# Pausing or not?: Examining the Service Walkthrough Technique

Johan Blomkvist and Mattias Arvola

**Linköping University Post Print**

Tweet

N.B.: When citing this work, cite the original article.

# Pausing or not? Examining the Service Walkthrough Technique

Johan Blomkvist
Linköping University
SE-581 83 LINKÖPING
johan.blomkvist@liu.se

Mattias Arvola
Linköping University
SE-581 83 LINKÖPING
mattias.arvola@liu.se

**The scope of service design calls for holistic design techniques that represent multiple service moments. One such technique is the service walkthrough that can be used to prototype and formatively evaluate services. A service walkthrough is an enactment of several consecutive service moments. This paper informs decisions about how to set up service walkthroughs by looking at two kinds of walkthroughs in a case study: with pauses for discussion and feedback after each service moment, and without pauses where the entire service journey is walked through before comments and feedback are collected. The case study did not show any differences in the content of the feedback, but more feedback was given in the walkthroughs with pauses. The feedback in the paused walkthroughs was also more detailed.**

*Case Study, Prototyping, Evaluation, Service Walkthrough, Content Analysis*

## 1. INTRODUCTION

Imagine being a tourist in the area around your hometown, finding places and stories you did not know existed. You explore them by walking around trails set up by others for you. Media appear on your smartphone at various stations along the way.

The case described in this paper explores ideas for computer-augmented trails, but the research focus is not on the design ideas, but rather on the methods that were used to evaluate them.

A formative evaluation technique called service walkthroughs has been outlined in earlier research on service design (Blomkvist, Åberg, & Holmlid, 2012). It has been devised to account for holistic aspects of services, a key issue when prototyping in service design (Blomkvist & Holmlid, 2010). Service design requires an understanding of how service moments work together (Koivisto, 2009). Service walkthrough can be seen as a conglomerate of pluralistic walkthrough (Bias, 1994), experience prototyping (Buchenau & Fulton Suri, 2000), and bodystorming (Burns, Dishman, Verplank, & Lassiter, 1994; Oulasvirta, Kurvinen, & Kankainen, 2003).

A service walkthrough is set up by first identifying a number of important sequences or significant moments in the service journey. Props are then gathered, or created, to facilitate an enactment of the service. The roles involved in the service are also identified and people are assigned to play those roles. They do not have to be played by representatives of the roles in the actual service. For instance, a designer can play a customer, and a customer can play a salesperson. The walkthrough then starts with a customer action. For example, a number of available offers can be shown to the person playing the customer and the enactment of the service can then start by whatever the customer chooses. A service can accordingly be walked through, step by step.

A walkthrough can be paused after each service moment for feedback and discussions, but discussions and feedback can also wait to the end. The purpose of this case study is to develop an understanding for the design feedback received from service walkthroughs with and without pauses. This will allow for designers to make a more informed choice on how to set up a walkthrough.

It has been suggested both in pluralistic walkthrough and in experience prototyping that it is a good idea to pause after each part of the experience to collaboratively evaluate one's impressions (Bias, 1994; Buchenau & Fulton Suri, 2000). Feedback collected immediately after has the potential of being more detailed.

There is however a risk that pausing to consider and provide feedback might interfere with the flow of events. A walkthrough without pauses have therefore the potential to generate feedback on a holistic level, concerning the long-term and overall user experience (UX) of how service moments work together. Long-term UX is the experience related to the outcomes of the activity that takes

place guided by the motives that drives it (Luojus, 2010). This can be contrasted to momentary UX, which is the experience related to the outcomes of actions and operations performed by people to reach their goals under the conditions of the situation (Luojus, 2010). Feedback in a walkthrough can be related to the long-term or momentary UX, where long-term UX comments concern whole experiences on the activity level, while comments about the momentary UX will be related to specific steps or actions and their outcomes in a larger context of the activity.

The level of detail in the design that a feedback comment focuses on is in this paper conceptualised using a model for interaction design levels suggested by Arvola and Artman (2007). The model divides the design space into five levels: (1) concept, (2) function and content, (3) structure, (4) interaction, and (5) presentation. The concept is the idea or purpose of the design. The function level concern what actions can be taken and the content materials used for the actions. On these two levels the issue is what the role of the design is and what it should allow people to do, how it supports or provides value for people. The concept and the function and content answer the question of what to design and why. The structure is "the arrangement and organization of functions and content" (Arvola & Artman, 2007). It refers to how information or objects are structured in time and space. Interaction means the ways that functions and content can be manipulated and interacted with. Presentation finally, concerns the look and feel of the design. The final three are thus related to more detailed information about how to design something rather than what to design or why.

Based on the earlier research, some differences in types of feedback between service walkthroughs with pauses and walkthroughs without pauses can be assumed to occur. We had the opportunity to study such differences in our case study where both kinds of walkthroughs had been conducted. We posed the following research questions for the case study:

(i)   What types of feedback are there in the service walkthroughs with pauses compared to the service walkthrough without pauses?

(ii)  Is there more detailed feedback in the walkthroughs with pauses?

(iii) Is there more feedback about long-term UX in the service walkthroughs without pauses?

## 2. METHOD

Our case study provided us with well-documented service walkthroughs of both kinds. The case itself is described in more detail in the next section. The feedback from the service walkthroughs was audio recorded and then transcribed.

To answer the first research question, the transcribed feedback was analysed using bottom-up qualitative content analysis (Graneheim & Lundman, 2004). The number of comments that were categorised as belonging to a certain type was also counted. The unit of analysis was feedback provided during the service walkthroughs. The meaning units were the documented comments about the service walkthrough. The meaning units were condensed into smaller units, which were turned into codes that signified different expressed sentiments. The analysis then proceeded by comparing the codes to find categories. That is, groups of content with similar meaning. These were further abstracted by identifying themes consisting of groups of categories, depending on the amount of data analysed. As a validation strategy, multiple independent researchers analysed the data before discussing it further. The number of comments belonging to different types was counted after the establishment of categories.

The second and third research question was addressed using quantitative content analysis where the level of details framework (Arvola & Artman, 2007) and long-term vs. momentary UX framework (Luojus, 2010) were used top-down as coding scheme.

## 3. CASE DESCRIPTION

The service walkthroughs analysed in this paper were conducted in a project aimed at developing a smartphone application for guiding people in the cultural and historical landscape near a medium sized Swedish city. The smartphone application was supposed to inspire primarily local people to visit a nearby nature reserve and also enable them to learn more about the cultural heritage of the area. Visits to the reserve were divided into stations where people could access media using their smartphone. The stations formed service moments that were formatively evaluated using service walkthroughs.

Four walkthroughs were conducted with two participants in each. Two participants were female and six were male. They were university students between 22 and 27 years old. Two of the four walkthroughs were paused for feedback after each station, while the other two were made without any pauses. Feedback from the users was then instead

elicited after the entire walkthrough. Two different semi-structured interview guides were used for the different kinds of walkthroughs. The walkthroughs including feedback sessions took between 40 and 60 minutes. The feedback was audio recorded and subsequently transcribed.

The smartphone application allowed users to either plan a route through the stations, or to choose a predesigned route. The first step for the participants was to choose whether to plan or not, based on what they imagined they would do in a real situation. They then walked through four stations. The prototype was a low fidelity prototype, which in this case also meant that the stations were much closer than they would be in reality. Audio information that was supposed to be included in the application was read to the participants. Video was shown on a separate smartphone.

## 4. RESULTS

Our first research question was what the types of feedback were in the service walkthroughs with pauses compared to the service walkthroughs without pauses. To answer this question the transcribed feedback was analysed using qualitative content analysis.

### 4.1 Types of Comments

Three types of comments were identified: suggestions (Table 1), negative comments (Table 2), and positive comments (Table 3). The three types were then analysed separately. All in all, 122 comments were analysed. There were 56 suggestions, 25 negative comments, and 41 positive comments. The walkthrough with pauses resulted in 73 comments in total, and the walkthrough without pauses resulted in 49 comments. There was accordingly more feedback given in the walkthroughs with pauses than the ones without pauses, but there were no differences in the proportions of suggestions, negative comments and positive comments. The question is then if there were any qualitative differences in the content of the feedback. The remainder of this section focuses on that issue.

*Table 1: Distribution of comments coded as suggestions*

| Suggestions | Without pauses | With pauses |
|---|---|---|
| Function | 6 | 12 |
| Content | 12* | 16* |
|   images | 1 | 2 |
|   information | 8 | 8 |
|   sound | 2 | 0 |
|   text | 1 | 6 |
| Interaction | 2 | 5 |
| Appearance | 2 | 1 |

*Table 2: Distribution of comments coded as negative*

| Negative comments | Without pauses | With pauses |
|---|---|---|
| Difficulties | 3 | 8 |
| Information | 5 | 3 |
| Miscellaneous | 1 | 5 |

*Table 3: Distribution of comments coded as positive*

| Positive comments | Without pauses | With pauses |
|---|---|---|
| Content | 5 | 5 |
| Control | 5 | 10 |
| Availability | 2 | 0 |
| Clarity | 1 | 1 |
| Experience | 5 | 6 |

**Without pauses:** Suggestions from the walkthroughs with pauses concerned issues like planning your own route beforehand, how to choose route, how to highlight things on the map, the addition of more nature related information, or more information in general about things like walked distance, as well as more accessible and clear information. The interaction could be improved by adding a return button and frames around clickable images in the interface.

Negative comments were made about content concerning the difficulties of getting started, using the zoom and knowing what to press. The information lacked clarity, and was confusing. A number of positive comments in relation to movies, audio and text, and how they could be accessed were made. As well as in regard to planning, interactivity, freedom to choose, and feedback. The experience was described as interesting, exciting, interactive, and modern.

**With pauses:** The paused walkthroughs generated suggestions such as adding the possibility to go back to previous stations and hear information about them, instructions about where to point the camera of the smartphone, and showing stations that have already been visited. Also, adding information about choices, symbols, stations, and locations would be good. On the negative side, unclear details like how to interpret information and how to choose stations and routes were mentioned. The application was also experienced as inconsistent and confusing at times, and difficulties included problems with hearing recorded messages and how to navigate the map.

Good functions, good amount of, and interesting information presented in various ways, as well as planning and choices were considered advantages. Clarity of choices, and variation were seen as positive for the experience.

Taken together we cannot see any substantial qualitative differences in the feedback.

## 4.2 Level of Detail

The second research question was if there was more detailed feedback in the walkthrough with pauses.

In the first two attempts at analysis, the two researchers independently categorised comments, but could not reach a satisfying level of agreement, which finally prompted a joint coding session. The two authors coded using a guide based on the model for interaction design levels. Differences in coding were discussed to reach a consensus result. The discrepancies in the individual coding were between content and presentation, between function and interaction, and how to view structure. Some additional problems were related to judgements where detailed comments propagated consequences on a more conceptual level. For instance, a comment about how a sequence of interaction should be designed was also a comment about what it would mean for the experience if the interaction was redesigned.

68 of the comments were, in the joint coding, coded as detailed, i.e. as relating to *how to design* and implement solutions (structure, interaction and presentation). The total number of comments in this analysis was 112 rather than 122, since 10 randomly chosen comments had been excluded after being used for training and checking the coding scheme. 44 (64% of comments) in the walkthroughs with pauses compared to 24 (56% of comments) in the walkthroughs without pauses contained detailed feedback. More detailed feedback can accordingly be seen in the walkthroughs with pauses.

The detailed feedback during the walkthroughs with pauses consisted of 27 suggestions, 11 negative, and 6 positive comments. The walkthroughs without pauses generated 13 suggestions, 5 negative, and 6 positive comments. If we compare these numbers to the total number of comments of each type, we can observe that the rate of suggestions (48% vs. 23%) and negative comments (44% vs. 20%) in the walkthroughs with pauses was substantially higher than they were in walkthroughs without pauses. There was no difference for positive comments (15% for both).

In this particular case, the additional comments generated during the walkthroughs with pauses were suggestions and negative comments regarding the prototype. The results of the analysis show that there were more comments on how to design and implement the solution in the walkthroughs with pauses than without pauses.

## 4.3 Long-Term UX

The third and final question was if there was more feedback about long-term UX in the service

walkthroughs without pauses. To answer that question we looked for comments on long-term UX by analysing whether a comment concerned the outcome of an activity, whether it was about several service moments at the same time, or if it was related to the cultural or social context. If it was neither of these it was considered to be a comment concerning momentary user experience.

Two researchers analysed the material separately. The interrater reliability (Cohen's Kappa) showed a substantial agreement (K = .61) (Landis & Koch, 1997). The whole material was 122 comments.

The coders agreed in all instances except 21. There was agreement on 19 long-term comments and on 82 momentary experience comments. Of the comments that there was agreement on, a vast majority related to operations or actions and their outcomes. Not many comments at all concerned the outcome of the activity. Of the 19 long-term UX comments, 10 were collected during the walkthroughs without pauses and 9 during the walkthroughs with pauses. 36 comments about momentary experience were identified in the walkthroughs without pauses and 46 during the one with pauses. Of the total agreed upon comments, 10 of 46 (22%) during the walkthroughs without pauses, and 9 of 55 (16%) during the one with pauses, concerned the long-term user experience. This result means that there was no difference to speak of between the walkthroughs without pauses and walkthroughs with pauses.

## 5. DISCUSSION

In the following sections, the results will be related back to the research questions.

### 5.1 There Was a Quantitative but No Qualitative Difference in Feedback

The first research question was: What types of feedback are there in the service walkthroughs with pauses compared to the service walkthroughs without pauses? In short, there were the same proportions of suggestions, negative comments and positive comments in both kinds of walkthroughs. There were, however, a larger number of comments raised in the walkthroughs with pauses compared to the ones without pauses. A likely explanation is that were more opportunities to provide feedback, since there was a stop for discussion after each station.

### 5.2 The Walkthroughs with Pauses had More Detailed Feedback

The second question was: Is there more detailed feedback in the walkthroughs with pauses? For the case we studied here, the answer to that question is 'yes, particularly suggestions and negative

comments'. The results accordingly suggest that there may be an advantage to pause the walkthrough if more detailed feedback is the goal of the prototype. This is in accordance with experiences from the pluralistic walkthrough method (Bias, 1994), and experience prototyping (Buchenau & Fulton Suri, 2000).

We believe the reason for getting more detailed feedback in the walkthroughs with pauses is a result of asking questions about a service moment directly after it. This made the feedback more situated, embodied, and easy to remember, since e.g. cues in the environment could guide the participants.

### 5.3 The Walkthroughs without Pauses did not Reveal More About Long-term UX

The third question was: Is there more feedback about long-term UX in the service walkthroughs without pauses? In relation to the particular case studied here, the short answer is 'no'. Most comments were not about the long-term experience of the walkthrough. Our expectation was that the walkthrough without pauses could have more comments about the whole long-term experience, since it would not interfere with the flow of events. 22% long-term comments in the walkthrough without pauses, and 16% in the one with pauses, do not provide any real support for our initial hypothesis. This result also implies that the pauses did not interfere with the overall experience. However, it could also mean that the walkthrough, as it was set up, did not generate much long-term UX feedback at all.

### 5.4 Improving the Theoretical Framework

There were discrepancies in the individual coding that forced us to use consensus coding. Discrepancies in coding are interesting to take note of since they indicate that the theoretical framework may require clarification or further development. The discrepancies were between 1) content and presentation, between 2) function and interaction, and 3) how to view structure.

The first discrepancy was content – presentation. Is the choice of media type (audio, video, text etc.) a matter of content or how it is presented? In the end we regarded it as content. Furthermore, when the participants said that the stations should be varied, is that a general matter of how content should be, or is a detailed matter of how it should be presented? In our final analysis we chose to categorise it as presentation.

The next discrepancy was function – interaction. When the feedback consisted of evaluations of how it was to use a function it could be interpreted as either function or interaction (since using a function is often done through interaction). In our final

analysis we chose however to categorise it as function if it was a matter of what people did and the response they got.

The third discrepancy was how to view structure. The structure level is close to both what to design and how to design. It is therefore sometimes difficult to identify, and easy to mix up with function and content, as well as with interaction and presentation.

Additional discrepancies were judgements where comments about details propagated consequences on a more conceptual level, as when a comment about interaction had consequences for the overarching experience. The theoretical framework of abstraction/detail levels in interaction design can in the light of this study be improved in the following way:

- The choice of media form (audio, video, text etc.) belongs to Functions and Content, whereas the exact design of it belongs to Presentation.
- The experience of similarity and difference between two or more moments in the customer journey belong to Presentation, if it is not a matter of media form.
- The experience of using a function belongs to the level of Functions and Content. It belongs to Interaction only if it is related to the operations people perform and the responses they get on those operations.

The theoretical frameworks used can be used to analyse verbal feedback from prototyping in other cases as well. This opens up the possibility to compare different techniques based on the feedback they generate.

### 5.5 Methodological Issues

There are some methodological issues to consider in this study. First of all, this paper reports a case study of a design project where four service walkthroughs were conducted. Two were of one kind and two of the other kind. This provided us with an opportunity to take a closer look at what the differences might be and put them into the light of earlier experiences in pluralistic walkthrough and experience prototyping. Our results largely support those earlier experiences, despite the differences between the methods. The results are based on case study research and generalisations should therefore be made with care. Experimental research is necessary to put cause and effect relations to test.

The case study gives credibility to results through ecological validity, even though it was students who participated in the walkthroughs. Students are one

possible target group in this particular design case. In general we need, however, more experimental research in design methods research. Experiments with groups rather than individuals require however quite a large number of participants, but it would be worthwhile for future research.

## 6. CONCLUSIONS

The purpose of this case study has been to develop an understanding for the design feedback received from service walkthroughs with and without pauses.

The results show no differences in the qualitative content. However, slightly more comments were generated in the walkthrough with pauses, and it provided more detailed feedback on how to design the service, but not on what the service should be. We expected that the walkthrough without pauses would give rise to more comments on long-term user experience, but that was not the case.

The lesson learned from this case study is hence that walkthroughs with pauses can elicit more detailed feedback on how to design the service, compared to a walkthrough without pauses. How far this lesson can be generalised beyond this particular case study is a matter for future research, but our results largely follow experiences from pluralistic walkthroughs and experience prototyping. Similar results for different design and evaluation methods give stronger support for the claim that pauses is good for more feedback in general and more detailed feedback in particular.

## 7. REFERENCES

Arvola, M., & Artman, H. (2007). Enactments in Interaction Design: How Designers Make Sketches Behave. *Artifact*, 106-119.

Arvola, M., Blomkvist, J., Holmlid, S., & Pezone, G. (2012). A Service Walkthrough in Astrid Lindgren's Footsteps. *Service design and service innovation conference, ServDes.* Espoo, Finland.

Bias, R. G. (1994). The Pluralistic Usability Walkthrough: Coordinated Empathies. In J. Nielsen, & R. Mack, *Usability Inspection Methods* (pp. 63-76). Wiley and Sons.

Blomkvist, J., & Holmlid, S. (2010). Service prototyping according to service design practitioners. *ServDes.2010.* Linköping, Sweden: Linköping university electronic press.

Blomkvist, J., Åberg, J., & Holmlid, S. (2012). Service Walkthrough to Support Service Development. *Proceedings of Service Design*
*and Innovation Conference, ServDes.* Espoo, Finland: LiU Electronic Press.

Buchenau, M., & Fulton Suri, J. (2000). Experience Prototyping. *Proceedings of the 3rd conference on Designing interactive systems: processes, practices, methods, and techniques* (pp. 424-433). New York: ACM.

Burns, C., Dishman, E., Verplank, W., & Lassiter, B. (1994). Actors, Hairdos & Videotape - Informance Design. *Conference Companion on Human factors in Computing Systems* (pp. 119-120). Boston, MA, USA: ACM.

Graneheim, U. H., & Lundman, B. (2004). Qualitative Content Analysis in Nursing Research: Concepts, Procedures and Measures to Achieve Trustworhiness. *Nurse Education Today, 24*, 105-112.

Holtzblatt, K., & Jones, S. (1993). Contextual Inquiry: A Participatory Technique for System Design. In D. Schuler, & A. Namioka (Eds.), *Participatory Design: Principles and Practices* (pp. 177-210). Hillsdale, New Jersey, USA: Lawrence Erlbaum Associates.

Koivisto, M. (2009). Frameworks for structuring services and customer experiences. In S. Miettinen, & M. Koivisto, *Designing Services with Innovative Methods* (pp. 136-149). Keuruu, Finland: Kuopio Academy of Design.

Landis, J. R., & Koch, G. G. (1997). The Measurement of Observer Agreement for Categorical Data. *Biometrics, 33*(1), 159-174.

Luojus, S. (2010). *From a Momentary Experience to a Lasting One: The Concept of and Research on Expanded User Experience of Mobile Devices.* Oulu, Finland: Acta Univ. Oul. Dissertation A 559.

Mattelmäki, T. (2006). *Design Probes.* Helsinki: University of Art and Design.

Nilsson, S., Arvola, M., Szczepanski, A., & Bång, M. (2012). Exploring Place and Direction: Mobile Augmented Reality in the Astrid Lindgren Landscape. *Proceedings of OzCHI 2012.* Melbourne, Australia.

Oulasvirta, A., Kurvinen, E., & Kankainen, T. (2003). Understanding contexts by being there: case studies in bodystorming. *Personal and Ubiquitous Computing, 7*, 125-134.

Schön, D. A., & Wiggins, G. (1992). Kinds of Seeing and their Functions in Design. *Design Studies, 13*(2), 135-156.

Wright, P., Wallace, J., & McCarthy, J. (2008). Aesthetics and Experience-Centered Design. *ACM Transactions on Computer-Human Interactions, 15*(4), 21.