
PERRON – Enhanced Pedestrian Routing and Navigation as well as Walkability Assessment of Pedestrian Ways

ERA-NET TRANSPORT III Flagship Call 2013
Future Travelling – Information and Services for Users

Start date of project: October 1st, 2014
Duration: 36 months

D6.1 Evaluation Plan

Cristina Olaverri-Monreal (Austrian Institute of Technology GmbH)

Stephanie Schwarz (Austrian Institute of Technology GmbH)

David Sellitsch (Austrian Institute of Technology GmbH)

Anita Graser (Austrian Institute of Technology GmbH)

Markus Straub (Austrian Institute of Technology GmbH)

Sebastian Naumann (Institut f. Automation und Kommunikation e.V. Magdeburg)

Eva Ganglbauer (Fluidtime Data Services GmbH)



Version: 2
Date: 05/01/2016
Dissemination level: (PU, PP, RE, CO): CO

Summary

In this deliverable a comprehensive testing and assessment plan for the PERRON project is presented. The plan includes testing scenarios, methods, metrics, selection and recruiting of users, with identified roles and actions for all evaluation activities as well as the definition of resources and a evaluation roadmap. In particular, the plan details iterative loops, through which the evaluation results are fed back to the respective work packages for further development.

Table of Contents

1. INTRODUCTION	4
1.1 BACKGROUND	4
1.2 USER INVOLVEMENT.....	4
1.3 GOALS AND RESEARCH QUESTIONS.....	5
2. EVALUATION PHASES.....	7
2.1 NAVIGATION INSTRUCTION STUDY	7
2.2 HEURISTIC EVALUATIONS	7
2.3 SECOND LAB STUDY.....	9
2.1 OUTDOOR STUDY.....	9
2.2 FIELD TRIAL.....	10
2.3 METHODS	11
2.3.1 <i>Thinking-Aloud method</i>	11
2.3.2 <i>Questionnaires</i>	11
2.3.3 <i>Interviews</i>	12
2.3.4 <i>Observations</i>	12
2.3.5 <i>Data Logging</i>	12
2.3.6 <i>Experience Sampling</i>	12
3. RISK MANAGEMENT	12
3.1 DAMAGE OF THE TEST DEVICES	12
3.2 RISK OF FINANCIAL OR PHYSICAL HARM FOR PARTICIPANTS.....	13
3.3 DROP-OUT RISK AVOIDANCE.....	13
4. EVALUATION ROADMAP	13
5. REFERENCES	14

1. Introduction

1.1 Background

ISO 9241-210 specifies four major design activities in UCD (user centered design): Understanding and describing the context of use, specifying requirements, design and evaluation of the solution in an iterative process. In the deliverables “D1.1 Description of Scenarios and User Groups” and “D1.2 Report on Requirements Document for Pedestrian Navigation and Quality Criteria of Pedestrian Ways”, we have reported about a detailed requirement analysis and specified three user groups for our solution. For user involvement in WP6 we will specifically target on smartphone power users, i.e. people who usually text and read on their phones in public and mobile situations e.g. while crossing a road. This might result in unsafe situations. Overall implications will be to supporting these users through routes with crosswalks or traffic lights for a safer crossing as well as persuade them to pay attention to the traffic. More specifically we will focus our investigation on the two navigation contexts:

- **Navigating in unfamiliar environment:** Pedestrians who use navigation tools for wayfinding to explore an unfamiliar environment **or** to find a specific place/ location e.g. shopping sites. Sightseeing might play an essential role for them. Implication:
 - There can be different motives for using navigation systems (other than saving time).
 - Joy of exploring a city without getting instructions from the smart phone device is essential to ensure an optimal user experience. Navigation instructions will be given only by demand.
- **Wayfinding in a shopping scenario:** Pedestrians who use navigation tools to find a specific place/ location e.g. shopping place. Implication:
 - We assume that people have several shopping destinations within the same trip.

In this document we will plan future evaluation activities while taking our target groups and navigation contexts into consideration.

1.2 User involvement

User will be selected according to the specified characteristics. We plan to invite different amounts of users for our lab studies and field trials in Vienna and Magdeburg (see Table 2).

The main selection criteria for test users are:

- Users must be at least 18 years old
- They must be smartphone users (mobile internet usage on a daily basis)
- They are familiar with using navigation systems (e.g. in vehicles or as pedestrians, including turn-by-turn instructions for pedestrians)
- They must be willing to sign an informed consent

Demographic data describing the study participants, such as: gender, age, profession, technical affinity, previous navigation system usage will be collected for every participant.

We will ensure users' privacy by anonymizing collected data. An informed consent form will be handed out to be signed by all involved participants. The informed consent ensures that the participant becomes fully informed about the research study and that he/she will agree based on the provided information. Participation is only possible if the informed consent is read and signed by the invited users. All participants will receive an allowance after the study from the facilitator.

1.3 Goals and research questions

For the evaluation of the pedestrian navigation system we will focus on

- Acceptance in terms of usefulness, ease of use, behavioral intention toward turn-by-turn vs. landmarks based instructions, acceptance differences depending on the trip context and preferences of landmark categories
- Usability in terms of performance, heuristics, understandability, timing of instructions
- User Experience of information and interaction design, appearance, layouts
- including satisfaction, joy of use, privacy, perceived safety as well as on experience feedback in terms of perceived way quality (safety, accessibility, attractiveness and comfort).

Furthermore objective safety, way quality and contextual factors will be investigated in later phases of the evaluation process in order to learn about the appropriate context for using a system for provided navigation instructions.

These will be evaluated for the turn-by-turn instructions as well as the landmark navigation hints according to the research questions shown in Table 1 below.

	Research questions	Evaluation phase
Content (turn-by-turn, landmarks)	<ul style="list-style-type: none"> • Acceptance: <ul style="list-style-type: none"> ○ How useful pedestrians perceive turn-by-turn ((better than navigation by line on map?) vs. landmarks based (better than ordinary/car-like turn-by-turn?) instructions? ○ Are there acceptance differences in terms of usefulness, ease of ease depending on the trip context? ○ Are preferences toward landmark categories dependent from the trip nature or the familiarity with the area? • Usability: <ul style="list-style-type: none"> ○ Were the provided turn directions correct? ○ Were any turn instructions missing? ○ Were there any superfluous turn-by-turn instructions? ○ Were the instructions timed well? (Not too early, not too late) ○ How useful were selected landmarks? ○ Which other landmarks should have been used instead? ○ Which types of landmarks proved particularly useful, e.g. buildings, characteristics of the path (such as e.g. steps, surface conditions), street furniture? ○ Were there problems with certain groups of landmarks? 	<ul style="list-style-type: none"> • 1st lab study - instructions • 2nd lab study - instructions, interaction design
presentation (turn-by-turn, landmarks)	<ul style="list-style-type: none"> • Acceptance <ul style="list-style-type: none"> ○ Do pedestrians prefer text to speech instructions or reading them on the screen? ○ Do users prefer visual and speech instructions combined depending on the context • Usability <ul style="list-style-type: none"> ○ Were the instructions clearly presented? 	<ul style="list-style-type: none"> • Heuristic evaluations (Expert evaluation of user interface mockups and click dummies) • 2nd lab study- instructions, interaction design • outdoor study • field trials
Overall (Content, presentation)	<ul style="list-style-type: none"> • User Experience <ul style="list-style-type: none"> ○ Satisfaction ○ Perceived safety and trust toward navigation instruction ○ Was the information presentation jeopardizing road safety? ○ Way quality and contextual factors (Which is the appropriate context for using a system for navigation instructions?) 	<ul style="list-style-type: none"> • 2nd lab study- instructions, interaction design • outdoor study • field trials

Table 1. Research questions and evaluation phases mapping

2. Evaluation Phases

In Austria, AIT and FLU will be responsible for the recruitment of the participants for the navigation instruction study and the lab study as well as the heuristic evaluation and the field trial in Austria. IFAK will lead the user trial in Germany. The number of the participants for the evaluation planned is shown in Table 2.

	Navigation instruction Study	Heuristic evaluations	Lab Study	Outdoor Study	Field Trial AT, DE
User involvement	8 - 12 power users	2-3 usability experts	8 - 12 power users	30 power users	2 x 12 - 15 power users
Focus of research	Acceptance (content)	Usability (content, presentation)	Usability, acceptance (presentation)	User experience (overall), Way quality	User experience (overall), Way quality
Organised by	AIT	AIT/FLU	AIT	AIT	IFAK/FLU/AIT
Timeline	10.-12.15 (M13-15)	11.15 – 04.16 (M14-19)	(M21 – 23)	01.17 (M28)	04.17 (M31)

Table 2. User involvement in PERRON evaluation activities

This sub-section presents evaluation methods that require the involvement of users. At the beginning of the design and development lifecycle heuristic evaluation ensure a “critical mass” of usability. This prevents expending resources for user-based methods in early stages. User-based evaluation methods will be used in lab studies and field trials.

2.1 Navigation Instruction Study

In the navigation instruction study we aim to investigate timing and frequency of directions when approaching a decision point/intersection and while walking along straight route segments. We also want to investigate how instructions should be presented in pedestrian navigation systems (direction sector assignment, phrasing, etc.).

12 users will participate in an outdoor study where they will individually walk 2 routes (10 minutes each) in the first district of Vienna. Researchers will show users print outs of different types of linguistic instructions at decision points, along the route and at intersections. Participants will be asked to point out the timing of confirmation instructions and provide feedback on the quality of instructions. Users will request confirmation instruction when they need them, while researchers will log the GPS data when on sections where users need confirmation. Additional data such as demographic information and walking routines will be collected in a post-test survey.

2.2 Heuristic Evaluations

At the beginning of the design and development lifecycle heuristic evaluation ensure a “critical mass” of usability. This prevents expending resources for user-based methods in early stages.

In heuristic evaluations usability experts aim to uncover most usability problems of software. Unlike other evaluation methods in heuristic evaluations there are no users involved. Selected experts (around 5 experts are ideal according to Nielsen) assign errors they have found to a list of heuristics. Niensens's heuristics are based on experience. The approach provides results fairly quick, thus supporting agile design and development processes. Heuristic evaluation does not replace user involvement in later stages of the evaluation, but it will help identifying usability issues in early stages efficiently. For our evaluations we will rely on the Nielsen's 10 heuristics. [Nielsen 1995]:

- **Visibility of system status**
Keep users informed about what is going on. Provide feedback within reasonable time.
- **Match between system and the real world**
Speak the users' language. Use relatable real world concepts rather than system-oriented terms. Make information appear in a natural and logical order.
- **User control and freedom**
Provide "emergency exits". Support undo and redo.
- **Consistency and standards**
Use consistent language. Follow platform conventions.
- **Error prevention**
Eliminate error-prone conditions or present users with a confirmation option before they commit to the action.
- **Recognition rather than recall**
Make objects, actions, and options visible. Make instructions visible or easily retrievable.
- **Flexibility and efficiency of use**
Create Shortcuts and allow users to tailor frequent actions.
- **Aesthetic and minimalist design**
Display only relevant information. Consider aesthetics
- **Help users recognize, diagnose, and recover from errors**
Display precise error messages in plain language and suggest solutions.
- **Help and documentation**
Provide efficient, easy to find documentation.

Heuristic Evaluations will be conducted by 2-3 experts in Vienna. Feedback from the experts will be structured according to the 10 heuristics by Nielsen.

In periodic iterations AIT will give feedback about designs, mockups and click dummies in their current state and will thereby uncover and "filter" usability issues before users are involved.

2.3 Second Lab Study

This study will evaluate the system in its current state, while the prototype should be implemented horizontally, meaning all the user interface elements and interactions should already be in place while functionality can be mocked.

AIT will evaluate the quality of the information and interaction design of the first and second lab studies using observation, semi-structured interviews, and questionnaires.

In Austria representatives of the defined target group who fulfill the inclusion criteria for participation will be invited. Invited users will evaluate the use of the PERRON prototype under standardized lab conditions. For this purpose, observations, quantitative measurements and interviews will be applied to gain realistic data that will be integrated into the subsequent modification and development process of the PERRON tool. The evaluation procedure includes three main stages:

- Pre-interview
- Prototype evaluation on the basis of selected test scenarios followed by post scenario questionnaires
- Post-interview and post study questionnaires

The beginning of the trial constitutes a detailed introduction from the facilitator who will explain the PERRON system and the functions it offers to the participants as well as the purpose of the user evaluation. Results from lab studies and interviews are analyzed and constitute the basis for the further technical development that will be implemented for the field trial phase.

2.1 Outdoor study

A researcher will observe participants on their routes, take notes and record a video. At certain sections the researcher will ask questions about landmarks named by the navigation system. Were they easily recognizable? Were there other landmarks which the user would prefer?

After the trip, the user will have to fill in a post task questionnaire and will be interviewed.

Route selection will cover roads with low, medium and high traffic volume, intersections, and straight route segments. The basic structure of a study session will be as follows:

- Introduction
 - Short Introduction of facilitator
 - Introduction of the project PERRON
 - Short Introduction of the session procedure
 - Administrative aspects and informed consent
- Setup of the eye tracking supply
 - Instruction
 - Calibration
 - Way finding tasks with the PERRON interface for mobile navigation
- Post-Interview
 - Post-Interview
 - Questionnaires (Self-assessment of spatial abilities, Spatial abilities power test)
- Closing Session
 - Closing remarks, thanking, conclusions
 - Handing out financial compensation

The outdoor usability and user experience study will take place in Vienna, AT within the evaluation site of the first Viennese district and defined dates (see Figure 3). The recruitment will be done by AIT. An informed consent document has to be signed by participants as well as by the facilitators in advance. Each session will last approx. 1.5 hours.

2.2 Field trial

Towards the end of the project the PERRON pedestrian navigation concepts and routing methods will be evaluated with users in Austria and Germany. Targeted number of participating users will be around 24 to 30 (see Table 2).

Field trials will be conducted in Magdeburg and Vienna.

Introduction Workshop

The introduction workshop will be held at each target site at the partner's facilities. Participants will be invited to appear at pre-defined time slots in small groups (4-6 people) in order to avoid too much crowding. Participants will fill a short questionnaire on basic demographic information and previous technology experience.

After a brief introduction to PERRON and the study background (10min), facilitators will hand out the test devices and tools. The final prototype application will be shown and usage will be trained. Then a pre-questionnaire will be distributed to gather first impressions from participants.

Participants will be instructed that when they receive text messages and from facilitators during the field phase they should follow the instructions (e.g. follow a link in a text message and fill a short online questionnaire in the smartphone).

Evaluation in the Field

Participants will use the PERRON application overall several weeks in their daily life. Users will be asked to plan and walk a route for a specific purpose within the study site in Vienna or Magdeburg. The purpose can either be:

- a shopping trip,
- a sightseeing trip
- a situation where they are busy texting on their phone while they are walking.

In this period facilitators from AIT will be available for giving support.

An experience sampling questionnaire will be displayed on the smartphone at the end of each navigation activity, asking users about how safe, attractive, accessible and comfortable they perceived the section. Trip purpose and familiarity with the environment regarding the route will also be assessed via experience sampling.

Data related to the interaction with the system will be logged for further analyses. Log files will include a for example a timestamp, the GPS position, as well as the current route and the personal route quality setting (see Table 3).

Route ID	Timestamp	Position: latitude	Position: longitude	Current Route: From	Current Route: To	Personal Route Settings:
111	14.12.2016 14:43:31	48.209137	16.372887	Stephansplatz	Schwedenplatz	Comfortable route

Table 3: Logging example

Preliminary Analysis of Results

At the end of the field phase, in a preliminary analysis AIT will screen the collected log files and data from weekly questionnaires to extract first results. Subsequently usage activity of participants will be screened according to their participation in the field phase. Motivated and interested participants will be notified to participate at the final workshops.

Final Workshops

After the preliminary analysis of the field study results final workshops will be conducted with a reduced number of participants (selected on the basis of a screening of user activities from the

preliminary analysis) in order to deepen and further elaborate the findings on the user experience of the PERRON application. In the workshops participants should be enabled to share thoughts and reflect on each other's activities.

2.3 Methods

To evaluate the content, presentation and user experience related to the routing and navigation tool that will be deployed, following methodology will be used during the design phase of the prototypes.

2.3.1 Thinking-Aloud method

Thinking-aloud is a method where users are encouraged to verbalize what they are thinking while interacting with the system in an evaluation setting. This provides insight about how users approach a task, reveals their thought process and misconceptions [Nielsen, 1993]

2.3.2 Questionnaires

In this section we present several different types of questionnaires which will be distributed to study participants.

Demographic information

Demographic data describing the study participants, such as: gender, age, profession, technical affinity, previous navigation system usage will be collected for every participant. The collected data allows a comparison between the users' understanding of the evaluated system and her demographical attributes.

Spatial abilities

Spatial abilities are involved in most of our daily orientation tasks in two and three dimensional surroundings (e.g. reading maps, using navigation tools). Beside psychometric approaches such as mental rotation tests, self-assessments for sense of direction such as the Santa Barbara sense of direction scale (SBSOD, Hegarty et al., 2002) have shown to provide valuable insights on a person's spatial skills. The SBSOD scale specifically is used to search for possible relationships between individual sense of direction and usage of routing and navigation tools.

Usability

In order to measure the perceived usability of a system we will use various usability measurement methods, such as the System Usability Scale (SUS) [Brooke 1996] and Single Ease Question (SEQ). SUS is used for measuring several different aspects of the usability of a system. It consists of ten (5-point) Likert scale questions (see Figure 1). The single ease question is a simplistic approach that can easily be applied in evaluations. By asking users to rate the question "Overall, how difficult or easy was the task to complete?" on a 7-point Likert-scale the SEQ provides a quick way to evaluate overall usability of a system.

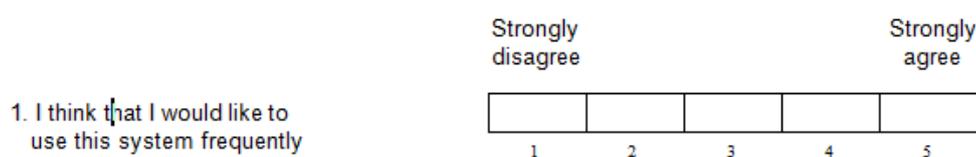


Figure 1: Example of SUS

User experience

AttrakDiff is used for measuring pragmatic as well as the hedonic quality of products [Hassenzahl et al., 2003]. The questionnaire shows how attractive the product is for users. The AttrakDiff uses several word-pairs (c.f., Figure 2) for measuring users' attitudes and emotional state regarding a product.



Figure 2: Example of AttrakDiff¹

2.3.3 Interviews

Interviews like questionnaires are an indirect method, since no data about a system or an interface itself is gathered. They allow gathering in-depth information regarding users' ideas, wishes and assumptions. Interviews can be categorized into unstructured, structured, semi structured interviews [Preece 2002]. In Perron we will use semi-structured interviews in which questions and topics are prepared, but the interviewer may ask spontaneous questions, thus creating a more natural feeling during the conversation.

2.3.4 Observations

We will observe users while they are interacting with the system. Users will be (video) recorded, which will allow us review certain parts of the evaluation in order to study details. We are also planning to use a mobile eye tracking system in the lab studies. Observations will provide insights how users interact with the system in the lab as well as in the field in real world contexts. This will also allow us to understand problems in the handling of the application, as well as where technology is successful in providing navigational instructions to support users.

2.3.5 Data Logging

All interactions users make during the field trials will be logged on the device including a timestamp. This allows detailed analysis and makes results from different user groups comparable.

2.3.6 Experience Sampling

Experience sampling [Larson & Csikszentmihalyi, 1983] is an in situ data collection method. Users are interrupted at certain times while using a device in order take notes or answer questionnaires about their experience. We will use experience sampling in the field trials and will trigger interrupts at the end of navigation segments.

3. Risk Management

In this section the risks especially related to the final field trials will be presented and the risk management strategies will be introduced.

3.1 *Damage of the Test Devices*

PERRON prototypes will be standard Android applications that do not pose any particular danger to the users' private devices. However, to avoid any claims for indemnification users will waive any rights for such claims as part of the informed consent they sign.

Evaluation Phase	Requirements	Organisation	Month
Navigation instruction study	Selected routes within the test site. Differently phrased instruction alternatives in printed form.	AIT, IFAK	13-15
Heuristic evaluation	User interface mockups and click dummies being developed with the User interface tool Invision ² . The screens include different examples of turn-by-turn and landmark information as generated by the router.	AIT, FLU	14-19
Lab study	Horizontal prototype covering GUI and interactions as wireframes where most important functions will be clickable.	FLU	21
Outdoor lab study	Prototype which is testable for UX, Acceptance and Usability	AIT	28
Field trial (Vienna & Magdeburg)	Working prototype including data logging and experience sampling questionnaire integration.	AIT, IFAK, FLU	31

Table 5: Requirements for evaluations

5. References

[Brooke, 1996]

Brooke, J. (1996). SUS-A quick and dirty usability scale. Usability evaluation in industry, 189(194), 4-7.

[Hassenzahl et al. 2003]

Hassenzahl, M., Burmester, M., & Koller, F. (2003). AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In Mensch & Computer 2003 (pp. 187-196). Vieweg+ Teubner Verlag.

[Larson & Csikszentmihalyi, 1983]

Larson, R., & Csikszentmihalyi, M. (1983). "The experience sampling method". New Directions for Methodology of Social and Behavioral Science, 15, 41-56

[Nielsen, 1995]

Nielsen, J. (1995). Ten usability heuristics. Website: http://www.useit.com/papers/heuristic/heuristic_list.html visited on July 22nd 2015.

[Nielsen, 1993]

Nielsen, J. (1993). Usability Engineering. Morgan Kaufmann, San Francisco, CA, USA, 1st edition.

[Preece, 2002]

Preece, J. (2002). Interaction design: Beyond Human-Computer Interaction. J. Wiley & Sons, New York, NY.

[Venkatesh & Bala, 2008]

² <http://www.invisionapp.com/>

Venkatesh, V., & Bala, H. (2008). Technology acceptance model 3 and a research agenda on interventions. *Decision sciences*, 39(2), 273-315.