ABSTRACT
There are many methods for the evaluation of ubiquitous computing (ubicomp) applications. These evaluations usually require an autonomous system, or use scenarios or storyboards instead.

We suggest Marionette Prototyping as a technique for ubicomp applications that can be used early in the design process. It allows participants to use a conceptual ubicomp application in a real-world context, followed by an evaluation that covers the participants understanding, experience and attitude with regard to the application. Marionette Prototyping is inspired by puppetry, especially the styles where the manipulator is in plain view. It combines principles from cardboard prototyping and Wizard-of-Oz and uses off-the-shelf tools and technology.

We have used Marionette Prototyping in the evaluation of a ubicomp application. This evaluation has shown that Marionette Prototyping provides input on the understanding, experience and attitude of the user with regard to the ubicomp application in question. From this first step, we can continue to develop this method as a technique for early, in context evaluation of ubicomp applications. In this, Marionette Prototyping can overcome some of the issues with current evaluation methods for ubicomp applications.
the miniaturisation of computational and communicational components, such interfaces also become more and more mobile. They are less linked to one particular context, but applicable in a variety of contexts.

We believe that only if you can actually use such an application and interface, and interact with it in these different real-world contexts, you can understand the application and evaluate all its aspects. However, developing a system that allows this takes time and money. Such an investment can be problematic given the uncertainty of whether the application will be understood and appreciated in the first place. On the other hand, evaluation using scenarios and storyboards includes a risk that participants misunderstand the application or that certain aspects of the application are not taken up in the evaluation.

The research described in this paper focuses on support for in context evaluation of ubicomp applications during early stages of their design process. More specifically, we use the ubicomp application in a way that it augments activity that is familiar to the user. We embed the prototype in this way partly to support the suggestion that it works autonomously, even though the computational technology is not fully functional.

BACKGROUND
Within the domain of ubicomp applications, various studies have been aimed at exploring possibilities for such applications (e.g. Iacucci et al., 2000, Iacucci & Kuuti, 2002; Consolvo & Walker, 2003). These studies are usually aimed at mapping situations and interactions that users encounter during the day, which in turn can provide inspiration for possible ubicomp applications. Furthermore, numerous (software) tools exist to develop prototypes for ubicomp applications (see e.g. Tang et al. 2010).

When it comes to understanding user experience and value of ubicomp applications, there is need for evaluation techniques: “Ubicomp researchers are beginning to explore evaluation techniques including field studies that drive invention, early-stage requirements gathering, and prototype iteration” (Carter and Mankoff, 2005). This has led to a discussion of a variety of methods for the evaluation of ubicomp application (e.g. Abowd et al., 2005; Consolvo & Walker, 2003; Carter & Mankoff, 2005; Dow et al., 2005; Reilly et al., 2005; Tang et al., 2010). Compared to the explorative studies into possible ubicomp applications, evaluative studies have the advantage that they do not need to cater for all possible applications, situations and interactions. Instead, they can focus on the evaluation of one specific application, with a limited set of situations of use and interactions.

On the other hand, there are also challenges in doing early stage evaluation of ubicomp applications (Carter et al., 2008), partly emanating from how the prototype, the prototyping situation, and the prototyping process are set up. The first challenge stems from the fact that such an evaluation often requires (parts of) the technology to ‘work’, in order to be able to properly evaluate the application. Besides, context seems to be important for ubicomp applications and since their mobility allows these applications to be used in different contexts, it is worth “to determine the contextual attributes for each application specifically” (Oulasvirta et al., 2003). Previous research on evaluation of mobile phone applications has shown that evaluation in context leads to identification of other issues than laboratory studies produce (e.g. Nielsen et al., 2006; de Sá & Carriço, 2008; Duh et al., 2006). This suggests that evaluating ubicomp applications in the real world is valuable and important, although this only applies if the evaluation takes place in a realistic context. If a laboratory test is replicated in the field, it is less realistic (Nielsen et al., 2006) and will thus not provide the added value of evaluation in context.

The in-context evaluation of prototypes that embody some functionality has proven to be a challenge for ubicomp applications. Carter and Mankoff (2005) showed that performing such an evaluation by deploying a more-or-less autonomous system over a longer time allows gathering information about a variety of usage situations. Yet, this comes at the cost of challenges with stability of the prototype and (software) updates, increasing the efforts of developing the prototype (Carter and Mankoff, 2005).

By applying the Wizard-of-Oz methodology in a way that the wizard has different roles in different stages of the design process, Dow et al. (2005) presented a possible solution to overcome this problem. However, their relatively strict application of the Wizard-of-Oz principles means that the wizard cannot see what participants do. Since the application in question was an audioguide, the location of participants was most important, which could be gathered remotely, using the participant’s GPS locations.

In ubicomp applications where users interact with ubicomp objects, on the other hand, it might be interesting—if not necessary—to see what participants do. In the evaluation of Rendezvous, an application that helps two persons to find a meetup location in a crowded area, Reilly et al. (2005) used a setup where a researcher would follow a participant during the evaluation. This allowed the researcher to observe what a participant was doing (and remotely manipulate the information on the participant’s device). They used Bluetooth technology for the remote manipulation. This required the researcher/wizard to be relatively close to the participant in order to make these remote manipulations, because the range for a Bluetooth signal is limited. A downside of this approach is that it the manipulations can become obtrusive.
In this paper, we add to the discussion on evaluation methods for ubicomp application methods by presenting Marionette Prototyping. This prototyping method can be used to evaluate ubicomp technology and applications in context, in early stages of the design process. Marionette Prototyping is developed to allow gathering feedback regarding three aspects of these applications: understanding, experience and attitude, each of which we will explain in more detail below.

Firstly, the application should be understandable, so that people know how to use it. They need to understand what the possibilities and limitations of the applications are and how they can interact with it. This aspect of an evaluation focuses on the rational elements of the interaction with an application.

Secondly, users will have a certain experience during interactions with the application. This aspect of an evaluation focuses on the emotional element of the interaction that the user has with the application.

Finally, one wants to capture the attitude that people have towards the application and their intention to use it. This aspect of an evaluation takes into account how the current design of the application influences their intention for use. In order to allow in context testing of ubicomp applications where users interact with ubicomp objects, we have included a technical setup in Marionette Prototyping that overcomes some of the issues described above.

MARIONETTE PROTOTYPING
Marionette Prototyping is meant to be used when prototyping is done for evaluation in use scenarios where the user is mobile and moves through contexts, in multiple device conceptual development, or where it is not feasible or desirable to instrument all devices with computational technology today. Then, the communication between ubicomp object and user-device—which would happen as a result of the user-object interaction—would be faked by a researcher in the Marionette Prototyping setup. That is, the interactive experience is supported by a researcher manipulating the user-device in correlation to the user’s interaction with the ubicomp object (Figure 2). This communication would take place using a wifi-connection, allowing the researcher/ manipulator to be far from the participant, helping to make the manipulations non-obtrusive for the participant.

THE TECHNICAL SETUP
Imagine a multi-device interaction context, where the future solution would involve direct wireless contact between a user-device and one or more ubicomp objects (Figure 1). The interaction might be active in the sense that the user scans an RFID tag in the ubicomp object with the user device, or passive, in the sense that an action of the user with the ubicomp object is communicated to the user-device (e.g. a swing with a golf club). Marionette Prototyping can be used if the object would be difficult to instrument with necessary computational technology today. Then, the communication between ubicomp object and user-device—which would happen as a result of the user-object interaction—would be faked by a researcher in the Marionette Prototyping setup. That is, the interactive experience is supported by a researcher manipulating the user-device in correlation to the user’s interaction with the ubicomp object (Figure 2). This communication would take place using a wifi-connection, allowing the researcher/ manipulator to be far from the participant, helping to make the manipulations non-obtrusive for the participant.

 SOME BACKGROUND ON THE MARIONETTE
A marionette is a puppet that is used in, for instance, marionette theatre. It is controlled by strings or wires by a marionettist, which sometimes is called a puppeteer or a manipulator. Operas and other dramas have been commonly performed with marionettes, often called marionette theatre. In Europe a common version of marionette theatre is where the puppeteer is hidden from the view of the audience. In contemporary drama, marionette styles where the puppeteer is seen by the audience have been used. In, for example, the musical "War horse", the life size horses are marionettes, with their puppeteer on stage, where they act together with ordinary actors. These ordinary actors have two roles: their own role (as e.g. soldier, farmer) and a role of puppeteer, manipulating (parts of) the horse.

In other marionette theatre styles, such as the Japanese Bunraku, or the Burmese Yoke thé, the puppeteers are visible to the audience. In Bunraku, for example, most of the marionettes have three puppeteers controlling them, often dressed in black dresses but in full view.

Figure 1. Imagined wireless interaction with a ubicomp object. The object sends information directly to the user’s device.

Figure 2. Marionette setup of the ubicomp interaction. For the user, the information seems to be coming from the object, whereas in reality, the researcher sends this information to the user’s device.

Thus, in the Marionette Prototyping method we create a situation that is similar to marionette theatre, where manipulator and spectator are in the same room and
visible for each other. In Marionette Prototyping, we will not hide manipulators physically, but we will conceal them by giving this manipulator another another role in the eye of the participants test. In other words, the manipulator is hidden by role. Furthermore, Marionette Prototyping includes some device being manipulated (the user device), and a set where the drama is taking place (the ubicomp object(s), the use-contexts and other objects). In both cases, there is a performative situation, which is fictional. In marionette theatre, the story is fictional, whereas in Marionette Prototyping, the interaction is fictional. This interaction develops as the researcher manipulates the user device, like the the story in marionette theatre develops as the puppeteer manipulates the puppets.

Marionette Prototyping is different from some forms of marionette theatre in the sense that the user is not just a spectator, but an actor that can influence the performance. It is thus similar to Theatre of the Oppressed (Boal, 1979) and the related concept of hidden theatre. In hidden theatre, the performance is staged, but the actors aim to hide this fact, which leads spectators to believe that the performance is unstaged.

In many imagined ubicomp applications it is common that the user interacts through the user-device with many different ubicomp objects and while moving in and through contexts. The specific user actions and interactive behaviour is adapted to the conditions of a situation, based on the context (e.g. a crowded space) as well as conditions at a micro or local level (e.g. changing body-position as a consequence of glare). To be able to correctly manipulate the user-device, this would require the manipulator to be present where the interaction happens. This is especially true if there are several different ways to interact with the ubicomp object, such as swiping, scanning, bumping.

THE PROTOTYPING PROCESS
The technical solutions to the issues of instrumenting objects and manipulating user devices discussed above allow us to do in-context prototyping of ubicomp applications. This allows to take into account context factors and the influence of the presence of other people into account in the evaluation. To evaluate the prototype we used a process that consists of three stages.

Firstly, a service walkthrough (see e.g. Blomkvist, Åberg and Holmlid, 2012) in the actual context, where twelve participants (5 male and 7 female, aged 31-49) could experience what it would be like to interact with the ubicomp object and with a mobile device in the real context for the application. The effects of these interactions were simulated using the Marionette Prototyping method. The aim with this was twofold: to understand the experience of using the application, and to uncover issues of using the Marionette Prototyping method. Participants were told that the researcher that would follow them during the walkthrough did so to take notes on the use of the application. In reality, this researcher also had the role of manipulator. This second role was not revealed to the participant. In other words, the manipulator was hidden by role to the participants.

The researcher/ manipulator carried a mobile device and a notebook (Figure 3) as support in his two roles. During the walkthrough, participants were able to interact with the ubicomp application under evaluation, in the real context of use. This application would be used in a public place and would involve a similar interaction with each of the ubicomp objects. Yet, these objects will be situated in slightly different circumstances (e.g. they could be placed on a high or low position and could be attached to different surfaces) and in different places within this public place. This could have an influence on the participant’s understanding, experience and attitude with regard to the ubicomp application. The ubicomp objects that were deployed for the walkthrough did not have the required computational functionality yet (this was mimicked by the manipulator instead). The graphical appearance of the ubicomp objects resembled the appearance of their non-ubicomp counterparts in the test location. Besides, the ubicomp objects were spread across the test location and placed alongside their non-ubicomp cousins.

Secondly, after the walkthrough, a questionnaire was distributed to evaluate the user’s experience, understanding and attitude with regard to the ubicomp application. This questionnaire focused on the participant’s perception of the application in terms of added values, expected problems, and conditions that would influence their adoption of the application if it became available.

Finally, an interview was conducted with the participant, where positive and negative experiences for each individual interaction with the ubicomp application during the service walkthrough was evaluated. This was done using a set of storyboard cards depicting the different steps and interactions the user had been doing. This evaluation consists of positive and negative aspects to the experience.

Figure 3. The setup for the researcher following along on the walkthrough. The tablet on the left, for manipulating the user-device in the manipulator role, and the notepad on the right, for taking notes in the researcher role.
DISCUSSION
Marionette Prototyping is inspired by both Wizard-of-Oz and cardboard computing. When using this prototyping method for the evaluation of the user’s understanding, experience and attitude with regard to a ubicomp application, we have observed the following things.

FUNCTIONAL ENGAGEMENT HIDES THE WIZARD
Marionette Prototyping is similar to Wizard-of-Oz in the sense that a part of the computational functionality of the ubicomp application is not developed to the level where it functions autonomously. Instead, a manipulator controls the feedback related to interactions that a participant has with the ubicomp object. In contrast to traditional Wizard-of-Oz, this manipulator is present in the same room as the participant. However, similar to the Wizard-of-Oz method, users are not aware of the presence of the wizard, albeit for a different reason. In Wizard-of-Oz, the wizard is physically hidden. In Marionette Prototyping, the manipulator is hidden by role because the manipulator is introduced to the participant as observer. Additionally, the manipulator is not in the line of sight of a participant. Furthermore, the wizard is concealed by the fact that the user is engaged elsewhere, drawing his or her attention away from the manipulator. This is comparable to some forms of marionette theatre (like Burmese Yoke thé or Japanese Bunraku), where the puppeteer is also present and visible to the audience, but occluded through emotional engagement of the audience in the story that is enacted. A skilful puppeteer can create an experience that is so immersive and engaging that the audience forgets that they look at a performance where the actors are in fact dolls that are manipulated by a puppeteer.

In Marionette Prototyping, we do not have the elements of empathy with the character(s) nor storytelling techniques that help create this occlusion through emotional engagement of the audience. Instead, we can only achieve occlusion through the position of the manipulator (outside the direct line of sight of the participant), hiding the manipulator behind the role of observer to the participants and through the participant’s functional engagement in the task at hand.

HIGH FIDELITY LOOK AND FEEL
The Marionette Prototyping method also shows similarities with cardboard mock-ups or paper prototyping in the sense that the interface of the ubicomp objects is not yet in the final stage of the design. They are not yet functional so only the appearance of the interface is representative for the design that is evaluated. In contrast to cardboard mock-ups and paper prototyping the appearance of the ubicomp objects were of the same standard as their non-ubicomp cousins. In addition, the feedback that users will receive, appears to be coming from an underlying, autonomous system, unlike in paper or cardboard prototyping, where feedback is clearly managed by a human computer (Snyder, 2003). Besides, the fact that the evaluation of the ubicomp application is embedded in a realistic activity and tested in the real context helps to make it more believable.

THE WIZARD WAS NOT SPOTTED
Despite the differences of Marionette Prototyping in relation to both Wizard-of-Oz and cardboard prototyping, we have not experienced that participants were influenced by these differences. Only one of the participants noticed that the application was not functioning autonomously. Furthermore, we have successfully collected the information regarding understanding, experience and attitude that we hoped to gather using the technique. This suggests that Marionette Prototyping can be used as prototyping tool for the evaluation of ubicomp applications during early stages of the design process. It fills a gap between evaluation of ubicomp applications using storyboard and scenarios on the one hand and the deployment of a more or less autonomous system on the other hand. As such, it does allow users to experience and interact with the application, without requiring the resources to develop an autonomously functioning prototype. An additional advantage of Marionette Prototyping is that it makes it possible to use tools and technologies that are already available today, such as smartphones, tablets and wireless internet. This also lowers the time and costs for developing prototypes of ubicomp applications.

CHALLENGES AND POTENTIAL SHORTCOMINGS
When designing the first Marionette Prototyping test, we noticed that the possibility to use existing tools and technology is not a guarantee for a smooth and easy design experience. For instance, even though wifi is commonly available these days, other aspects—such as security—may form an obstacle for using this technology. Furthermore, we focused first and foremost on making sure that it was technically possible to manipulate the app on the participant’s smartphone using the app on the tablet. As a result, we could not put much time and thought in the design of the user interface of the smartphone app, which therefore ended up being relatively basic and stripped down. Even though we have no reason to suspect that this influenced the participant’s evaluation of the ubicomp application as a whole, it would be interesting to explore the effects of different levels of detail in the design of the smartphone app on the participant’s understanding, experience and attitude with regard to the ubicomp application.
FURTHER DEVELOPMENT
An aspect that we believe is interesting for future development for Marionette Prototyping concerns data collection. Currently, the technique relies completely on the researcher for both manipulation of the technology and data collection. Being engaged in these two tasks at the same time could for instance lead to an incomplete experience (the researcher could fail to observe certain actions that should have triggered feedback) or omissions in the data collections (the researcher does not notice reactions from the participant because (s)he is engaged in the manipulator task). On the other hand, it has been argued that using video to observe and record the user's actions also influenced their experience (e.g. Isomursu et al., 2004).

Furthermore, the software tools that are used in Marionette do not allow to make adjustments on the fly, such as Memento (Carter et al., 2007) allows for prototypes or applications that only run on a smartphone. Being able to make small adjustments between tests would make it possible to correct simple issues that (re)occur during each test (as Snyder (2003) suggests for paper prototyping).

MARIONETTE PARTICIPATORY INNOVATION
Marionette Prototyping, in the form it has been presented here, is limited in its participatory stance. However, as we succeeded with a visible manipulator, we may now start to experiment with the actions of the manipulator, and the involvement of the prototyping participants in the development of prototypes.

For example, with the above mentioned possibilities to make adjustments, it would be possible to provide additional opportunities for participatory design of ubicomp applications. Instead of simply evaluating applications, it would become possible to iterate and improve the design together with the participants, to develop “prototype-driven specifications” (Bogers and Horst, 2014) for the ubicomp application. This would make Marionette Prototyping not just a method for evaluating ubicomp applications, but also for co-creating new or alternative versions of such applications. One interesting variant would be to do the kind of stop-motion iterations that are suggested by Blomkvist & Arvola (2014). Another would be to facilitate multi-stakeholder collaborative prototyping (Terweisch and Loch, 2004; Bogers and Horst, 2014).

In addition, there is a space for exploration regarding how actively the manipulator interacts with the participants, and the content of these interactions. We are interested to learn what actions of the manipulator do not create (unwanted) breakdowns in the participants engagement and which amount of participation in changing and building the application will still keep the participant within a suspension of disbelief. Investigating this would rely on the assumption that the innovation process builds on contextual engagement, as well as dramaturgical techniques.

CONCLUSION
In this paper we have presented Marionette Prototyping as a method for prototyping ubicomp applications in context, during early stages of a design process. This method is aimed at gathering feedback on the user’s understanding, experience and attitude with regard to the application. Data on these aspects is gathered through a combination of a service walkthrough followed by a questionnaire and an interview.

Marionette Prototyping overcomes some of the technical issues regarding evaluating ubicomp applications early in the development process by taking inspiration from Wizard-of-Oz methodology and cardboard prototypes. The Marionette Prototyping method is different because the wizard is not physically hidden from view. Instead, the wizard is hidden by role, because the manipulator is introduced only as observer to the participants. In addition, it uses functional engagement to direct attention towards the goal of using the application. The appearance of the prototypes used in Marionette Prototyping also has a higher fidelity than the appearance of cardboard prototypes.

A first application of the methodology has provided the desired insights regarding the user’s understanding of, experience with and attitude towards the ubicomp application. It has also shown that the puppeteering setup and the (hidden) double role of the wizard helped to convince users that they were interacting with an autonomous application.

We intend to continue exploring the possibilities and limitations of Marionette Prototyping and further improve the shortcomings that were discovered during the initial use of the method.

ACKNOWLEDGEMENTS
We wish to acknowledge the invaluable support from the participants in the Marionette tests performed.

We wish to acknowledge Malin Niklasson and Johan Blomkvist that helped run some of the tests. We are grateful for the support from Magnus Berggren and his research team, as well as the technical teams at Ericsson, Acreo and Sogeti. This research was supported by a grant from VINNOVA, dnr 2012-01172 “Printed Electronics Meets Mobile”.

REFERENCES


