On the Use of Virtual Environments for the Evaluation of Location-Based Applications

Arief Ernst Hühn[1], Vassilis-Javed Khan[1], Andrés Lucero[2], Paul Ketelaar[3]
[1]NHTV Breda, University for Applied Sciences, P.O. Box 391, 4800 DX Breda, The Netherlands
[2]Nokia Research Center, P.O. Box 1000, FI-33721 Tampere, Finland
[3]Radboud Universiteit, P.O. Box 9104, 6500 HE Nijmegen, The Netherlands
{huhn.a, khan.j}@ nhtv.nl; andres.lucero@nokia.com; p.ketelaar@ru.nl

ABSTRACT
User experience (UX) research on pervasive technologies faces considerable challenges regarding today's mobile context-sensitive applications: evaluative field studies lack control, whereas lab studies miss the interaction with a dynamic context. This dilemma has inspired researchers to use virtual environments (VEs) to acquire control while offering the user a rich contextual experience. Although promising, these studies are mainly concerned with usability and the technical realization of their setup. Furthermore, previous setups leave room for improvement regarding the user's immersive experience. This paper contributes to this line of research by presenting a UX case study on mobile advertising with a novel CAVE-smartphone interface. We conducted two experiments in which we evaluated the intrusiveness of a mobile location-based advertising app in a virtual supermarket. The results confirm our hypothesis that context-congruent ads lessen the experienced intrusiveness thereby demonstrating that our setup is capable of generating preliminary meaningful results with regards to UX. Furthermore, we share insights in conducting these studies.

Author Keywords
User experience evaluation; mixed reality; location-based services; pervasive computing.

ACM Classification Keywords
H.5.m [Information interfaces and presentation (e.g., HCI)]: Miscellaneous.

INTRODUCTION
The advent of new mobile devices like smartphones and tablets has contributed substantially to the momentum of the adoption of pervasive computing. Although these devices share many characteristics with their static counterpart, the desktop computer, their mobile nature enables applications to bring their functionalities into the user's dynamic context. Apart from location independence, applications also gain the ability to interact directly with the specific context of the user, due to context-sensitive technologies (e.g. camera, GPS, compass, accelerometer, etc.). Augmented reality and location-based services (or a combination of these) on a mobile device are popular examples that tie in closely with the ever-changing environment of the user and have attracted much attention from advertisers and marketing scholars [10]. However, advertising research is limited when it comes to the evaluation of context-sensitive applications because of the lack of means to deal with the context-prone and dynamic research setting pervasive technology poses. Because of the lack of better alternatives advertising researchers mostly focused on simple forms of advertising (simple push SMS/MMS-advertising) and their methods generally relied on surveys with scenario descriptions [4,6,40,42,43], thereby generating questionable results.

In contrast, to the field of human-computer interaction (HCI), mobility and context-sensitivity are already familiar challenges when it comes to the evaluation of pervasive applications. Until relatively recently mainly two types of studies were engaged to evaluate mobile applications: laboratory and field-test studies [19]. Researchers have been divided about which method is more effective and efficient [30]. Main objections to field studies are concerned with the lack of control and the expensive and time-consuming endeavours associated with this type of research [21, 35]. In contrast, lab studies do not offer an immersive and interactive context for the user, thereby overlooking possible important influences that determine the everyday practice of these pervasive applications [21]. Furthermore, some mobile applications are currently difficult to deploy because the supporting technology (e.g. indoor positioning, RFID, etc.) is not yet in place [5].

These shortcomings have given rise to a new line of research where lab studies have been extended with the utilization of virtual environments (VEs). This offers participants a dynamic and interactive context during their experience while researchers gain a controllable and malleable research setting [24]. Although very promising, the aforementioned studies are mostly limited to usability studies. Moreover, the setups from these studies leave room for improvement in terms of display methods, navigation, and interactivity between the VE and the mobile app.

This paper builds on this line of work by introducing a working concept with a distinct combination of features: a mixed reality setup consisting of an interactive 360° CAVE (Cave Automated Virtual Environment) installation interacting with a location-based advertising application running on a smartphone [18]. To show the potential relevance of this approach we present a case study in which we conduct two experiments with this setup. These studies evaluate the UX of a location-based advertising application within the context of a virtual supermarket. Based on literature and empirical data, it is our aim to tentatively answer the following explorative research question: *Can we generate meaningful research conclusions about the UX of pervasive advertising apps in our mixed reality setup?* The answers and provided insights should help advertising and HCI researchers to conduct virtual reality aided evaluation studies of their own, thereby benefitting the design process of location based advertising applications and pervasive applications in general.

Before we discuss the specifics of our case study, we present the limitations in advertising research and review the different mobile evaluation methods in HCI.

BACKGROUND

Mobile Advertising Research

Since the advent of pervasive computing the advertising business has increasingly concerned itself with mobile technology as means to disseminate ads. As a consequence, mobile advertising has attracted the attention of advertising scholars to scrutinize its effectiveness. These studies investigated several mobile ad properties like for example text versus graphic [43], push versus pull [40], permission based versus non-permission based [8] situational context [42], location congruent versus location incongruent [4] and their influence on subjective experience either based in cognition or attitude (e.g. ad irritation, ad intrusiveness and attitude towards the mobile ad, perceived usefulness). Subsequently Theory of Reasoned Action [14], Theory of Planned Behavior [3] and Technology Acceptance Model [9] offer a theoretical framework which links cognition and attitude to the behavioral intentions and behaviors that are of relevance to advertisers (e.g. intention to use a technology, intention to buy).

Results are usually obtained through surveys using scenario descriptions of use-context. Marketing studies using real encounters with mobile applications (e.g. [23]) are on the other hand scarce. Although these scenario-based studies still yield interesting results regarding expectations people have, they are questionable since the participants have not interacted with the application itself, let alone within the use context [42, 43].

This lack of data based on real contextual interaction signals the need for new approaches to study these phenomena. In this respect HCI methodology could be valuable to the field of mobile ad research since it is already familiar with the challenges of evaluating mobile applications. Mobile advertising on the other hand serves as an appropriate vehicle to illustrate the potential of these evaluation methods within the field of HCI.

Mobile Application Evaluation: Field Versus Lab

Field studies are praised for taking the real life context into account. When it comes to studying the UX of mobile context-sensitive apps, where a dynamic context is salient, field studies seem to be even more appropriate. However, methods examining phenomena in a real life context are not widely used [19]. A possible reason is that these methods are time consuming, costly [35] and sometimes rely on third parties to acquire the necessary infrastructure [42]. Moreover, some future concepts are hard to deploy in the real world, due to current technical constraints [5].

In addition to these practical issues, there are also some methodological challenges that keep researchers from using field studies on a frequent basis when evaluating mobile apps. Kjeldskov and Stage [20] argue that the mobile nature results in an unstable context with detrimental consequences. They report difficulties in terms of observation, manipulation and control of the use-context. In line with this conclusion Roto et al. [35] report that investigating mobile phone use in the wild with a quasi-experiment *"is a laborious method and requires careful planning and vast technological resources."*

Instead, a great amount of user studies is conducted within laboratory settings. Kjeldskov & Graham [19] conclude that 71% of the mobile evaluation studies they reviewed take place in laboratories. The main advantage is the control researchers have on the use context thereby excluding confounding variables and securing reproducibility. Certain evaluations are reasonably executable in lab settings when they only take the necessary environmental aspects into account. Lumsden [27] for example evaluated the quality of three mobile speech recognition systems where the environmental noise was artificially recreated with a 7.1 surround system.

Nonetheless there are drawbacks to laboratory studies, such as the lack of a realistic and immersive context [19]. Although real life environments can be recreated physically within lab studies to a certain extent, most mobile use contexts ask for an extensive and dynamic environment instead of a small sized living room or shop, as is the case with most lab studies (e.g. [23]). In the case of evaluation of context-aware mobile systems the limitations of lab studies are even more prevalent since context not only serves as a backdrop, but also plays an active role in the use of the application [1,36]. Contrary to the findings of Kaikonen et al. [16] and Kjeldskov et al. [21], Nielsen et al. [30] and

Duh et al. [12] found more usability problems in the field than in the lab attributed to a different interaction style, cognitive load and a richer environment in the field.

These shortcomings on both field and lab studies, point out the need of an approach that provides on the one hand a controllable setting and on the other hand a rich context users can relate to and interact with while keeping costs, time and organizational effort as low as possible [20,24,36].

Evaluation with Virtual Environments (VEs)

Inspired by the shortcomings of these traditional methods recent studies have tried to address these issues by using virtual environments (VEs). Ubiwise [5] is one of the first research efforts to evaluate pervasive applications with the assistance of a VE. Since that first attempt there have been several other prototypes presented with the same approach [5,24,32,34,36,38,39].

Not only do VEs improve control [24,39], but they also give the researcher the possibility to test mobile concepts that are hard to deploy in the real world [5]. Furthermore, the opportunities regarding measurement will improve when conducting studies with the use of VEs. Advanced methods like psychophysiological measurements, video recording and behavioral tracking are easier to implement because the researcher is given a stable research setting which simplifies the placement of measuring equipment. In the study by Schellenbach et al. [36], VEs enabled them to use a motion capture system to measure the effect of specific interactions with a mobile navigation system on the gait of participants. To set up such a motion capture system in the field or recreate the use-context (in their case study a museum) would be a daunting task.

When comparing the different setups reported in literature, we noted important similarities and differences. Most of these setups have in common that they use hybrid simulations, i.e. "*setups that rely on the integration and combination of the real and virtual world.*" This enables participants to directly interact with a physical mobile device [24], whereas full virtual simulations would have represented the device virtually, forcing its interaction to take place indirectly (e.g via keyboard and mouse), which could disrupt the usage [24]. In terms of differences we noted that some evaluation studies used desktop setups (e.g. [5,24]) while others have moved to large-scale panoramic screens [39,38,22], which increases the feeling of presence and immersiveness [26,39]. Although not necessarily detrimental for their research goal, these screens displayed only environments using prerecorded photos or video lacking the possibility to navigate through the represented environment. In contrast, evaluations that relied on a navigable environment used 3D computer rendered environments [5,24,32,36]. Navigation style within these studies ranged from traditional input methods to more natural forms like treadmill based movement. The latter seems preferable since controls involving the mapping of body movement not only evoke higher presence [37] but also seem to result in a better representation of the cognitive state during the use of applications in the real world [39]. In addition, audio, haptic and olfactory stimuli in a VE could also increase the sense of presence even further [11]. Nonetheless the development of these test beds cannot only rely on theory but should also be validated in the field to strengthen its external validity. Concordantly, several studies have already planned to replicate their evaluation study in the field [22,36,39].

Based on these comparisons we can reason that our CAVE-smartphone interface apparently extends earlier approaches by having a distinct combination of features: a 360° view CAVE, surround sound, navigation through full-body movement and the direct interaction between virtual location and the mobile application.

Furthermore, we have not yet encountered a study within VE aided evaluation focusing on the UX of pervasive applications. Instead most studies have focused on usability evaluation and the technical realization of these test beds. Our case study gives some preliminary insights into how VEs can contribute to the UX evaluation of mobile apps.

UX and Ad Intrusiveness

UX in HCI is a very broad concept and still undergoes semantic negotiations. Generally there are two approaches to UX: reductionist and holistic [17]. The reductionist approach chooses to identify distinct psychological constructs, while the holistic approach looks at experience as unique, highly depending on the specific situation and people. Since our case study is grounded in the motivation of advertisers, we are specifically interested in the way location-based applications can improve the user's experience of advertising. This involves a construct that takes into account the subjective experience of the user in relation to: 1) the medium (application), 2) the content of the medium (in our case advertising), and 3) the context. Moreover, the construct needs to be a good predictor for behavioral intention/behavior to be of use to advertisers. These requirements obviously place us within the reductionist approach.

The concept of perceived intrusiveness, originally used within persuasive communication, concerns itself with the experience people have when their cognitive processes are interrupted by an advertisement, which have proven to be good indicators for behavioral intentions (e.g. ad avoidance, [25,28]. Intrusiveness regarding context is originally defined by Ha [15] as 'the degree to which ads in a media vehicle interrupt the flow of an editorial unit'. Li et al. [25] expands the concept by redefining "editorial unit" as "all new environments in which ads appear", thereby including advertising formats that not only interrupt media content but also other contexts, as is the case with mobile advertising. Since mobile devices are ubiquitous due to their mobile nature, mobile ads can potentially interrupt the user anytime and anywhere [6]. Location-based applications offer the opportunity to adjust the ad to the physical environment of the user. Edwards et al. [13] state that context-congruent ads can lower the intrusiveness by reducing the divergent knowledge structures that would normally be activated by context-independent ads. Based on the existing theory, our hypothesis for our case study is:

H1: Location-based ads will be perceived as less intrusive compared to location-independent ads.

METHOD

Before presenting the two experiments, we describe the general setup and procedure that both experiments shared.

General Setup

The experiments took place within a virtual supermarket simulated by a CAVE. A supermarket seemed to be the most suitable setting since mobile ads are more effective if they are aimed at low-priced and frequently bought products [8]. The VE (modelled in Maya and rendered in OGRE) is projected onto four rear- projection screens (each 3.6 meters wide by 2.6 meters high). The screens form a closed space, thereby offering the participant a 360° view of the environment. Participants can move in the virtual setting with the help of a head-tracking system based on four Wii-mote IR cameras. In contrast with head-mounted displays, the CAVE does not block out the physical world, which offers the opportunity to use physical objects and the representation of the participant's own physical body. In this particular case, it gave us the possibility to use an actual smartphone, with which participants were able to receive location-based ads. The tracking device (Figure 1, left) determines the participant's head position in the physical room, which is then used to control motion in the virtual supermarket. In essence, the participant acts as a "human joystick": when the participant stands in the centre of the CAVE the virtual camera stands still, whereas when the participant takes one step in a certain direction, the virtual camera moves accordingly, thus giving the illusion of movement within the virtual space. The participant is able to turn and step in every direction relative to the CAVE's centre. The simulation is also sensitive to the magnitude of the participant's distance from the centre of the physical room. This distance, determines the speed with which a participant walks within the environment. The simulation also corrects the first person view for the vertical axis. Thus, in the case in which one, for example, jumped or ducked the simulation corrected the perspective according to the vertical position of the participant's head. When it comes to shopping within the supermarket our simulation does not yet support interaction with virtual products. To simulate the act of selecting a product, participants were asked to make a grabbing gesture, without actually attempting to interact with the virtual product (Figure 1, left). When the participant made this grabbing gesture, they received auditory feedback.

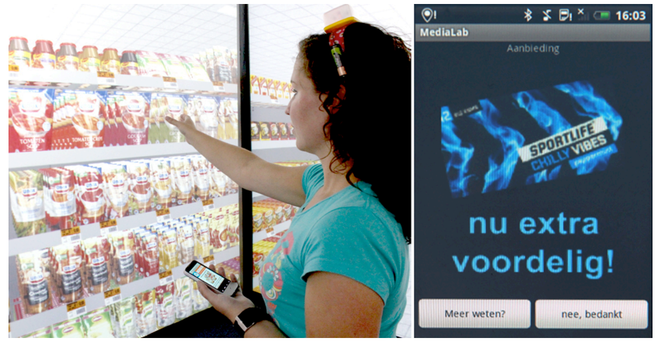

Figure 1. Left: Participant (wearing a headtracker and holding smartphone) makes a 'grabbing' gesture in the virtual supermarket Right:. The chewing gum ad (study 1) after entering trigger area. With 'know more' & 'no, thanks' buttons the participant can pull more info.

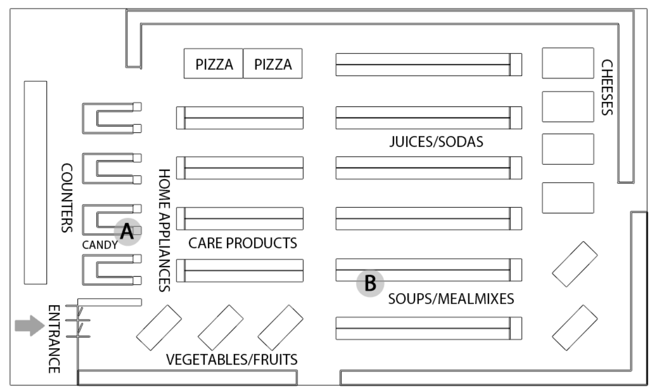

Figure 2. Floor plan of the virtual supermarket; location-congruent trigger area study 1 (A). Location incongruent area study 1 (B). Location (in)congruent area study 2 (B)

The interior of the supermarket (Figure 2), with regards to corporate style and spatial arrangement, was based on common denominators found in stores of popular Dutch supermarkets. The entire supermarket inventory was apparent and about one sixth of the shelves were filled with products at the time of the experiments. In the future it will be filled with a representative product set.

An Android application was developed, which connected the phone with the CAVE through Bluetooth. The application would play an audio message, vibrate and present the ad to participants. The application offered the option of receiving or rejecting more information about the product offer by using yes/no buttons. This ad was presented when the user was within a specific trigger area of the virtual supermarket.

General Procedure

During the briefing of both experiments we covered different topics: interacting with the smartphone and the CAVE, shopping instructions, notice for the slight risk of physical inconveniences (dizziness and nausea) due to disorientation in the CAVE, the length of the experiment (~30 minutes) and the confidentiality clause regarding the collected data. In addition, participants received tasks they had to complete during their virtual supermarket experience. These tasks involved doing shopping and varied slightly for each study to serve the specific research goal. After the briefing, participants were led to the CAVE where the head tracker was mounted and calibrated according to the specific height of the participant. Depending on the group the participant belonged to, the researcher had to start the corresponding supermarket version (congruent or incongruent). After the experiment the participant was asked to fill out a questionnaire.

STUDY 1: PILOT

Our aim with our pilot study was to investigate the user's experienced intrusiveness when receiving a location-based ad. For clarity reasons we repeat our hypothesis:

H1: Location-based ads will be perceived as less intrusive compared to location-independent ads.

Experimental Design

A between-subjects design was used to test the hypothesis. Participants were randomly assigned to two conditions. Location-congruent: The ad with product X was presented when the user was on location A, in proximity of product X. Location-incongruent: The ad with product X was presented when the user was on location B, not in proximity of product X. Proximity was set to 0.5 meter distance to the product. Product X was chewing gum; location A was at the counter close to the chewing gum; location B was at the soup shelf instead (Figure 2). The ad advertised chewing gum (Figure 1, right).

Participants

While having 27 participants, for the results described in this paper we decided to include the 12 participants whose experiment proceeded according to protocol; i.e. they noticed the ad on the mobile phone. These 12 consisted of 8 who witnessed the location-incongruent setup (4 male and 4 female; average age: 23) and 4 who witnessed the location-congruent (2 male and 2 female; average age: 21). All of them were international university students. Moreover, they all owned a mobile phone and were familiar with smartphones equipped with a touch screen interface. For their effort participants received a 5€ gift coupon.

Tasks

The initial task was to navigate through the supermarket for a few minutes to get familiar with the controls and the interaction. Next, participants were asked to go to the main entrance of the supermarket to receive the second task: buying a pizza. While executing the second task participants were given a pre-scripted phone call before reaching the pizza refrigerators. During this call the extra task of getting a soft drink and a soup package was given. After completing the task participants had to go to counter number 3. The shelves above the conveyor belt at that counter were filled with products including the advertised chewing gum. Depending on the condition a participant was assigned to, the chewing gum ad was triggered at one of two locations. The location-congruent condition triggered the ad within 0.5 meters of the chewing gum shelf (Figure 3, left) while the location-incongruent condition triggered the same ad within 0.5 meters of the soup shelf (Figure 3, right). The half-meter radius was determined by the researchers based on the observation that people usually enter this area when they want to look at products, whereas people that just walk by in the aisle do not trigger the ad.

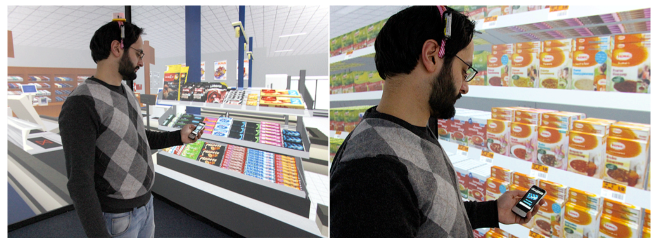

Figure 3. Left: Participant from the context-congruent group receives a chewing gum mobile ad in front of the chewing gum shelf. Right: Participant from the context-incongruent group receives a chewing gum mobile ad in front of the soup shelf.

These tasks ensured that both groups visited their trigger area without giving away our research goal. In the case of the location-congruent group we did not need to ask participants to seek the chewing gum shelf, since they would visit it while checking out at counter 3. For the location-incongruent group the task to pick up soup ensured all participants within the incongruent group experienced similar conditions while receiving the mobile ad.

Measurement

Immediately after the virtual shopping experience, participants filled out a questionnaire that assessed their perceived ad intrusiveness of the received mobile ad. In order to take into account the complexity of an everyday commercial setting we also measured several control variables that have proven to affect the evaluation of advertising and location-based advertising. Based on our literature study we included the following scales: product involvement [41], privacy concerns [7] and the attitude towards advertising in general [29].

Results

The 12 respondents rated homogeneous with regards to the control variables, thereby excluding the possibility of these variables explaining the measured effect. The median scores on the ad intrusiveness scale for the congruent and incongruent group were respectively 3 and 5 (7 point scale). Based on the Mann Withney U test we conducted we can conclude that the median scores on the intrusiveness scale differed significantly ($z=-2.722$ $p=0.003$). It can be further tentatively concluded that mobile context-congruent ads lead to less ad intrusiveness than mobile ads which are not context-congruent.

Figure 4. 'Incongruent Group' shelf setup (left): all six shelves offer only meal mixes. 'Congruent Group' shelf setup (right): top three shelves offer soup, including the advertised soups (highlighted); bottom three shelves offer meal mixes

STUDY 2: LARGE SCALE

Our second study approaches the same research hypotheses as our pilot. For clarity reasons, we state once again:

H1: Location-based ads will be perceived as less intrusive compared to location-independent ads.

For this study we also extended our theoretical scope to the consequences of intrusiveness for behavioural intentions. Theory of Reasoned Action [14], Theory of Planned Behavior [3] and Technology Acceptance Model [9] predicts that positive perceptions lead to a favourable attitude, which in turn has a positive influence on behavioural intentions. This theoretical framework has already been successfully applied to intrusiveness in the context of web ads [28]. Thus, we hypothesize:

H2: Lower level of perceived ad intrusiveness results in more favourable attitude toward the mobile advertising application

H3: Attitude towards the mobile advertising application is positively related to the intention to use the mobile advertising application

In the following section we present the motivations for the changes we introduced compared to our pilot (Study 1).

Experimental Design

For this study we again used a between-subjects design to test our hypotheses. Participants were randomly assigned to two conditions: 1) Location-congruent: the mobile ad with product x was presented when the user entered location B where product x, product group X and product group Y were available, 2) Location-incongruent: the mobile ad with product x was presented when the user entered location B where only product group Y was available (Figure 4). Contrary to study 1, we kept location constant and manipulated the products on the shelf to keep as many circumstantial factors constant as possible. Location B could be found approximately in the centre of the supermarket and covered the space with again a radius of 0.5 meter from the manipulated shelf (Figure 2). For product x, a well-known Dutch soup product was used. Accordingly, the ad showed an offer with this specific soup product (Figure 5). Further, product group X consisted of soups and product group Y consisted of meal mixes.

Figure 5. The soup ad for study 2 (left) the participants received when they reached the trigger location. Subsequently, participants could pull more information about the offer by using 'yes' and 'no' buttons (middle, right).

Setup

Compared to the pilot study, we improved the realism of our scenario by creating a simple UI of the application using the supermarket's corporate style. Furthermore, participants had to start the application themselves in contrast to the pilot where the application was already started and was running in the background. In this way we wanted to convey the idea to the participant that this is an opt-in service within the environment of an application instead of a simple push message.

Participants

We recruited 70 participants through a marketing research company. From the 70 participants 15 participants received the ad outside the perimeter as a consequence of an unstable Bluetooth connection and 2 participants did not notice the ad. The remaining 53 participants (31 male, 22 female; age range 17-64 years, M= 28.57 years, SD=11.60) consisted of 26 who witnessed the location-incongruent setup (15 males and 11 females; M=28.96 years, SD=14.22) and 27 who witnessed the location-congruent (16 male and 11 female; M=28.19 years, SD=11.22). Moreover, they all owned a mobile phone and were familiar with smartphones equipped with a touch screen interface. For their effort participants received a 5€ coupon.

Tasks

Participants were given two tasks during the briefing: first, they freely navigated in the supermarket for a few minutes to get familiar with the controls and the interaction. Second, they had to pick five different food products and then to go to the counter. The central location of our trigger area, the task of shopping five products and the limited product set of the virtual supermarket ensured that both groups visited the trigger area. In this way we did not have any specific instructions for participants to visit the manipulated shelf, thereby keeping our research goal concealed.

Measurement

After the virtual shopping experience, a questionnaire assessed the perceived intrusiveness (C's alpha=.841), based on the scale of Li et al. [25], attitude toward the mobile application (single item) based on a scale of Bergkvist and Rossiter and measured their intention to use the application (C's alpha=.921) based on a scale of

Kowatsch and Maas [23]. As was the case with our pilot study we also took into account relevant control variables: product involvement (C's alpha=.842) [41], the attitude towards advertising in general (C's alpha=.841) [29] and personal innovativeness (C's alpha=.845) [2].

Results

The variable *intrusiveness* shows a statistically significant higher value with the incongruent-group (M=3.15, SD=0.90) than with the congruent-group (M=2.45, SD=1.07); (t(51, N=53)=-2.546, p<.05), thereby pointing in the direction expected by H1. The ANCOVA (see Table 1) confirms the significant effect of the congruent/incongruent condition when controlled for *product involvement* and *attitude toward advertising in general*. It also shows that the congruent/incongruent condition has the greatest effect (η^2=0,094). H1 is thereby supported by the data. Furthermore, we tested the remaining hypotheses by conducting a regression analysis. We once again took the congruent/incongruent condition (incongruent=0, congruent=1) into account and found a significant effect in the expected direction (β=-*0*.286, p<.05). The congruent/incongruent condition explains together with the control variables 23.2% of the variance of *intrusiveness*. *Intrusiveness* in turn has a significant effect on the *attitude toward the app* (β=-0.173, p<.05) and is responsible for 14.1% of the variance, thereby supporting H2. *Attitude toward the app* has a positive influence on 'intention to use the application' (β=0.393, p<.001) when controlled for *innovativeness* (H3 supported). Together they explain 48.1% of the variance of *intention to use the application*.

Variable	*df*	*MS*	*F*	*partial* η^2
Product involvement	1	1.621	1.826	.036
Attitude ad in general	1	4.369	4.921	.091*
Context congruency	1	4.514	5.085	.094*
Error	49	.888		
Total	53			
Corrected Total	52			

Table 1. ANCOVA for 'intrusiveness'

GENERAL RESULTS AND LESSONS LEARNED

The results from both studies show us that location congruency is a significant factor when it comes to lowering the intrusiveness people experience when using a mobile (advertising) application. Furthermore, users of a context-congruent application tend to form more positive attitudes toward the application and as a consequence are more likely to use it in the future.

In addition to the gathered quantitative data, these studies also provide insight into the specifics of conducting pervasive application evaluations using VEs. With regards to the setup, scenario and experimental design we had to be creative to deal with the challenges that this kind of research poses. Firstly, given the fact that the setup did not support any interaction with the virtual products, the participants were asked to make a 'grabbing' gesture towards the product of their choice. To give participants feedback on whether they were successful in their attempt, a 'wizard of Oz' procedure was implemented: as soon as the participant made the grabbing gesture the researcher manually played a confirmation sound. Since the CAVE installation formed a closed space, separating participants from the researcher, the use of video cameras was essential in observing the participant's actions within the CAVE.

When it comes to the experimental design, the concept of 'context-congruency' can be operationalized by manipulating one of the following factors: 1) by altering the product in the advertising application, 2) by altering the location where the ad is triggered (thereby obviously altering the context in which the ad is received) or 3) by replacing the advertised product in the shelf with a different product. In our pilot study (Study 1) we altered the trigger location thus manipulating the aforementioned factor 2. This resulted in the 'context-congruent' group receiving a chewing gum ad at the counter where chewing gum was displayed while the 'context incongruent' group received the chewing gum ad at the soup shelf. Although altering the location seemed a rational choice, we realized that it compromises the requirement for experiments to keep circumstantial factors constant. Receiving an ad at the supermarket counter compared to receiving an ad at a supermarket aisle not only manipulates context-congruency but also possibly includes other situational factors that could unintentionally affect our dependent variable. Therefore, in our second study we decided to keep the location of the trigger area for both groups the same and to change the products that are within the vicinity of the trigger location instead. Because of the plasticity of our VE it was fairly easy to replace the congruent products (soup) with incongruent products (meal mixes). The third aforementioned factor we could manipulate, i.e. changing the product within the ad, was discarded as an alternative since it could affect the participant's product involvement towards the advertised product, which in turn influences their experienced intrusiveness [42].

When scrutinizing the effect of context-congruency, it is important for the research goal that participants reach the trigger area and notice the products. The importance lies in controlling any biases participants might experience in case their task description includes products in the ad's location-triggering area. Since the empty shelves were already common in the virtual supermarket, we were able to empty the opposing shelf to ensure that the participants would only notice the (in)congruent products while receiving the ad with their application. In addition, we placed the trigger area in a central location of the supermarket to increase its accessibility. This once again illustrates the benefits of the dynamic configuration of the VE in helping the researcher to adjust the setting to the needs of the study. Furthermore, participants can also be 'guided' towards the trigger area through the task scenario. In our first study we used the shopping list and counter to draw people towards their respective trigger areas. The second study had a different approach. During our pilot we experienced that the limited product set of the supermarket already 'forced' the participants to visit all the filled shelves, since there were not many other places of interest. Just by giving a generic task where participants had to shop for five products sufficiently increased the chance they reached the trigger area. This caused all participants to visit the trigger area.

CONCLUSION

This case study presented a setup with a CAVE-smartphone interface that was used to evaluate the UX of a location-based advertising application. We conducted two experiments that evaluated the perceived intrusiveness of location congruent versus location incongruent mobile ads. The results from both studies show a statistical significant difference between the two groups; participants receiving a location-congruent mobile ad perceived it as less intrusive compared to participants receiving a location-incongruent mobile ad, thus confirming our central hypothesis. Furthermore, our second study confirmed our other hypotheses: users of a context-congruent application tend to form more positive attitudes toward the application and as a consequence are more likely to use it in the future and to buy the advertised product. Thereby the results seem to support the underlying theory [3,9,14,25] and contribute to the knowledge of persuasive communication by applying the construct of intrusiveness to pervasive advertising.

A more abstract, yet tentative methodological conclusion (see Limitations and Future Work section) is that such a setup enables mobile and pervasive computing researchers as well as advertisers/advertising scholars to conduct experimental studies to yield meaningful results regarding the UX of mobile applications. Our literature study showed that such a setup has substantial advantages when compared to field and lab studies. When compared to field studies, such a setup can help researchers control and manipulate context parameters, enable the replication of studies and apply extensive measurement. When compared to lab studies, this setup can introduce malleable, immersive and interactive experiences with the envisioned context of use [39]. As described earlier, we were able to manipulate location (study 1) and context (study 2) fairly easily when compared to the effort one would need in the field. In contrast, Nurmi et al. [31] conducted a field study in a real supermarket and reported difficulties in modifying the environment due to restrictions from the supermarket. Compared to the existing literature our setup is a useful expansion as it offers a distinct combination of features like an interactive virtual environment (VE), 360° view, navigation controls with body movement and a virtual localizing system interacting with a physical smartphone. Nonetheless, its value needs to be verified by comparing this setup with other approaches.

Furthermore, based in our experience we provided guidelines for researchers planning to use such setups for evaluating context-aware pervasive computing applications.

Naturally, developing such an installation would initially require a considerable investment, but in the long run it would save time and money when it comes to performing user evaluations [24]. Furthermore, since the setting as well as the pervasive technology infrastructure takes place within the simulated environment, it does not rely on third parties for the operationalization.

LIMITATIONS AND FUTURE WORK

Although our results resonate with theory and expectations, they remain tentative due to certain limitations. Since we have not yet conducted a validation study, there is still uncertainty with regards to the external validity. We see two important factors that could have decreased the generalization of our results. In the first place we observe a potential tradeoff between the ecological validity and the internal validity. To strengthen the internal validity of our study we had to investigate the relationship between the UX and context-congruency within a vacuum of some sort. Conducting the experiments within a VE helped us achieve that goal, due to high control and manipulability; we were able to customize the shelves and filter out confounding variables. However, these adjustments could have affected the outcome in such a way that the results cannot speak for the use context we wanted to investigate in the first place (i.e. real world supermarkets). This obviously has little to do with the use of VEs but more with the inherent properties of an experimental setup that investigates a specific causal relationship.

Secondly, we can speculate about the possibility that the limitations of our setup and simulation affected the ecological validity. The interaction with the VE, like the navigation style and manipulation of objects differed from physical reality. The absence of a complete product set is also a noteworthy difference. Furthermore, the presence of supporting equipment (head tracking system, beamers etc.) could have impaired the level of immersiveness. Validation studies that compare the results of a CAVE with a real life supermarket, could give some insight in to what extent the results are affected by these limitations. However, we must note that these shortcomings are not necessarily inherent to the principle of a virtual reconstruction of UX, but are caused by our specific operationalization and by the limitations of current technology. As technology progresses certain aspects of the setup could be improved. In our case we are already in the process of extending our product set, improving the graphics and have already replaced the Wii-mote based head-tracking with the help of a Microsoft Kinect, thereby making the facilitating technology transparent to the user. Finally this issue raises the philosophical question whether such setups can become truly invisible and match field studies in realism. A pragmatic approach would be that the facilitating technology should be as unobtrusive as possible so that the participant can focus better on the reconstructed UX. Results from validation studies should guide us in this process of improvement.

In addition, the setup could also affect the outcome as a consequence of Hawthorne effects. Even though participants got the chance to get used to the setup, there is the risk that they behaved differently because the setup still reminded them they were taking part in an experiment.

Apart from the external validity we also see some room for improvement in terms of measurement. In our case we relied on questionnaires to gather our data. Even though the construct of intrusiveness is usually measured with self-report Likert scales, the delay between the stimulus and measurement could have affected the outcome. Since the use of the CAVE setup allows for elaborate measurements, direct measures like observations or psychophysiological measurements should be considered in the future.

ACKNOWLEDGMENTS

We would like to extend our gratitude to our financial sponsors: the Dutch Ministry of Education, Culture and Science (OCW), RAAK scheme (RAAK Pro) and Popai Benelux. Moreover, we would like to thank Hans Bouwknegt, Koos Nuijten and Marnix van Gisbergen for their essential contribution to this project. A special thanks goes out to our colleague Nils Deslé for his exceptional work on developing the virtual supermarket.

REFERENCES

1. Abowd, G. and Mynatt, E. Charting past, present and future research in ubiquitous computing. *ACM Trans. Comput.-Hum. Interact. 7*, 1 (March 2000), 29-58
2. Agarwal, R. and Prasad, J. The role of innovation characteristics and perceived voluntariness in the acceptance of information technologies. *Decision Sciences 28,* 3 (1997), 557–582.
3. Ajzen, I. The Theory of Planned Behavior. *Organizational Behavior and Human Decision Processes 50,* 2 (1991), 179-211.
4. Banerjee, S. and Dholakia, R.R. Mobile advertising: does location-based advertising work? *International Journal of Mobile Marketing 3,* 2 (2008), 68-74.
5. Barton J.J. and Vijayaraghavan, V. Ubiwise: A ubiquitous wireless infrastructure simulation environment. Tech. report HP Labs, 2002.
6. Bauer, H.H., Reichardt, T., Barnes, S.J. and Neumann M.M. Driving Consumer Acceptance of Mobile Marketing: A Theoretical Framework and Empirical Study. *Journal of Electronic Commerce Research 6*, 3 (2005), 181-192.
7. Barkhuus, L. and Dey, A. Location-based services for mobile telephony: A study of users' privacy concerns. *Proc. INTERACT '03,* IOS Press (2003).
8. Barwise, P. and Strong, C. Permission-Based Mobile Advertising. *Journal of Interactive Marketing 16*, 1 (2002), 14-24.
9. Davis, F.D. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly 13* (1989), 319-340.
10. Dhar, S. and Varshney, U. Challenges and business models for Mobile Location Based Services and Advertising. *Commun. ACM 54, 5 (May 2011)*, 121-129.
11. Dinh, H.Q., Walker, N., Song, C., Kobayashi, A., and Hodges, L.F. Evaluating the Importance of Multi-sensory Input on Memory and the Sense of Presence in Virtual Environments. *Proc. VR'99,* IEEE (1999), 222.
12. Duh, H.B.-L., Tan, G.C.B. and Chen, V.H.-h. Usability Evaluation for Mobile Device: A Comparison of Laboratory and Field Tests. *Proc. Mobile HCI '06,* ACM Press (2006), 181-186.
13. Edwards, M., Li, H. and Lee, J-H. Forced exposure and psychological reactance: Antecedents and consequences of the perceived intrusiveness of pop-up ads. *Journal of Advertising 31* (2002), 83-95.
14. Fishbein, M. and Ajzen, I. *Belief, Attitude, intention and behavior: An introduction to theory and research.* Addison-Wesley, 1975.
15. Ha, L. Observations: advertising clutter in consumer magazines: dimensions and effects. *Journal of Advertising Research 36*, (1996), 76-83.
16. Kaikkonen, A., Kallio, T., Kekäläinen, A., Kankainen, A. and Cankar, M. Usability testing of mobile applic*ations: A comp*arison between laboratory and field testing. *Journal of Usability Studies 1*, 1 (2005), 4-16.
17. Karapanos, E. *Quantifying Diversity in User Experience.* PhD Thesis, Eindhoven University of Technology, 2010.
18. Khan, V.J., Nuijten, K. and Deslé, N. Pervasive Evaluation Application within Virtual Environments. *Proc. PECCS '11*, 261-264.
19. Kjeldskov J. and Graham C. A Review of MobileHCI Research Methods. *Proc. MobileHCI '03,* LNCS Springer-Verlag (2003), 317-335.

20. Kjeldskov, J. and Stage, J. New techniques for usability evaluation of mobile systems. *International Journal of Human-Computer Studies 60*, 5-6 (2004), 599-620.
21. Kjeldskov J., Skov M.B., Als B.S. and Høegh R.T. Is it Worth the Hassle? Exploring the Added Value of Evaluating the Usability of Context-Aware Mobile Systems in the Field. *Proc. MobileHCI '04,* Springer-Verlag (2004), 61-73.
22. Klompmaker, F., Stern, C., Reimann, C. and Santelmann, H. A Mock-up System for the Early Testing of Location Based Services. *In: Mobile Interaction with the Real World 2007/5th Workshop on HCI in Mobile Guides. Singapore: ACM*
23. Kowatsch, T. and Maass, W. In-store consumer behavior: How mobile recommendation agents influence usage intentions, product purchases, and store preferences. *Computers in Human Behavior 26*, (2010) 697-704.
24. Leichtenstern, K., André, E. and Rehm, M. Using the Hybrid Simulation for Early User Evaluations of Pervasive Interactions. *Proc. of NordiCHI '10*, ACM Press (2010), 315-324.
25. Li, H., Edwards, S.M. and Lee, J. Measuring the Intrusiveness of Advertisements: Scale Development and Validation. *Journal of Advertising 31* (2002), 37-47.
26. Lin, J.J.W., Duh, H.B.L., Abi-Rached, H., Parker, D.E., and Furness, T.A. Effects of Field of View on Presence, Enjoyment, Memory, and Simulator Sickness in a Virtual Environment. *Proc. VR '02,* IEEE (2002), 164-171.
27. Lumsden, J., Kondratova, I. and Durling, S. Investigating Microphone Efficacy for Facilitation of Mobile Speech-Based Data Entry. *Proc.BCS-HCI '07,* BCS Swinton (2007), 89-97.
28. McCoy, S., Everard, A., Polak, P. and Galletta, D.F. An experimental study of Antecedants and consequences of online ad-intrusiveness. *Journal of Human-Computer Interaction 24*, 7 (2008), 672-699.
29. Muehling, D.D., Stoltman, J.J. and Grossbart, S. The impact of comparative advertising on levels of message involvement. *Journal of Advertising 19*, 4, 41-50.
30. Nielsen, C.M., Overgaard, M., Pedersen, M.B., Stage, J. and Stenild, S. It's worth the hassle!: the added value of evaluating the usability of mobile systems in the field. *Proc. NordiCHI '06*, ACM Press (2006), 272-280.
31. Nurmi, P., Salovaara, A., Bhattacharya, S., Pulkkinen, T. and Kahl, G. Influence of Landmark-Based Navigation Instructions on User Attention in Indoor Smart Spaces. *Proc. IUI'11,* ACM Press (2011), 33-42.
32. O'Neill, E., Klepal, M., Lewis, D., O'Donnell, T., O'Sullivan, D. and Pesch, D. A testbed for evaluating human interaction with ubiquitous computing environments. *Proc. Tridentcom '05*, IEEE (2005), 60-69.
33. Pramanik, A. The use of a virtual environment as a method of wayfinding research in architecture. Master's Thesis, Texas Tech University, 2006.
34. Reynolds, V., Cahill, V. and Senart, A. Requirements for an ubiquitous computing simulation and emulation environment. *Proc. InterSense '06,* ACM Press (2006).
35. Roto, V. Oulasvirta, A. Haikarainen, T. Lehmuskallio, H. and Nyyssönen, T. Examining mobile phone use in the wild with quasi-experimentation. *HIIT Technical Report 2004-1*, August 13 2004.
36. Schellenbach, M., Krüger, A., Lövdén, M. and Lindenberger, U. A laboratory evaluation framework for pedestrian navigation devices*.* *Proc. Mobility '07,* ACM Press (2007), 495-502.
37. Slater,M, Steed,A., McCarthy, J. and Marinelli, F. The Influence of Body Movement on Presence in Virtual Environments. *The Journal of the Human Factors and Ergonomics Society 40,* 3 (1998), 469-477.
38. Singh, P., Ha, H.N., Kuang, Z., Olivier, P., Kray, C., Blythe, P. and James, P. Immersive video as a rapid prototyping and evaluation tool for mobile and ambient applications. *Proc. MobileHCI '06,* ACM Press (2006), 264-264.
39. Snowdon, C. and Kray, C. Exploring the use of landmarks for mobile navigation support in natural environments. *Proc. MobileHCI '09*, ACM Press (2009), Article 13, 10 pages.
40. Unni, R. and Harmon, R. Perceived effectiveness of push vs. pull mobile location-based advertising. *Journal of Interactive Advertising 7*, 2 (2007), 28-40.
41. Verbeke, W. and Vackier, I. Profile and effects of consumer involvement in fresh meat. *Meat Science, 67,* 1 (2004), 159-168.
42. Wehmeyer, K. Mobile Ad Intrusiveness – The Effects of Message Type and Situation. *eMergence: Merging and Emerging Technologies, Processes, and Institutions 6* (2007), 1-18.
43. Xu, H., Oh, L-B. And Teo, H-H. Perceived effectiveness of text vs. multimedia Location-Based Advertising messaging. *International Journal of Mobile Communications 17*, 2 (2009), 154-177.