Evaluating Mobile Usability: The Role of Fidelity in Full-Scale Laboratory Simulations with Mobile ICT for Hospitals

Yngve Dahl[1], Ole Andreas Alsos[2], and Dag Svanæs[2]

[1] Telenor Research & Innovation,
Otto Nielsensvei 12, 7004 Trondheim, Norway
Yngve.Dahl@telenor.com
[2] Department of Computer and Information Science,
Norwegian University of Science and Technology,
Sem Sælandsvei 7-9, 7491 Trondheim, Norway
{Ole.Andreas.Alsos,Dag.Svanes}@idi.ntnu.no

Abstract. We have applied full-scale simulations to evaluate the usability of mobile ICT for hospitals in a realistic but controllable research setting. Designing cost-effective and targeted simulations for such a purpose raises the issue of simulation fidelity. Evaluators need to identify which aspects of the research setting that should appear realistic to simulation participants, and which aspect that can be removed or represented more abstractly. Drawing on research on training simulations, this paper discusses three interrelated fidelity components—equipment/prototype fidelity, environmental fidelity, and psychological fidelity. These components need to be adjusted according to which design aspects evaluators want to gather feedback on. We present examples of how we have configured the components in various simulation-based usability assessments of mobile ICT for hospitals. The paper concludes by providing a set of guiding principles concerning the role of fidelity in simulation-based usability evaluations.

Keywords: Clinical information systems, fidelity, evaluation, human factors, mobility, simulation, training simulation, usability, user-centered design.

1 Introduction

In this paper we investigate the concept of fidelity in the context of full-scale laboratory simulations and usability evaluations of mobile ICT for hospitals. Fidelity, as it is commonly understood in HCI literature, is the extent to which a computer application or system reproduces visual appearance, interaction style and functionalities during evaluation [1]. In conventional desktop-based usability testing the physical and social aspects of the use situation are of little concern for evaluators. Interaction with desktop-based systems is highly uniform and static with regard to the physical and social aspects of the use situation—one user sitting in front of a PC, with his or her attention directed mainly on the computer screen.

J.A. Jacko (Ed.): Human-Computer Interaction, Part I, HCII 2009, LNCS 5610, pp. 232–241, 2009.
© Springer-Verlag Berlin Heidelberg 2009

As human-computer interaction moves "beyond the desktop" and into highly dynamic work settings, such as hospitals, the old standards for usability testing arguably no longer hold water. Work situations in hospitals often involve mobility and bodily work [2, 3]. This makes the usability of mobile ICT more subject to external factors that are not related to the GUI and the software being evaluated as such. These are factors that fall beyond what evaluations conducted in conventional usability laboratories can reveal.

We have attempted to meet these challenges by means of full-scale laboratory simulations—Simulated, natural-like hospital environments, in which nurses and physicians act out clinical scenarios using mock-ups or prototype systems. Such an approach arguably raises the need for evaluators to think about the fidelity, or level of detail, of the research setting. A critical issue for applying simulations as a usability evaluation methodology in a cost-effective way is deciding on the right level of fidelity.

This paper aims first to show that fidelity in usability evaluations of mobile ICT is a concept that extends beyond the software prototype being evaluated. Particularly, when addressing hospital settings, where the technology is likely to be used as part of work activities requiring manual labor with hands and feet in addition to high situational awareness, the physical environment and the work tasks become vital components of the total system being simulated.

The second objective of this paper is to demonstrate how the test environment, the prototype, and the test scenario can be adjusted to achieve targeted evaluations and induce behavior among participant that is desirable in terms of informing specific design aspects relevant for the usability of mobile systems in hospitals.

2 Methodological Motivation

The work conditions that ICT supporting clinical care is used in are very different than those for office settings and desktop-based computer interaction. The call for mobile ICT in hospitals essentially stems from the distributed nature of clinical work, and rapidly changing work situations. In order to conduct valid usability assessments of mobile interactive technology for such environments, then, the design solutions must be evaluated in a relevant use context [4]. Consequently, aspects that are characteristic for clinical work, such as mobility, clinician-patient interaction, and frequent context shifts, must be reflected in the setting in which the evaluation takes place.

Conducting usability evaluations in actual clinical situations is challenging. Firstly, the hospital is a high-risk environment, in which it can be critical to avoid affecting ongoing work. Secondly, patient information confidentiality is likely to prevent video and audio recording. At the same time, conventional laboratories intended for controlled desktop-based usability evaluations are unsuited for reconstructing the rapidly changing conditions of hospital work.

The combined need for a realistic, yet controllable, research setting, has motivated us to build a customized full-scale model of a hospital ward section, with advanced video and audio recording facilities.

3 Simulation Fidelity Principles

HCI literature in general provide little practical guidance on how to compose full-scale simulations with the aim of evaluating usability of mobile ICT. This has motivated us to study literature on training simulations in search of guiding principles. Given the relatively long history of training simulators, we will first highlight some of the central principles described in related research literature. Next, we will provide a brief overview of studies within the field of mobile HCI, where simulations have been employed to evaluate prototypes and early concepts.

3.1 Simulations Applied for Training Purposes

Within high-risk industries such as aviation, naval shipping, health care, and nuclear power production there has been a long tradition for using simulations for training purposes in risk-free environments. Obviously, the objective of training simulations and simulations for usability purposes differ. Simulations applied in the context of usability assessment aim at gathering data about the effectiveness, efficiency, and user satisfaction of a product used by specific users in a realistic situation. The focus is on product performance. Training simulations, on the other hand, are typically used for educational purposes and for maintaining or enhancing human work-related skills. In this sense, they are human-centric rather than product-centric.

Despite these differences in focus, many of the concepts developed from research on training simulations are highly relevant when designing simulation-based usability evaluations. Among the most central concepts described in research literature on training simulations are *equipment fidelity*, *environment fidelity*, and *psychological fidelity* [5].

Equipment fidelity refers to the extent to which the appearance and feel of real tools, devices, or systems that simulation participants operate on are duplicated. For example, aircraft cockpit procedures have been trained both with hi-fi representations of aircraft instruments and with lo-fi mock-ups [6].

Environment fidelity concerns the extent to which physical characteristics of the real-world environment (beyond the training devices) are realistically represented in the simulation. High-fidelity aircraft simulators are full-size replica of cockpits that duplicate of the operational aircraft environment and motions to great detail [7]. In flight training environments of less fidelity (e.g., desktop evaluation environments), visual and motional cues are typically reduced or lacking.

Lastly, psychological fidelity, relates to the realism of the simulation as perceived by its participants. In other words, it is extent to which participants are able to engage in the simulated situation, as they would have done in the natural setting. This is intimately dependent on the psychological demands the simulated tasks place on the participants. Human perception, attention, decision-making, memory, and action are factors that may influence psychological fidelity [8, p. 420]. Developing scenarios that replicate task demands of the real-world system is a common technique for enhancing psychological fidelity [5].

In training simulations, psychological fidelity is often considered the most important, because it is the most relevant attribute for learning. Prophet and Boyd [6] found that the transfer of training was equal for students practicing ground cockpit procedures in real airplanes, and student practicing the same routines on lo-fi representations of the relevant devices.

Each of the three components described above can be set along a continuum ranging from low to high fidelity (Fig. 1). As pointed out by Beaubien and Baker [5], the level of fidelity for each component typically depends on the purpose of the training simulation. For example, low-fidelity role-plays have been used to train teamwork related attitudes and skills, while simulations of higher fidelity are required if the goal is to learn the specific consequences of actions. Taken together equipment, environment, and psychological fidelity form the overall simulation fidelity (Fig. 1).

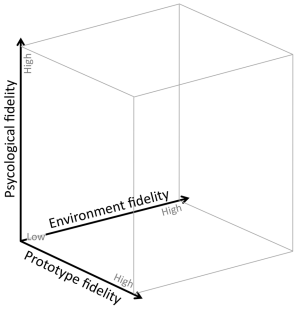

Fig. 1. Three interrelated simulation fidelity components. Each component can be set along a continuum ranging from low to high fidelity.

3.2 Mapping the Simulation Fidelity Concepts to HCI Terminology

In the context of HCI and usability assessments we consider equipment fidelity to be the equivalent of computer system or application fidelity (we will refer to this as prototype fidelity). As previously noted, this component encompasses physical appearance, interaction style, and functionalities. Environmental fidelity is the realism of the physical use setting (i.e., the point of interaction), while psychological fidelity corresponds to the user-perceived realism of the tasks they are given as part of a test.

3.3 Simulations Applied in Mobile HCI

The use of simulations in mobile HCI is in many ways a result of the recognition that conventional usability laboratories and testing do not duplicate factors affecting mobile usability [9, 10]. It can also been considered an efficient approach to overcome many of the challenges related to studying mobile systems in the field [11]. Examples of usability studies in which contextual features have been simulated in laboratory settings are decribed in work by Bohnenberger et al. [12], Pirhonen et al. [13], and Kjeldskov and Skov [14].

Some simulation-based usability studies of relevance to mobile ICT in hospital settings can be found in Refs. [4, 15, 16].

4 Applying the Fidelity Principles in Practice

In the current section we will provide examples from conducted simulations, showing how we have attempted to carefully adjust simulation fidelity to promote reflection

among participants regarding specific design aspects. Our examples also highlight the close interrelationship between the various simulation fidelity components presented earlier.

4.1 Case Study: Point-of-Care Scenarios

There are many hospital scenarios in which mobile ICT may prove helpful [2]. We have concentrated on a limited set of situations that have been found appropriate for usability testing in our full-scale hospital ward model. In particular, we have focused on situations where nurses and physicians are located at the patient bedside, i.e., at the point of care. Examples of hospital routines were such situations occur include ward rounds, during administration of medicine, and in response to patient calls. These situations form suitable test candidates for mobile ICT because they occur frequently in hospital wards, require mobility, and involve personnel who need quick and effortless access to patient related information.

4.2 Environment Fidelity

As previously pointed out, the circumstances under which mobile ICT supporting hospital work are used are radically different than that for office work and conventional desktop computer interaction. Human-computer interaction in clinical settings is more physical in nature both in the sense that hospital workers are using them while on the move, and in situations requiring physical interaction with patients (e.g., assistance and examination).

We have observed that to capture physical and bodily usability aspects of mobile ICT used at point of care, prototype solutions must be evaluated in research environments that closely mimic the physical environment of real patient rooms. For example, realistically propositioned rooms furnished with patient beds makes it possible for participants to move naturally in the model both around and between patient beds. For simulations addressing the usability of point-of-care systems this is essential, because it can help give participants a realistic impression of how well different solutions are physically adapted to the care situations. Examples of physical design factors, which according to simulation participants can increase the usability of mobile ICT at point of care, include the possibility easily to share screen content with patients (Fig. 2, left) and the opportunity to have digital media ready at hand to be to be used and put aside depending on what the immediate care situation calls for (Fig. 2, right). Both examples illustrate the intimate relationship between the physical environment and the physical placement and form factors of interactive devices.

Fig. 2. Examples showing the close relationship between the physical environment and the physical placement (left) and form factors (right) of digital media

There are often subtle details that need to be in place to capture physical and bodily usability factors. For example, to form a realistic impression of how well an interaction design solution accommodates the dialogue between clinicians and in-bed patients, patients need to be represented by human actors in real hospital beds (Fig. 2, left). Likewise, simulation participants need to wear their daily work uniforms, to bring forth or temporarily put aside digital media (Fig. 2, right).

4.3 Prototype Fidelity

The Spectrum of Prototype Fidelity. Prototypes can generally be divided into low, medium, or high fidelity. In low-fidelity prototyping, props (e.g., foam and cardboard models, paper and post-it notes, etc.) are often used as physical representations of interactive devices with rough sketches of envisioned graphical interfaces. Medium-fidelity prototypes are functional (computer-based) models of systems. They generally have simplified GUIs, but little or no functionality behind the GUI elements. High-fidelity prototypes are sophisticated and functional versions of envisioned designs. They may show sample information content.

Prototype Fidelity for Point-of-Care Scenarios. In our studies of mobile ICT applied in hospital settings we have used prototypes of different fidelities depending on design phase. As part of specifying user requirements for mobile ICT for point-of-care usage, we have conducted simulations using lo-fi props in realistic models of patient rooms [17]. The main rational for this is to put focus on the context in which the technology will be used, rather than details concerning GUIs and software functionality. By applying mock-ups one can avoid restricting reflection among participants by committing to particular hardware and interfaces.

In simulations with functional prototypes, we have attempted to be more particular about design aspects we wanted to address. For these simulations we have applied *mixed fidelity prototypes* [18], i.e., models that combine different fidelities with regard to GUIs, functionalities, interaction styles, and information content. For example, in simulations addressing the usability of location-based access to medical information at the point of care (the full study is described in Ref. [16]) we found it useful

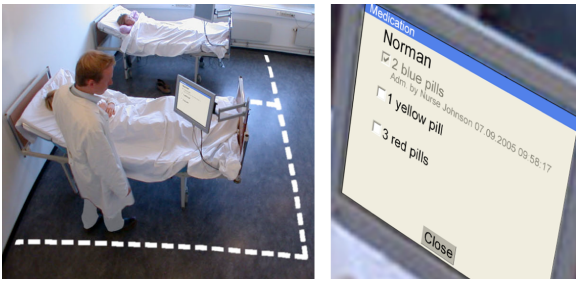

Fig. 3. Mixed fidelity prototype providing location-based access to a patient's medical record. High-fidelity sensors and radio tags detect the physical position of a test participant and near patients. The graphical representation of the medical record, which is automatically retrieved and presented on the bedside terminal, is of low fidelity.

to implement the interaction techniques using high-fidelity hardware and sensors to give participants a realistic impression of actual use. At the same time, we deliberately kept the details of the GUI at a low level and avoided to link the prototype to realistic medical sample data. The main rational for this delimitation is that we primarily wanted to help participant focus on and give feedback regarding the usability of interaction styles, rather than GUI related aspects of the design. Fig. 3 shows the mixed fidelity prototype that we applied in the example described above.

4.4 Psychological Fidelity

This section gives a brief summary of different techniques we have employed to increase the psychological fidelity of our simulations.

Domain Expertise. As pointed out above, developing scenarios that mimic the task demands of the real-world system can increase psychological fidelity. To facilitate this, and make sure that the simulations reflected a sufficient degree of realism, the scenarios were designed with assistance from domain experts (physicians and nurses).

Baseline Scenarios. To help participants relate the prototypes and evaluated concepts to their everyday workday, we have learned that running an introductory baseline scenario reflecting current paper-based practices are useful (Fig. 4). These "as-is" scenarios can effectively act as a reference or benchmark for the participants when they later act out the same scenario using functional prototypes. This help participants "make sense" of scenarios involving information media they are not familiar with.

Fig. 4. Baseline scenarios reflecting current paper-based practices (left) can act as references for test subjects when they later try out digital solutions for the same purposes (right)

Targeted Simulations. Because the conducted simulations mainly have focused on evaluating early concepts and partial prototypes, we have tried to tailor the scenarios to help promote feedback on the particular design aspects. In some cases, this has resulted in delimitations that, depending on the purpose of the simulation, possibly might reduce psychological fidelity. Using only mock-up representations of electronic medical records in some of the simulations, as explained above, is an example of such a limitation of scope. This, however, we consider necessary compromises to achieve targeted evaluations. Feedback from the participants indicated that they found the simulations realistic in spite of such simplifications.

Scripted Simulations. We have also experimented with ways to increase the participants' perceived realism of the simulations by trying to integrate their professional experience. One technique we have applied is to use patient actors instructed to reveal

certain pieces of information during the scenario, but leave it for the participants (i.e., clinicians) to decide how to act on that information [19]. For other types of simulations we have found it sufficiently to give the patient actors a more passive role, e.g., acting as physical markers in the simulations.

5 Setting the Scene Right

A fundamental dilemma related to simulations, whether applied for usability assessment or training purposes, is specifying a priori a sufficient fidelity level. Typically, approximating mobile use contexts involves trade-offs between control over the research setting, realism, and available resources [4]. In the current section we will present and discuss some guiding principles for "setting the scene right" for simulation-based usability evaluations of mobile ICT for hospitals.

5.1 Psychological Fidelity First

In Sect. 3.1 we pointed out that psychological fidelity is often considered the most significant fidelity component in training simulations. This is because it represents the transfer of skills learnt in the simulated setting back into the real world. Based on our experiments, we argue that the same component is also the most critical when designing full-scale simulations for usability assessment purposes, but for different reasons. We see psychological fidelity first and foremost as a key premise for provoking reflection among simulation participants. This is especially valuable in early phases of the design process before critical design choices are taken, and when end user feedback is most likely to help inform the actual design. In order to motivate reflection among simulation participants it is essential that they see the on-the-job relevance of the simulation. As discussed in Sect. 4, a realistic test environment and prototypes with certain functional features can help evoke the central psychological mechanism triggered in everyday clinical work.

5.2 How Much Fidelity Is Enough?

In our full-scale simulations we have attempted to follow a "just enough" principle with regard to the amount of realism reflected. That is, duplicating aspects that we, along with domain experts, consider most likely to affect the perceived usability of the prototypes that are to be tested. This, as we have pointed out earlier, depends closely on the object of the simulation. Because the usability of mobile ICT for hospitals are highly contextual, there is no "one size fits all" approach when it comes to fidelity of simulation-based usability evaluations.

6 Summary and Conclusions

In this paper we have investigated the role of fidelity in full-scale laboratory simulations used for usability assessment of mobile ICT for hospitals. Drawing on training simulation research, we identified three components of relevance—Environment fidelity, equipment or prototype fidelity, and psychological fidelity. We have shown

by examples from practical simulations how the different components can be modified to match the focus of different evaluations.

The key principles the current paper has suggested regarding fidelity of simulation-based usability assessments of mobile ICT for hospitals are as follows:

- Simulations need to be specific about the design aspects that are being evaluated. Cost-effective and targeted simulations should replicate features of the mobile ICT solution, physical environment, and work tasks, that are considered relevant for the design aspect one wants to gather feedback on.
- There is no direct correlation between the overall fidelity of experimental simulations and effectiveness in terms of informing design. Realistic prototypes and work environments may enhance the perceived realism of the simulation, but this does not guarantee that feedback from participant are valuable in terms of informing design.
- Simulations for usability assessment purposes should prioritize psychological fidelity. This component is most relevant for provoking reflection in design among simulation participants. To provoke such reflections it essential that participants are able to relate the design concept to their everyday work.
- The requirements to simulation fidelity will typically increase as the mobile ICT solution is developed.

We expect that future simulations in our full-scale ward model will enable us to further explore the role fidelity plays in simulation-based usability evaluations of mobile ICT for clinical work.

Acknowledgements

The current work has been supported in part by the Norwegian Research Council by grant 176761 (POCMAP) of the VerdIKT program, DIPS ASA, The Industrial Research Fund for NTNU, St. Olav University Hospital, Akershus University Hospital, NTNU, and Telenor R&I.

References

1. Rudd, J., Stern, K., Isensee, S.: Low vs. high-fidelity prototyping debate. Interactions 3, 76–85 (1996)
2. Sørby, I.D., Melby, L., Nytrø, Ø.: Characterising cooperation in the ward: framework for producing requirements to mobile electronic healthcare records. International Journal of Healthcare Technology and Management 7, 506–521 (2006)
3. Bardram, J.E., Bossen, C.: Mobility Work: The Spatial Dimension of Collaboration at a Hospital. Computer Supported Cooperative Work 14, 131–160 (2005)
4. Kjeldskov, J., Skov, M.B., Als, B.S., Høegh, R.T.: Is It Worth the Hassle? Exploring the Added Value of Evaluating the Usability of Context-Aware Mobile Systems in the Field. In: Mobile HCI 2004, pp. 61–73 (2004)
5. Beaubien, J.M., Baker, D.P.: The use of simulation for training teamwork skills in health care: how low can you go? Qual Saf Health Care 13 (2004)

6. Prophet, W.W., Boyd, H.A.: Device-Task Fidelity and Transfer of Training: Aircraft Cockpit Procedures Training. Tech. Report. Human Resources Research Organization, Alexandria, VA (1970)
7. Rehmann, A., Mitman, R., Reynolds, M.: A handbook of flight simulation fidelity requirements for human factors research. Tech. Report No. DOT/FAA/CT-TN95/46. Wright-Patterson AFB, OH: Crew Systems Ergonomics Information Analysis Center (1995)
8. Patrick, J.: Training. In: Tsang, P.S., Vidulich, M.A. (eds.) Principles and Practice of Aviation Psychology, CRC Press, Boca Raton (2002)
9. Johnson, P.: Usability and mobility; Interactions on the move. In: Proceedings of the First Workshop on Human-Computer Interaction with Mobile Devices (1998)
10. Graham, R., Carter, C.: Comparison of speech input and manual control of in-car devices while on-the-move. In: Mobile HCI 1999 (1999)
11. Pascoe, J., Ryan, N., Morse, D.: Using while moving: HCI issues in fieldwork environments. Transactions on Computer-Human Interaction 7, 417–437 (2000)
12. Bohnenberger, T., Jameson, A., Krüger, A., Butz, A.: Location-aware shopping assistance: Evaluation of a decision-theoretic approach. In: Paternó, F. (ed.) Mobile HCI 2002. LNCS, vol. 2411, pp. 155–169. Springer, Heidelberg (2002)
13. Pirhonen, A., Brewster, S., Holguin, C.: Gestural and Audio Metaphors as a Means of Control for Mobile Devices. In: Proceedings of the SIGCHI conference on Human factors in computing systems (2002)
14. Kjeldskov, J., Skov, M.B.: Creating Realistic Laboratory Settings: Comparative Studies of Three Think-Aloud Usability Evaluations of a Mobile System. In: Interact 2003, pp. 663–670. IOS Press, Amsterdam (2003)
15. Alsos, O.A., Svanæs, D.: Interaction techniques for using handhelds and PCs together in a clinical setting. In: NordCHI, Oslo, Norway, pp. 125–134. ACM, New York (2006)
16. Dahl, Y., Svanæs, D.: A comparison of location and token-based interaction techniques for point-of-care access to medical information. Personal and Ubiquitous Computing 12, 459–478 (2008)
17. Svanæs, D., Seland, G.: Putting the users center stage: role playing and low-fi prototyping enable end users to design mobile systems. In: Proceedings of the SIGCHI conference on Human factors in computing systems, Vienna, Austria, ACM, New York (2004)
18. Petrie, J.N., Schneider, K.A.: Mixed-fidelity prototyping of user interfaces. In: Doherty, G., Blandford, A. (eds.) DSVIS 2006. LNCS, vol. 4323, pp. 199–212. Springer, Heidelberg (2007)
19. Alsos, O.A., Dabelow, B.: Stylus, Finger, or Buttons? A Comparative Evaluation Study of Interaction Techniques for PDAs in Point-of-Care Situations (submitted manuscript)