipe card, add it for two people*"). The suggested navigation of the shopping assistant, for example, led directly to the location of the product that was next on their shopping list. However, efficient customers can either navigate through the entire store, as they get the appropriate product upon recognition or go directly to the product that they need, without crossing every store department. Hence, the mock-up flow only meets with the first type of efficiency.

In the post-test and the evaluative interview contextual elements were explored that might have hindered their shopping experience. The bugs of the shopping assistant and social context, such as the other customers, were evaluated as the most hindering contextual elements (see Table 1). Although the interpretation of statistical analyses warrants caution because of the small sample (N = 18), Spearman's correlation coefficient (r) was used to investigate if the perceived hindrances of contextual elements related to the technology acceptance model. It is found that only perceived time pressure is negatively correlated to the attitude towards the shopping assistant (r = -.723, p = .002). Based on these preliminary insights it is found that almost none of the contextual elements are linked to the acceptance of the shopping assistant.

Table 1 Evaluation of disturbing contextual elements during user test and Spearman's bivariate correlation coefficients with the Technology Acceptance Model
(N = 18, 1 = not at all disturbed - 5 = disturbed a lot)

Disturbing contextual elements	*Median*	*Perceived usefulness*	*Perceived ease of use*	*Attitude*	*Intention to use*
Infrastructure context	1,5	-0,044	-0,263	-0,334	-0,09
Shopping assistant tablet	2,0	-0,036	0,164	-0,149	0,191
Social context	2,5	0,009	-0,446	-0,274	-0,165
Network connection	2,0	-0,063	-0,194	-0,313	-0,11
Bugs of the shopping assistant	2,5	0,156	0,075	0,171	0,172
Eye-tracking glasses	2,0	0,325	0,196	0,299	0,41
Atmosphere	1,0	-0,32	-0,152	-0,323	-0,284
Time pressure	1,0	-0,155	-0,268	-,723**	-0,313
Concern about surveillance	1,0	-0,036	0,046	-0,311	-0,235

4 Discussion

The early evaluation of contextual concept testing leads to a better matching of the needs from users with the solutions from designers. Although users evaluated the context-aware shopping assistant as innovative and useful in the co-creation workshop, once implemented in a real supermarket they were less enthusiastic about the innovation. This proves the assumption of Lettl (2007) that it is hard to evaluate radically new ideas, as people acted differently than they supposed in the co-creation workshop (saying versus doing). Researchers and stakeholders have underestimated the social experience of shopping since the shopping assistant is perceived as too intrusive by the participants.

In addition to concept enhancement, also a number of methodological consequences resulted out of this study. First, participants reported a lack of agency during the contextual testing, as only two customer journeys (with and without a shopping list) were programmed in the clickable mock-up. Moreover delay, due to incorrect wizard behavior, is perceived as one of the most disturbing contextual elements during the testing (Dow, MacIntyre and Lee, 2005). These findings suggest that future studies using the Wizard of Oz methodology should rather evaluate the human-computer interaction than the concept and its features. This implies that participants should fulfill a more central role than the technology, which also reflects in more general, but relevant tasks (Dahlbäck, Jönsson and Ahrenberg, 1993). Second, participants were observed by some sources in this study, i.e., GoPro mounted on the front of the shopping cart, observer, "the wizard", and the eye-tracking glasses. Although participants did not evaluate "concern about surveillance" as a disturbing contextual element, this resulted in socially desirable behavior.

Based on the abovementioned issues, we suggest that virtual reality as an empirical research tool might counter most of the issues associated with the Wizard of Oz methodology. The added value of VR lies in its low threshold to evaluate a digital product before launch in a controllable and replicable environment to drive the user-centric dimension of innovation. Because of its automation, alternative contextual uses and responses can be explored (Bauer, 2016). Currently, some retailers already developed a virtual experience of their brick-and-mortar store for shelf planning, etc. Once this virtual environment (VE) is developed, the threshold to continuously evaluate context-aware innovations is lower, and afterward designers are still able to conduct significant changes in the prototype to gain competitive advantages (Soares and Soares, 2015). Moreover, automated data collection is enabled such as eye-tracking through the head-mounted display, which counters the perceived surveillance of participants (Stengel *et al.*, 2015).

5 Areas for feedback & development

Areas for feedback and development cover guidelines on the use of VR for prototype testing in innovation research, as well as methodological frameworks to evaluate user experience in a (virtual) context.

6 References

Bauer, Jared S. "Communicating Meaning in Context-Aware System Design." Ph.D. diss., University of Washington, 2016.

Dahlbäck, Nils, Arne Jönsson, and Lars Ahrenberg. "Wizard of Oz studies - why and how", *Knowledge-Based Systems* 6, no. 4 (1993): 258–266.

Dey, Anind K. and Gregory Abowd D. "Towards a Better Understanding of Context and Context-Awareness," *Computing Systems* 40, no. 3 (1999): 304–307.

Dow, Steven, Blair MacIntyre, and Jaemin Lee. "Wizard of Oz support throughout an iterative design process," *Pervasive Computing, IEEE CS and IEEE ComSoc* 4, no. 4 (2005): 18–26.

Evers, Christoph, Romy Kniewel, Kurt Geihs, and Ludger Schmidt. “The user in the loop: Enabling user participation for self-adaptive applications,” *Future Generation Computer Systems*. 34 (March 2014): 110–123.

Falcao, Soares C. and Marcelo Soares M. “Application of Virtual Reality Technologies in Consumer Product Usability,” *International Conference of Design, User Experience, and Usability* (june 2013): 342-351.

Gutiérrez, Mario, Frédéric Vexo, and Daniel Thalmann. *Stepping into virtual reality*. Lausanne, Switzerland: Springer, 2008.

Hassenzahl, Marc. “The Thing and I: Understanding the Relationship Between User and Product,” in *Funology: From Usability to Enjoyment*, eds. M. A. Blythe, A. F. Monk, K. Overbeeke and P. C. Wright (Kluwer Academic Publishers, 2005) pp. 1–12.

In The Pocket, "The State of Mobile." In The Pocket Company Web site, https://inthepocket.mobi/blog/2016/this-was-the-state-of-mobile-16, accessed January 2017.

Lettl, Christopher. “User involvement competence for radical innovation,” *Journal of Engineering and Technology Management - JET-M* 24, no. 1–2 (2007): 53–75.

Li, Yang, Jason I. Hong, and James. A. Landay. “Design Challenges and Principles for Wizard of Oz Testing of Location- Enhanced Applications,” *IEEE Pervasive Computing* 6, no. 2 (2007): 70–75.

O'Neill, Eleanor, David Lewis, Kris Mcglinn, and Simon Dobson. “Rapid User-Centred Evaluation for Context-Aware Systems,” in *Interactive Systems: Design, Specification, and Verification*. eds. G. Doherty and A. Blandford (Ireland: Dublin: Springer-Verlag Berlin, 2007) pp. 220–233.

Olsson, Thomas, Else Lagerstam, Tuula Kärkkäinen, and Kaisa Väänänen-Vainio-Mattila. “Expected user experience of mobile augmented reality services: A user study in the context of shopping centres,” *Personal and Ubiquitous Computing* 17, no. 2 (2013): 287–304.

Rebelo, Francisco, Emília Duarte, Paulo Noriega, and Marcelo M. Soares. “Virtual Reality in Consumer Product Design: Methods and Applications,” *Human Factors and Ergonomics in Consumer Product Design* (June 2011): 381–402.

Rebelo, Francisco, Paulo Noriega, Emília Duarte, and Marcelo M. Soares. “Using virtual reality to assess user experience,” *Human Factors- The Journal of the Human Factors and Ergonomics Society* 54, no. 6 (2012): 964–982.

Schmidt, Albrecht. “Implicit Human Computer Interaction Through Context,” *Personal Technologies* 4 (2000): 191–199.

Stengel, Michael, Steve Grogorick, Martin Eisemann, Elmar Eisemann, and Marcus Magnor. “An Affordable Solution for Binocular Eye Tracking and Calibration in Head-mounted Displays,” *Proceedings of the 23rd ACM international conference on Multimedia* (2015): 15–24.