DELIVERABLE REPORT
D9.1.2
"Field Trial and Evaluation Plan"

Collaborative project

MASELTOV
Mobile Assistance for Social Inclusion and Empowerment of Immigrants with Persuasive Learning Technologies and Social Network Services

Grant Agreement No. 288587 / ICT for Inclusion

project co-funded by the
European Commission
Information Society and Media Directorate-General
Information and Communication Technologies
Seventh Framework Programme (2007-2013)

Due date of deliverable:	December 31, 2013 (month 24)
Actual submission date:	December 23, 2013
Start date of project:	January 1, 2012
Duration:	36 months

Work package	**WP9 – Field trials and Evaluation**
Task	**T9.1 – Preparation of Field Trials and Evaluations**
Lead contractor for this deliverable	**CURE**
Editor	**Stephanie Deutsch (CURE)**
Authors	**Stephanie Deutsch, Jan Bobeth (CURE)**
Quality reviewer	**Lucas Paletta (JR)**

Project co-funded by the European Commission within the Seventh Framework Programme (2007–2013)		
Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

partner		Organisation	ctry
01		JOANNEUM RESEARCH FORSCHUNGSGESELLSCHAFT MBH	AT
02		CURE – CENTER FOR USABILITY RESEARCH AND ENGINEERING	AT
03		RESEARCH AND EDUCATION LABORATORY IN INFORMATION TECHNOLOGIES	EL
04		FUNDACIO PER A LA UNIVERSITAT OBERTA DE CATALUNYA	ES
05		THE OPEN UNIVERSITY	UK
06		COVENTRY UNIVERSITY	UK
07		CESKE VYSOKE UCENI TECHNICKE V PRAZE	CZ
08		FH JOANNEUM GESELLSCHAFT M.B.H.	AT
09		TELECOM ITALIA S.p.A	IT
10		FLUIDTIME DATA SERVICES GMBH	AT
11		BUSUU ONLINE S.L	ES
12		FUNDACION DESARROLLO SOSTENIDO	ES
13		VEREIN DANAIDA	AT
14		THE MIGRANTS' RESOURCE CENTRE	UK

CONTENT

1. INTRODUCTION

The evaluation plan presented in this deliverable describes how the MASELTOV services will be evaluated at the various stages of the project. Early and continuous feedback is very important to avoid the development of undesired functionalities and not usable services. Especially when designing for a vulnerable target group such as immigrants who just arrived in a new country and have a different cultural background (see MASELTOV target group definition in D2.3.1). For this reason, immigrants are involved at any stage of the design process of the MASELTOV services. For a better understanding this user-centred approach is described in chapter 2. presenting the overview and status of the accordant tasks in Figure 1.

This deliverable is updated twice in the course of the project. In the first version the iterative evaluation of user interfaces was explained (See D9.1.1.). The actual document represents the second version with the detailed setup of the first field trials, while the third version will describe the planning for the assessment of the final integrated prototype of the MASELTOV project (to be described in D9.1.3.). Furthermore, the actual as well as the upcoming version of this document contains an updated version of Figure 1 showing the status within the user-centred design process in the project.

2. USER-CENTRED DESIGN

In this chapter we describe the general approach of involving users in the design process and the service development in the MASELTOV project.

2.1 DEFINITION AND ORIGIN

User-centred design is an approach to software development which has been evolving in the past decades. The goal of user-centred design is to make a product easy to use for its intended users which, however, might be a difficult task (Vredenburg et al. 2001). The process has emerged from traditional software development approaches to help keeping the focus on the user's needs and not to get lost in solving technical challenges first (Lowdermilk 2013). The main idea is to involve users from the start and in all stages of a software project in an iterative manner: to move from a technology-driven approach to a user-driven approach (Vredenburg et al. 2001). User-centred design is not only about understanding the users of a system but also requires an understanding of the tasks that users will perform with the system and of the environment in which the system will be used (Stone et al. 2005). This process is based on four main principles about how interactive systems should be developed as certified in the ISO standard 9241-210 (2010):

- The active involvement of users
- An appropriate allocation of function between user and system
- The iteration of design solutions
- The engagement of a multidisciplinary design team

This means to ask or observe users for their needs, to present design ideas frequently to users for their feedback and to update the design iteratively, to evaluate the functional prototypes under real conditions (Petrie and Bevan 2009). These authors provide a good overview on the important concepts of human-computer interaction and user-centred design such as usability, accessibility and user experience (Petrie and Bevan 2009). Based on these concepts and the aforementioned principles we set up the user-centred design process within MASELTOV as described in the upcoming section 2.1.

An important precondition for applying a user-centred design process is to know who the users of a system will be in order to involve participants as close as possible to the intended target group (Grudin and Pruitt 2002). However, the identification of users is particularly demanding if there are large numbers of heterogeneous users (Kujala and Kauppinen 2004). In the special case of MASELTOV where we work with immigrants who form a large and very heterogeneous group in Europe we needed to specify the target group carefully. An additional challenge in the work with immigrants is mistrust towards the researchers which hinders the recruitment of study participants and might influence study results (Hynes 2003). For these reasons, the target group definition needed to narrow down from the whole group of immigrants to a subset with similar attributes but might not be too limited to gain access to members of the target group. The MASELTOV target group is defined and updated according to recommendations of the reviews (Szwochertowska et al. 2013) in the two versions of the MASELTOV deliverable D2.3 Use Cases and Service Scenarios.

When it comes to concrete user involvement in ICT projects the question how many users are needed for the current task is raised and has been discussed extensively (Nielsen 1993, Spool and Schroeder 2001, Hwang and Salvendy 2010). In user-centred design, sampling can be

based on groups that are identified by the main user characteristics (Kujala and Kauppinen 2004).

For the elicitation of cultural-specific user requirements Aykin et al. (2006) recommended qualitative methods. Vulnerable target groups like newly arrived immigrants may be reluctant to take part in formal research studies (Atkinson & Flint 2001). Building up trust to those groups can require more than application of anonymity, confidentiality, and the use of ethical principles, for example establishing credentials by working voluntarily with a refugee community organization (Hynes 2003). Potential research participants have to have the possibilities to protect themselves in terms of their social relations with the researchers and they have to have the power to decide over whether to participate (Lammers 2005). The immigration status is very relevant regarding vulnerability and the status may change. Researcher should focus on a smaller amount of participants that are studied more intensively (e.g. observing and interviewing the concrete target group). In qualitative research, it is enough to pick up representative users from each group based on the main user characteristics which allows all necessary users to be represented (Kujala and Kauppinen 2004). Beyer and Holtzblatt (1998) recommend that between six and twenty users should be visited depending on the scope of the study. Accordingly in MASELTOV we conducted interviews and focus groups within the requirements analysis in WP2 with selected representatives of target groups.

To assess the design ideas the most important aspect is to work in iterations and to present updated design frequently to users (Stone et al. 2005). Therefore a small amount of users per iteration is enough. The precise number of necessary users is difficult to define, and it depends on the case (Kujala and Kauppinen 2004). Nielsen (1993) states that the majority of usability issues of a system (around 80 %) will be identified by the first five users. However, for more complex applications up to 15 participants were needed to identify 80% of the usability issues (Spool and Schroeder 2001). Later on, Hwang and Salvendy (2010) proposed 10±2 participants for usability studies. Schmettow (2012) argues that usability studies differ so much that a "magic number" of involved users can never be defined in a reliable way. In MASELTOV we stick to usual practice and conduct three formative evaluation iterations with 5-10 users per iteration during the design phase in the lab (T2.5 and T9.2) as experts also suggested recently (Petrie and Bevan 2009, Hwang and Salvendy (2010), Lowdermilk 2013).

For the summative testing of the working system which will take place in the field Petrie and Bevan (2009) suggest to involve between 8 and 30 users. In MASELTOV, we plan to involve 36-72 users (see DOW Part B Table 11). A more detailed planning of the final field studies will be reported in the last version of this document (D9.1.3).

A particularity of the user-centered design process in MASELTOV which distinguishes this project from other ICT project with vulnerable target groups is the external ethical approval of all user involving tasks. This means that an independent expert from the ICMPD (International Centre for Migration Policy Development) reviews all guidelines for user studies before these studies take place. More details can be found in the Ethical Manual (D1.4).

2.2 THE USER-CENTRED DESIGN PROCESS IN MASELTOV

Figure 1 presents an overview and the current status of the tasks forming the user-centred design process in MASELTOV. The user involvement was started with the requirements analysis in which we have collected specific service requirements of immigrants in order to profoundly understand their specific service needs (see D2.3.1). The next step was to design the services and discuss them with immigrants. To benefit from their ideas we conducted two

participatory design sessions (see D2.4). The goal of the high user involvement was to identify potential differences in design and solution approaches and to discuss them with end users. Subsequently and within the course of iterative interface design as part of WP2, in WP9 three usability evaluation studies were conducted, each at a different stage, respectively in January, May and in October 2013 (see Figure 1). Accordant results of these activities are reported in D9.2.1, D9.2.2, and D9.2.3.

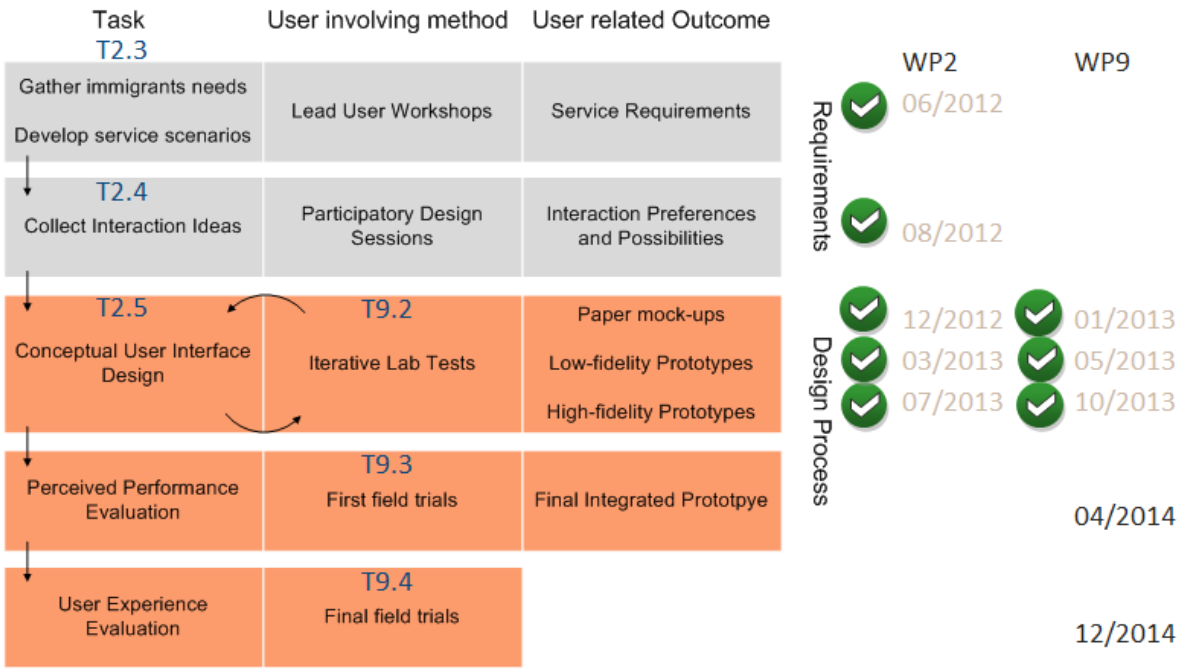

Figure 1: Overview and status of tasks within the user-centred design process of MASELTOV.

The MASELTOV user interface was designed in an iterative manner (see T2.5) based on the outcomes of the participatory design sessions and the identified user needs. After the elaboration of the user interface concepts, the resulting mock-ups were presented first to usability experts and then to the users in form of usability tests (see T9.2). Thus, the design of the user interface could be directly linked to the iterative usability testing that took place in Task 9.2, where the interface concepts were evaluated frequently from an early stage onwards to ensure user validation.

The concepts of the first iteration were visualised in a simple wireframe format. They were created rather fast and easy as necessary changes and updates were expectable after the first evaluation. As shown in Figure 1 the feedback flew directly into the next iteration step. There the concepts were refined to low-fidelity and within the third iteration to high-fidelity click-dummies. Graphics and a more sophisticated visual representation were added during this later phase of the design process when the interaction flow was specified. The high-fidelity prototypes represented the majority of the services and functions offered by MASELTOV, so that the immigrants got a real impression of how MASELTOV would work. Meanwhile the creative design phase has been achieved and the user interfaces are currently implemented, improved and connected with the accordant services (WP6, WP7 and WP8). The several service components remain separately and thus can be evaluated individually. They will be tested in the first field trials taking place in three cities (London, Madrid/Barcelona, Vienna/Graz) of the NGOs with the accordant group of immigrants (see T9.3).

After the final iteration and revision the several service components will be integrated into the final prototype of the MASELTOV system by beginning of autumn 2014. This final prototype shall cover all specified use cases and scenarios and thus provide service support in various situations. Between September and November 2014 we will conduct the final field trials in order to evaluate the developed MASELTOV services under real world conditions (see Task 9.4). Therefore, methods such as e.g. "Remote Usability Testing" (e.g. Andreasen et al., 2007) or "Experiences Sampling Methods" (e.g. Larson et al., 1983) will be applied in order to gather real time information about the usage of MASELTOV that can be analysed to determine the user experience.

3. PLANNING OF THE FIRST FIELD TRIALS (T9.3)

The functional prototypes of the MASELTOV services will be evaluated with end users in the field, in particular with the most vulnerable immigrants. Immigrants often do not speak the language of their host country, and often do not have sufficient technological skills. In this study the prototypes for the access to the several services developed in WP 7, 8 and 9 will be tested the first time under real conditions. These first field trials are for testing the functionalities of the main service components under real conditions. Although the user interfaces have been evaluated before and the technological functions have been verified before, this is important as the MASELTOV services leave the controlled lab conditions. Hence the purpose of this task is to evaluate the functionality of the services and the infrastructure they are based upon. Table 1 depicts the services and components, which are planned to be evaluated in T9.3. With the support of CURE, the MASELTOV NGOs will conduct the testing in their according country with the according users in the according language (see Table 1). For the trials the services will be available in the respective language of the immigrants.

Table 1 Overview of location and language settings for the planned service evaluations.

Services and Modules	To be evaluated in (by participants)		
	UK (Latin American)	Spain (Arabic)	Austria (Turkish)
Navigation service		x	x
Information service	x		x
Social network service	x		
Serious game	x		
GeoSocialRadar	x	x	x
Text Lens service	x	x	x
User profile	x	x	x
Dashboard	x	x	x

Reaching representatives of the primary target group in MASELTOV turned out to be challenging in the previous studies. This is why we plan to adopt a modified procedure for the conduction of the first field trials in two evaluation steps:

1. A booth-supported testing approach
2. A field trial of one week duration

A detailed guideline for the field evaluation will be provided by CURE. This test protocol will be reviewed for ethical issues by ICMPD before the user tests take place.

In the following we describe the research questions and planned methodological approaches in more detail.

3.1 RESEARCH QUESTIONS

With the following main research questions we will investigate the user experience, problems, barriers, needs, etc. immigrants have using selected MASELTOV modules over one week duration. The following main research questions will be investigated in more depth with sub research questions:

1. Which functions/features of the MASELTOV services do immigrants use?
   1.1. How useful are the provided MASELTOV services?
   1.2. How satisfied are immigrants with the usage of the different MASELTOV services over the time?
   1.3. How curious are immigrants about the different MASELTOV services over the time?
   1.4. What are the missing features/functions?
2. What problems occur with the MASELTOV services?
   2.1. How do they solve the problems (e.g. whom do they ask)?
   2.2. What kind of support would they like to have?
   2.3. What kind of improvement idea for the MASELTOV services do they have?
3. What do participants like/dislike about the MASELTOV services?

3.2 STUDY DESIGN

For the first field trial a method mix will be used in order to address the exact target groups and to gather qualitative and quantitative data through self-documentation and observation by researchers (e.g., in interviews and workshops).

3.2.1 RECRUITING AND EVALUATION AT THE BOOTH

In a first step a new approach will be realized in order to reach the real target group in the field. To do so, a booth will be set up in a public area within the defined cities in UK, Spain, and Austria, which is supposed to be highly frequented by our target groups (e.g. specific market area, public event for immigrants). The booth will be used to approach immigrants and to invite them to use selected MASELTOV services (see Table 1) on the spot. Two facilitators (with at least one speaking the mother tongue of the target group) will be at the booth for approximately one week with the goal to collect user feedbacks and contact data from potential study participants for the evaluation step 2 (see below). They will support interested immigrants using the services right away and collect their immediate feedback in terms of satisfaction and curiosity. Figure 2 depicts the planned procedure to approach interested immigrants and to invite them to use the MASELTOV services and accordant give feedback on the spot. Following questionnaires will be provided:

- In order to determine whether people match the targeted criteria a few questions on their demographics will be answered.
- To assess the aspects of usability the standardized 4-item questionnaire UMUX will be used (Finstad, 2010).
- In order to collect data regarding the feelings of the participants during the interaction they will be asked to express them with the help of the SAM figure for pleasure (Bradley & Lang, 1994).

Furthermore some open questions on their general impression and possible issues will be asked. Finally, for the interested participants, the affordability questionnaire (developed in T2.2) will be handed out.

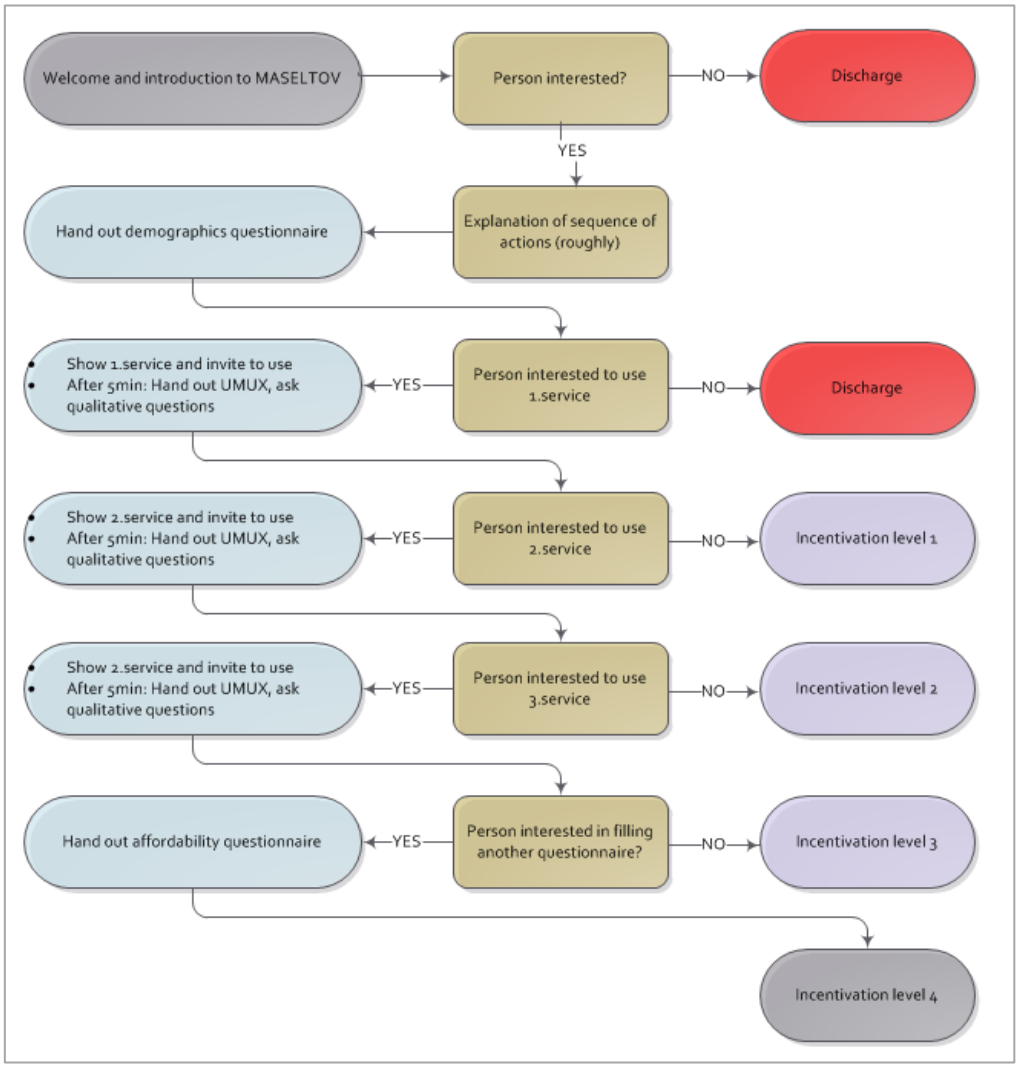

Figure 2: Procedure for the service evaluation at the booth

Facilitators will also invite participants and interested pedestrians, who match the defined target group criteria, to use selected MASELTOV services as part of the second evaluation step i.e. use the services for a week and give feedback on their extended use by attending a discussion round at the end of the field phase.

3.2.2 ONE WEEK FIELD TRIAL

In addition to the user involving evaluation activity at the booth, selected services will be evaluated then by users in the field for one week. Therefor and according to the description of work we will build on the initial plan to recruiting six key participants in every involved country who should use these components for one week under real conditions in order to detect bugs and report usability problems especially regarding the provided content.

With the support of CURE, the MASELTOV NGOs will conduct the testing in their according country with the according users in the according language (see Table 1). Per country six participants will evaluate the MASELTOV services in February and March 2014.

As defined in the description of work different foci for the prototype testing in every involved country will be set:

- **Information Services**: This field trial will focus on the mobile assistance of information needs (WP6). These trials will take place at Madrid (assisted by FUN).
- **Navigation**: In Vienna and Graz (assisted by DAN), the field trials will focus on the navigation functionality including the public transportation service provided by FLU.
- **Language and Communication Services**: This field trial is focused on the language learning aspect and the effects of the serious game (WP7) provided by COV. In particular, we concentrate on the learning of the English language by the two major target groups (assisted by MRC) and their progresses in understanding the cultural differences of the English society fostered by the serious games component. Another focus will be set on experiencing communication skills motivated by the community services (WP8).

In spite of having defined the different foci in every involved country we will not only assess the correspondent services but also the respective others, although to a minor degree. Hence several services namely the Text Lens, the Geo-Social Radar as well as the user Profile and the Dashboard view will be evaluated in all three cities, as illustrated in Table 1.

3.2.3 FINAL WORKSHOPS

After a preliminary analyse of the field study results final workshops will be conducted with some of the participants in order to deepen and further elaborate the findings on the user experience of the MASELTOV services, as a preparation for future methodological decisions to be effectuated in preparation of the final long term field trial.

3.3 EXPECTED RESULTS

The results will give insights on the impressions immigrants have at first glance about the MASETLOV Services, which functions/features they use and what they like/dislike. Additionally, a list of found problems with the services in terms of usability, user experience and content availability will be provided and ideas for improvement from the immigrants. The motivation to use the services will also be investigated and how it changes over time. Moreover, it should be found out why immigrants would (not) recommend others to use the MASELTOV services.

3.4 EVALUATION ROADMAP

Once all functional prototypes have been analysed regarding their functionality, detailed guidelines for the booth evaluation as well as the one week trial will be elaborated by beginning of February 2014, before the first field phase will be initiated in the three European cities.

In the third and last year of the MASELTOV project it is planned to integrate all modules and components in the final MApp system. The results of the first field trials to be conducted in February and March 2014 will flow back to the partners in order to update the MASELTOV

services in terms of content provision and stability for the finalisation of the final integrated MApp system. This final prototype and fully integrated system will be evaluated under real world conditions and over long time duration in the final trials in September 2014.

4. SUMMARY AND OUTLOOK

The evaluation plan of MASELTOV and the status of the user-centred design process we follow within this project are documented in this deliverable. This is the second of three versions and will be updated twice before the first field trials to be conducted in February 2014 and before the final field trials to be conducted in September 2014.

In this version of the deliverable we described how the first evaluation of the service components in the field will be organised, building the base for the conduction of the evaluation activities in T9.3. The results of the evaluation will be documented in the upcoming deliverable D9.3. Based on the detected problems and elicited recommendations the prototype implementations will be updated by responsible partners for the final integration of the MApp system.

5. REFERENCES

Aykin, N., Honold Quaet-Faslem, P., Milewski, A. E. (2006): Cultural Ergnomics. In: Salvendy, G. (ed.) Handbook of human factors and ergonomics. Hoboken: John Wiley, pp. 1418-1458

Andreasen, M. S., Nielsen, H. V., Schrøder, S. O., & Stage, J. (2007). What happened to remote usability testing? An empirical study of three methods. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems CHI '07. ACM, New York, NY, 1405-1414.

Atkinson, R., & Flint, J. (2001): Accessing hidden and hard-to-reach populations: snowball research strategies. Technical Report. Issue 33., Department of Sociology, University of Surrey, Guildford, England.

Beyer, H. and Holtzblatt, K. (1998): Contextual Design: Defining Customer-Centered Systems. San Francisco: Morgan Kaufmann Publishers.

Bradley, M.M. and Lang, P.J. Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry 25*, 1 (1994), 49–59.

Finstad, K. The Usability Metric for User Experience. *Interacting with Computers 22*, 5 (2010), 323–327.

Grudin, J. and Pruitt, J. (2002): Personas, participatory design and product development: An infrastructure for engagement. Proc. PDC 2002, 144-161.

Hwang, W. and Salvendy, G. (2010): Number of people required for usability evaluation: the 10±2 rule. Communications of the ACM 53(5), 130–133.

Hynes, P. (2003): The issue of "trust" or "mistrust" in research with refugees. New Issues in Refugee Research, Working Paper 98, Geneva: UNHCR Evaluation and Policy Analysis.

ISO (2010): Ergonomics of human- system interaction. Part 210: Human-centred design process for interactive systems. Geneva, International Standards Organization

Kujala, S., & Kauppinen, M. (2004). Identifying and selecting users for user-centered design. In Proceedings of the third Nordic conference on Human-computer interaction - NordiCHI '04 (pp. 297–303).

Lammers, E. (2005): Refugees, asylum seekers and anthropologists: the taboo on giving. Global Migration Perspectives 29.

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

Lowdermilk, T. (2013): User-Centered Design. O'Reilly.

Nielsen, J., Molich, R. (1990): Heuristic evaluation of user interfaces. In: Proceedings of the SIGCHI conference on Human factors in computing systems: Empowering people (CHI '90), ACM, 249-256.

Nielsen, Jakob and Landauer, Thomas K. A. (1993): Mathematical Model of the Finding of Usability Problems. INTERCHI'93, 206-213.

Nielsen, J. (1994): Heuristic evaluation. In: Nielsen, J., and Mack, R. L. (Eds.): Usability Inspection Methods, John Wiley & Sons, New York, NY, 25-62.

Petrie, H., & Bevan, N. (2009). The evaluation of accessibility, usability and user experience. In Constantine Stephanidis (Ed.), The Universal Access Handbook. CRC Press.

Schmettow, M. (2012). Sample size in usability studies. Communications of the ACM, 55(4), 64–70.

Spool, J., & Schroeder, W. (2001). Testing web sites: Five users is nowhere near enough. In CHI'01 extended abstracts (pp. 285–286).

Stone, D., Jarret, C., Woodroffe, M., Minocha, S. (2005): User Interface Design and Evaluation. Morgan Kaufmann Publishers

Szwochertowska, M., González Enríquez, C., Johnson, J., Sabour, M. (2013): MASELOTV YR1, Technical Review Report.

Vredenburg, K., Isensee, S., Righi, C. (2001): User-Centered Design: An Integrated Approach. Prentice Hall