A Mixed-Fidelity Prototyping Tool for Mobile Devices

Marco de Sá, Luís Carriço, Luís Duarte, Tiago Reis
LaSIGE & Department of Informatics
Faculty of Sciences, University of Lisboa
Campo Grande, 1749-016, Lisboa, Portugal
{marcosa,lmc}@di.fc.ul.pt, {lduarte,treis}@lasige.di.fc.ul.pt

ABSTRACT

In this paper we present a software framework which supports the construction of mixed-fidelity (from sketch-based to software) prototypes for mobile devices. The framework is available for desktop computers and mobile devices (e.g., PDAs, Smartphones). It operates with low-fidelity sketch based prototypes or mid to high-fidelity prototypes with some range of functionality, providing several dimensions of customization (e.g., visual components, audio/video files, navigation, behavior) and targeting specific usability concerns. Furthermore, it allows designers and users to test the prototypes on actual devices, gathering usage information, both passively (e.g., logging) and actively (e.g., questionnaires/Experience Sampling). Overall, it conveys common prototyping procedures with effective data gathering methods that can be used on ubiquitous scenarios supporting in-situ prototyping and participatory design on-the-go. We address the framework's features and its contributions to the design and evaluation of applications for mobile devices and the field of mobile interaction design, presenting real-life case studies and results.

Categories and Subject Descriptors

H5.2. [Information interfaces and presentation] (e.g., HCI): User Interfaces–*Evaluation, Prototyping, User-centered Design.*

General Terms

Design, Experimentation, Human Factors.

Keywords

Mobile Interaction Design, Prototyping, Usability, Evaluation.

1. INTRODUCTION

Designing for mobile devices is an increasingly demanding challenge. Besides the hardware constraints that are imposed by their size, interaction modalities, diversity and portability, their pervasiveness and multi-purpose functionality imply an entire new set of usage paradigms.

As a consequence, new design approaches are required, particularly for the evaluation and prototyping phases. The absence of specific methods and techniques is patent [14], which leads to none or to incomplete, and definitely inadequate, evaluations. In fact, these stages are usually supported by common methods which are impracticable or not suited to mobile scenarios, generally neglecting their ubiquitous nature.

During the design of a few applications directed to PDAs we were faced with several recurrent problems: (1) the prototyping techniques found in the literature and commonly used for desktop applications were inadequate to the ubiquitous nature of our applications; (2) prototypes started to mislead users due to the used material hindering, at times, their participation on the process and were limited regarding some of our goals; (3) lab experiences proved to be insufficient while determining usability issues with the developed prototypes; (4) techniques such as the Wizard of Oz or direct observation posed restraints to our evaluation since they were extremely difficult to apply on real world settings and (6) methods such as Experience Sampling Method (ESM) [5] or diary studies required extra effort and, although providing qualitative data, did not cover the interaction details that we wanted to evaluate.

These difficulties propelled the adoption of alternative techniques and experiences that brought out very positive results [23]. One of the main contributions that came about from this process was the integration of several functionalities and tools into a specific framework that entangles various techniques for mobile prototyping and evaluation purposes. The framework takes into account previous work within this area, the lessons that we learned and introduces new contributions that foster participatory and continuous in-situ design supporting designers, and the design's evolution, through the initial stages of user-centered design of mobile applications.

We start by addressing the existing work in this area. Afterwards we describe our tool's concept, its goals and novel contributions also detailing its architecture and features. We then present some already achieved results and delineate future work directions.

2. MOTIVATION/RELATED WORK

Design methods and techniques for mobile devices, albeit being recent and somewhat immature fields of research are increasingly being addressed by researchers, leading to the appearance of different approaches for a wide range of problems [13]. Unsurprisingly, given their differences from desktop systems, most efforts have been directed towards prototyping and evaluation, with some references also pointing to the generation of UI design guidelines specific for small screens [2].

Regarding prototyping, new techniques and orientations, particularly for low-fidelity prototypes, have been introduced [23]. These suggest the need for more detailed and carefully built

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
AVI'08, May 28-30, 2008, Napoli, Italy
Copyright 2008 ACM 1-978-60558-141-5…$5.00.

prototypes that offer a more resembling picture of final solutions and their characteristics [10]. In fact, the adopted prototyping technique can be determinant during the consequent evaluation stages, allowing users to freely interact with them, improve them and use them on realistic settings without misleading users [23]. Furthermore, to assert on various details that might be relevant at different stages of prototyping, the concept of mixed prototyping has emphasized the need to create different prototypes to evaluate different dimensions of usability [20]. On these aspects, prototyping tools can play a paramount role, allowing designers to maintain their sketching and writing practices while creating prototypes that can actually run giving users a more tangible and realistic feel of the future application.

DENIM [19] and SILK [17] are two prototyping tools that give designers the ability to quickly create sketch-based prototypes and interact with them on the computer, also including the possibility of replacing drawn components with actual programmatic components. More recently, systems such as SketchWizard [6] or SUEDE [16] have also emerged, supporting new modalities and interaction modes such as pen-based input on the former and speech user-interfaces on the latter. Ex-A-Sketch [9] also allows designers to quickly animate sketches drawn on a whiteboard. On a different level, the BrickRoad project [23] also supports the design of location-enhanced applications, especially during early design stages.

However, although these tools have useful functionalities and features, and provide sketching and quick prototyping mechanisms, the integration with the evaluation stages is rarely addressed. Moreover, the evolution from early based sketches to more advanced prototypes is only present on SILK and DENIM which lack crucial components (e.g., sound) and active behavior or are deeply focused on specific domains. Furthermore, none addresses the specific needs of mobile devices or provides usability guidelines and aids to designers while creating their prototypes. Nevertheless, the automatic support for Wizard-of-Oz prototypes and the ability to animate hand drawn sketches has shown very positive results.

As aforementioned, problems are felt again when evaluating the developed prototypes. Although some recent studies reflect an increasing amount of attention towards contextual evaluation, out of the lab, its relative inexistence contrasts with the importance and benefits it presents to mobile devices [7],[21]. Existing examples usually point guidelines on how to emulate real world settings within labs [1],[15] or provide solutions [5] that are useful as a complement but, even if obtaining positive results, do not address specific usability problems, do not provide quantitative data and focus mainly on user satisfaction. Furthermore, they show little regarding user interaction towards the applications.

Some recent approaches have also addressed this stage of design, focusing methods to gather usage data remotely through active – requiring user intervention - (e.g., ESM, Diary Studies) and passive modes – without user intervention - (e.g., Logging). For instance, with close goals to our framework regarding evaluation, the Momento [4], and the MyExperience [8] systems provide support for remote data gathering. The first relies on text messaging and media messaging to distribute data. It gathers usage information and prompts questionnaires as required, sending them to a server where an experimenter manages the received data through a desktop GUI. On the second, user activities on Mobile Phones are logged and stored on the device. These are then synchronized depending on connection availability. The logging mechanism detects several events and active evaluation techniques can be triggered according to contextual settings.

However, and although some goals or used techniques are similar, our approach intends to provide qualitative and quantitative information that can be easily understood by non-expert users, focusing on interactions that directly relate to the developed prototypes on very early stages. Our goal here is to integrate the prototyping and evaluation stages seamlessly, facilitating user involvement and the design process. Moreover, none of these approaches integrates the prototyping and evaluation on real devices, also including means adjust the prototypes while evaluating them or to analyze them (e.g., various alternatives to one user interface), individually or simultaneously, on an easy-to-read video-like mode. Furthermore, most depend on server-client architectures, requiring a constant connection or frequent synchronizations. Still, these systems and recently conducted experiments [12],[18] validate the need to undertake evaluation on real-life settings using both passive and active data gathering techniques, even at a very early design stage.

3. CONCEPT, GOALS AND FEATURES

To cope with early design stage difficulties, which pertain both to prototyping and consequent evaluation, the developed prototyping framework's features cover both these stages, supporting an iterative and participatory design that facilitates the transition between them.

Its umbrella goal is to support the early design stages of applications for mobile devices. Like some of the aforementioned frameworks [16],[19] it provides designers with tools to quickly create prototypes and evaluate them, focusing specifically mobile and handheld devices. It supports in-situ and participatory design and enables designers to use both passive and active evaluation methods. The framework allows the construction of low, mid and high-fidelity prototypes and extends its automatic Wizard of Oz usage through their evaluation, also providing means to analyze the gathered data.

More concisely, on the prototyping stages we aim at: (a) supporting a visual, quick and easy design of realistic mobile prototypes, with flexibility regarding their fidelity (b) offering expert users or users without any programming knowledge the possibility of building or adjusting their prototypes; (c) allowing and promoting participatory design and prototyping during outdoor evaluation sessions within realistic settings.

For the evaluation stage our goals are: (a) retrieving reliable usage information without intrusive equipment, without the designer or usability engineer's presence and using seamless/passive techniques; (b) supporting the analysis of usage patterns and usability concerns through the visualization of the user's activities and (c) the integration of methods such as probing [11], ESM [5] and diary studies extending the scope of the evaluation process.

Our main contributions over previous work are the convergence of prototyping and evaluation techniques into one end-user tool, supporting several degrees of fidelity, allowing the comparison of design alternatives, and suggesting new ones if available, facilitating the detection of usability problems or design flaws on early design stages. This coverage is extended to their evaluation on various stages of design within the context in which they are most likely to be used (e.g., device, location, environment), always centering its procedures on the user. Within these, the framework also supports in-situ participatory design, directly on the targeted devices. Globally, this can be achieved through the following features:

1. Prototyping with mixed-fidelities (thus analyzing different usability dimensions). Hand drawn sketches or interactive visual pre-programmed components can compose different prototypes with varied levels of visual refinement, depth of functionality and richness of interactivity [20], comparing different design alternatives, evaluating button sizes, screen arrangements, element placement, interaction types, navigation schemes, audio icons, and interaction modalities, among others.

2. Direct prototyping on the mobile devices. Users are able to update prototypes on mobile devices, re-arranging simple details and improving the prototypes during evaluation sessions on real settings, out of the lab. The overall building mechanism is simple, visual or wizard based allowing experienced designers or inexperienced final users to adjust their own prototypes. This enables its usage for probing purposes [11], promoting experimentation and on-the-fly design of new solutions for and on the context in which the user is interacting with the tool.

3. Integrated usability guidelines for mobile devices on mid-fidelity prototypes. When prototypes are created using the visual primitives and components, usability guidelines can be automatically enforced, if chosen by the designer/user. For instance, the location of each component, the actual size of the component or even the amount of information per screen can be automatically arranged. These guidelines are configurable and can be domain oriented (e.g., e-health – special icons, education - limited content, media players). The framework is also able to provide alternative versions of the created prototypes (e.g., displaying a similar prototype that uses radio buttons instead of a combo-box).

4. Avoid cargo cult syndrome [10]. By using actual devices, problems regarding the device's characteristics (e.g., size, weight, screen resolution, shape) emulation are solved, allowing their utilization on realistic settings. This provides users a much more tangible and realistic usage experience.

5. Automatically support the Wizard-of-Oz technique. By adding behavior to the digitalized sketches or by using visual components, users can navigate through the prototype without having to explicitly replace the screens by hand or without the presence of a designer to do so.

6. Gather data through passive and active techniques. On the former, every action that the user takes is automatically logged with customized granularities. On the latter, the use of ESM and diary studies, integrated within the tool, provides another source of data and usability information. Integrated questionnaires can be popped during or immediately after using the prototype, or even automatically during the day according to specific settings (e.g., if the user is unable to achieve a specific goal or is continuously failing to press a small button).

7. The framework also includes a log player which re-enacts (through a video-like mode) all the users' activities with accurate timing and interaction details, attenuating the need for direct observation.

4. FRAMEWORK/ IMPLEMENTATION

In order to support the aforementioned functionalities, the prototyping framework is divided into several tools.

4.1 Prototype Building Tools

The first tool, the prototype builder is divided into two modes. The first is a wizard-based user interface that guides users to create a prototype screen by screen. It supports the definition of the prototypes' fidelity, degree of functionality and behavior (Figure 1). It allows users to create each screen individually, organizing them sequentially and customizing them according to their needs. The second mode is the advanced mode. Here, designers can easily drag and drop the selected components, use hand drawn sketches, pictures or images for multiple screens, also arranging the "prototype's wireframe" or storyboard (Figure 2).

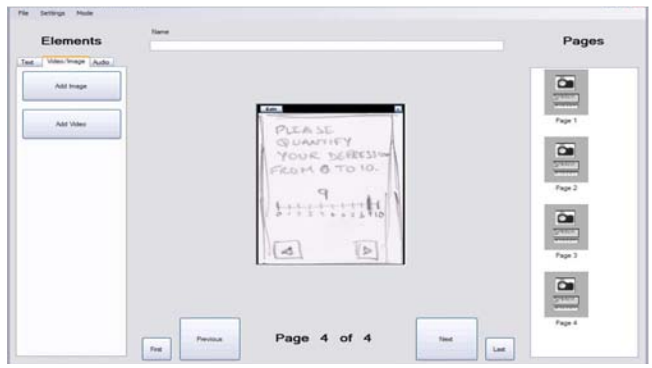

Figure 1. Prototype building tool – Wizard Mode.

On both modes, interactive output components (e.g., combo-boxes, labels, images and audio files) can be used. Input components (e.g., text data entries, sound recorder, and video recorder) are also available. These components are used to create mid to high-fidelity prototypes. For low-fi prototypes, sketches (hand-drawn and scanned or digital drawings) can be easily imported and their behavior adjusted. For each component, different configurations are also available (e.g., multiple-choices through radio buttons or combo-boxes).

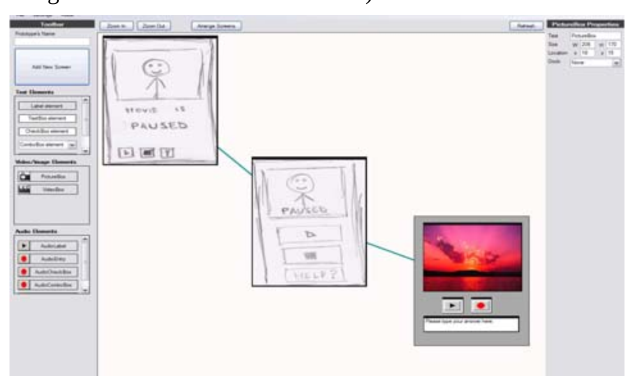

Figure 2. Prototype building tool – Advanced Mode.

The prototype's behavior can be defined within three levels: a component/element, a screen behavior and a global behavior. On the first users can define the behavior when using an individual component (e.g., a button press displays a warning). The second defines the behavior for the entire screen (e.g., the user missed two of the screen's components and these are highlighted) and the third for the entire prototype (e.g., a questionnaire is popped once the user reached the fifth screen).

Each prototype is specified in XML and stored within a file that contains its specification which can be transferred and updated even without using any specific tool. The tool's modularity allows different components to run on different devices and systems.

4.2 Runtime Environment and Logging

The counterpart of the previous tool is the runtime environment. This tool is responsible for materializing the prototypes on the targeted device. Currently we have a runtime environment for Windows Mobile, Palm OS and SymbianOS. It is composed by a straightforward user interface that displays a list of the available prototypes for users to select. Once a prototype is chosen, it will be displayed and users can interact with.

The runtime environment also offers options to edit the prototype, save usage information at any given point or to define the granularity of the saved data. On the editing mode, every component's location, size and some content can be updated or changed. Screens and components can be deleted or their sequence arranged (e.g., card/screen-sorting and in-situ design).

Integrated within this runtime environment there is also the logging engine which stores every event. Events range from each tap on the screen, each button press or even each character that was typed by the user. Events are saved with a timestamp, allowing its reproduction for the re-enactment of the usage behavior. Other details such as the type of interaction, location of the screen tap, etc., are also stored for filtering purposes.

4.3 Analysis

The final tool pertains to the analysis of the logs generated by the logging engine. The log player resembles a "movie player" which re-enacts every action that took place while the user was interacting with the prototype. Several analysis granularities are provided ranging from each character that was typed to every visited screen. Pausing, stopping or adjusting the speed in which events are (re)played is also possible (e.g., fast-forward; double speed) through the options shown at the bottom of Figure 8.

The tools are available for desktop computers and, on a simpler version, for the abovementioned mobile platforms. The runtime environment is also available for desktop devices so that, if needed, designers can quickly review their prototypes before sending them to the mobile devices.

5. CREATING A PROTOTYPE

Following the traditional approach of low-fidelity prototyping, each prototype is composed by a set of screens (e.g., traditionally composed by paper cards). Each screen can be composed by a sketch, hand drawn and scanned to the computer or drawn using specific software. These are the lowest-fidelity prototypes where the screen is based solely on a digital version of a hand drawing made by the designer. Alternatively, as already mentioned, the framework includes visual components (e.g., drop-boxes, buttons, text-fields, track-bars, images, videos, sounds) that can be used to create a screen, alike the commonly used post-its. Screens are added as necessary and arranged on a storyboard to define their sequence (Figure 2).

At this stage, the degree and depth of functionality of the prototype can also be configured. If using a hand-drawn sketch for a low-fidelity prototype, "clickable" areas can be configured, generally over a drawn button or list. To do so, the designer visually drags a resizable rectangular area to the element he/she wants to make "clickable". Afterwards, these areas can be added with behavior (e.g., once they are clicked something happens).

On a higher fidelity prototype, the elements of the screen can be activated (e.g., drop-box contains a number of items, text-field receives an amount of characters) or de-activated (e.g., used solely for screen arrangement purposes). This allows us to test several dimensions. For instance, we can compare drop-boxes against lists or to text-fields or evaluate the location of each of these elements. Thus, it is possible to add or remove functionality to some degree or simply use the prototype for screen navigation, color, or button size tests.

Common components assume their traditional functionalities. Text-boxes allow text input; track-bars the selection of a numeric value, sound and video recorders record sound and video, etc.

5.1 Replacing the Wizard of Oz

Globally, the prototype's behavior is defined by selecting what buttons trigger the appearance of which screens, providing an automatic Wizard of Oz approach. Alternatively, these are arranged sequentially according to their location on the advanced mode, and their sequence on the wizard mode. Users can create warnings that can be popped up and shown according to specific triggers (e.g., selection from a drop-box or typing of a password). This mechanism is supported by three different types of rules.

The first ones are content-based rules, triggered when certain content, within a component, is chosen (e.g., if the user chooses yes or no from a list or a high or low value on a track-bar). Figure 3 on the left shows an initial PalmOS version of the mobile prototype builder. It depicts the adjustment of a rule that triggers a warning when the same answer is repeated 4 times.

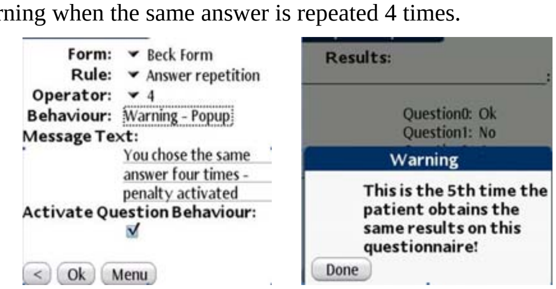

Figure 3. Rule definition and warning on a PalmOS PDA.

Time-based rules, on the other hand, are activated according to time limits (e.g., the user takes more than one minute to press a button or too long to answer a question within a questionnaire). Finally the interaction-based rules can be triggered according to the amount of taps on the screen, the location of those taps or the number of times a button is pressed.

To complement these rules and to function in concert with them, there are three types of behaviors. The first one is the "jump to" action. As the name indicates, once activated, it will automatically force a jump to a designated screen. For instance, using a "click area" that triggers a "jump to" behavior allows a user to configure an active hand drawn button, on a sketch-based prototype, to jump to the following, previous or any other screen/sketch once it is selected. This mechanism allows designers or users to define navigational constraints without writing code or programming, replacing the designer on his sketch and component removing/inserting activities (e.g., Wizard of Oz technique).

The second type of behavior is composed by warnings (Figure 3 – on the right). Popping a warning alerting the user that he/she selected the top value, or did not select any value from a track-bar is a simple example. The third type of rule hides or shows components (e.g., if the user selects an option from a combo-box, the correspondent data entry field is shown).

These rules and correspondent behaviors, when used together, allow designers to compose fairly elaborated prototypes. However, they still maintain the necessary simplicity to be easily specified by end-users as well, through a simple to use, selection-based wizard interface.

Prototype files can be dragged directly into the device or can be transferred automatically through the building tool, if a connected device is detected. Once on the device they can be directly used on the runtime environment.

5.2 Reviewing Logs and User Behavior

Since one of the main goals of mobile evaluation is to evaluate the users' behavior on real scenarios, we intended to replace, as far as possible, direct observation with a similar mechanism. Therefore, several visualization options for the usage logs are available (e.g., event lists, selection tables). However, the most interesting one presents an exact replica of the users' behavior, emulating the mobile device and re-enacting every tap on the screen, every typed character and so on (Figure 8). Although these logs are limited to the direct interaction that the user has with the device, they still present enough detail to compare different design choices, evaluating navigation options, component placement and size, audio icons, audio volume, synthesized-text, the prototypes' feasibility and other questions that designers face on the early design stages.

6. CASE STUDIES

We have used the prototyping framework to generate and evaluate a set of prototypes on two different domains. On the first, psychotherapy, we developed low and high-fidelity prototypes for several therapeutic tools [3]. On the second, education, teachers used the framework to create different elaborated prototypes [22].

6.1 Psychotherapy

The first case study involved a team of mobile interaction designers and a team composed by a group of cognitive behavioral researchers and practicing psychotherapists. The main goal was to continue an on-going project which aimed at the support of cognitive behavioral therapy through the use of mobile e-artifacts [3]. Given the highly ubiquitous tasks that are encompassed within such type of therapies and the critical domain of healthcare in which we were working, the introduction of the framework and its functionalities aimed at facilitating the expert team to participate on the process and the quick construction of prototypes that could be easily evaluated and tested by therapists and patients. Moreover, initial tests with paper prototypes were misleading therapists regarding usage possibilities resulting on a constant rejection of most of the design team's ideas. Therapists had difficulties imagining and materializing the end result based on sketches and paper-based prototypes.

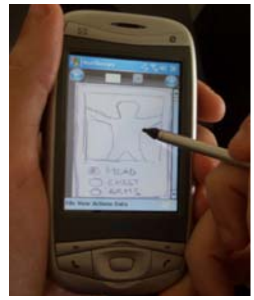

Figure 4. User interacting with a low-fi prototype for a pain therapy application on a SmartPhone.

Accordingly, several iterations of low-fidelity prototypes, that had been previously drawn were digitalized and used by the designers on the framework. "Click Areas" were defined and their behavior configured. These sketch-based prototypes were tested by therapists and some psychotherapy students, on smartphones (Figure 4) and later evolved to higher-fidelity ones, that allowed user input and reacted to usage behavior (Figure 5).

At this stage, therapists and researchers from the psychotherapy team started to use the framework as well, mainly to adjust the already developed prototypes (e.g., changing some interaction types). As the prototypes started to refine, the therapists handed the new versions to some students and used them on experimental therapy sessions within the research laboratory. Throughout this process, tools for anxiety, depression, pain therapy and associated disorders, were created and thoroughly evaluated.

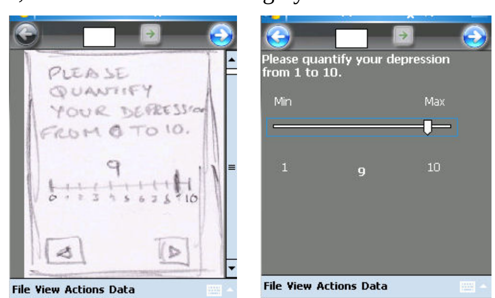

Figure 5. Left: Sketch-based low-fi prototype for a psychotherapy tool and its evolved high-fi software version.

Once most of the created prototypes had been experimented and adjusted in-situ by both therapists and researchers and on some experimental sessions, all the logs were carefully reviewed by both teams. Whereas the design team was focused on interaction details and on usability assessment, the therapists started to detect hesitations and symptomatic behaviors while users interacted with the prototypes. For instance, using the log player, therapists were able to detect questions where users spent more time or thoughts that were constantly written and deleted. Some of these behaviors led to the identification of critical subjects and to the detection of underlying problems that patients faced but did not mention during their face-to-face therapeutic sessions.

Overall, the prototypes were very well accepted and the tested versions, with some adjustments, were even used as final applications.

6.2 Mobile Learning

The second domain in which the framework was used was education [22]. In this case, the main goal was to design and evaluate a possible application for students to use, while at school or at home, to complete tests, homework, to review content provided by them. A team of 3 designers and one composed by 4 teachers were involved on the entire process as well as students for the final evaluation sessions.

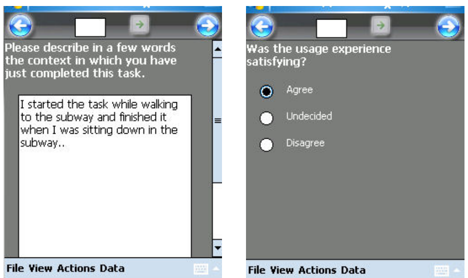

Figure 6. Questionnaire shown while using the prototype.

Teachers aimed at creating an easy to use tool that could convey the possibilities of assessment and task completion by students of various ages, as well as the access to relevant content that would be provided to students as necessary. The design process started with a set of meetings where requirements were established and ideas started to emerge, especially from the teachers' side. Given the successful experience with the previous case study, teachers were provided with the prototyping framework since the beginning, and a short tutorial (1 hour – wizard mode) was given to all the involved teachers. The functionalities were explained and the results gathered from the previous experience were described. Accordingly, teachers were given the framework and created a set of prototypes for applications that would allow students to achieve various different activities (e.g., watch a short movie, read a short book, complete tests or homework).

Given the functionalities that were explained, teachers started by creating low-fidelity prototypes for all the tools and were concerned mainly with the aesthetics, content organization, vocabulary and features of each tool. Once the prototypes were created, the design team conducted a series of evaluation sessions, with the teachers, in order to assess each of the low-fi prototypes for the targeted tasks.

The evaluation sessions were carefully planned, including a detailed description of the goals, the tasks that had to be performed, the student profiles that would be used and the locations and settings in which all the sessions would take place. In this last aspect, particular care was taken to select scenarios and settings with different conditions, regarding light, noise, user posture (e.g., walking, seating, etc) and the introduction of casual distractions (e.g., interrupting the user to ask a question, requesting the user to walk on a busy corridor, and so on). Overall, we tried to conduct the evaluation sessions on the most realistic settings possible. On some of the evaluation sessions, the students that tested these low-fi prototypes even took the devices and prototypes home with some pre-determined tasks to complete. In these situations, in order to have a glimpse of the context of use, specific questionnaires were included on the prototypes and automatically shown while using the prototype (Figure 6).

Once all the initial evaluation sessions, with the low-fi prototypes were complete, both teachers and designers started to look at the logs. Results from this process were naturally taken into account on the higher-fidelity versions of the prototypes.

On the following design cycle, teachers and designers continued to collaborate and began to create high-fidelity prototypes for the same tools. Based on the low-fi prototypes and using the available components, teachers replaced sketches with pictorial based tools for younger children and more textual (e.g., track-bars, textboxes) prototypes for teenagers or adults. Again, after a set of prototypes for each tool was created, evaluation sessions were conducted.

Half of the total of 6 tests, involving 36 students, took place at the university campus while the rest was done at various locations, including students' homes. The campus tests were filmed using a low-cost mobile kit developed specifically for this purpose (Figure 7). The initial kit used a shoulder camera. However, this approach, although capturing the user interaction with the device, provided little information regarding the context and the user's interest points. Various mobile devices (e.g., with and without keyboards were handed to the students). On one of the selected tasks, students were required to complete a test at school and another at home. Students used the prototypes to respond to tests and were free to use the devices to whatever they wished. After the tests were completed and usability questionnaires responded, students returned the devices to teachers and logs started to be analyzed, together with all the footage that was captured.

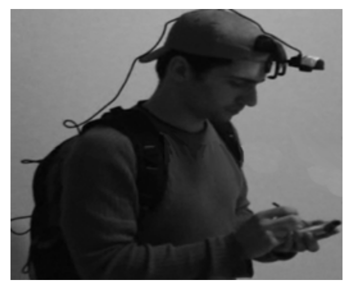

Figure 7. Mobile Video Capturing Kit.

The design team quickly detected some problems with the prototypes, particularly referring to the selected interaction modalities and the locations in which the prototypes were used. For instance, track-bars and text-boxes were difficult to use while walking or on the bus/subway whereas lists and multiple-choices (e.g., radio buttons) were easily handled. While for text-boxes this was already expected since keyboards, either physical or virtual, have to be used, for track-bars this came as a surprise.

From an educational point of view, and based on the suggestions made by the designers, teachers also tried to detect student difficulties while using the prototypes. To achieve so, and to isolate difficulties that could pertain to the components, or to the user interface itself, teachers connected three different aspects to detect learning issues.

Accordingly, to identify possible problems, it was necessary to search for questions that were often revisited, that took a long time to respond and where values were frequently changed/updated. These three aspects together excluded situations where the student could have left the device unattended or questions where students had to write instead of selecting an option. It also excluded questions that were only revised instead of edited and so on. This process allowed teachers to identify difficult subjects, preferred content and component adequacy to each age level or provided material.

Overall, both teachers and students were very pleased with the prototypes suggesting new case studies and features. Teachers appreciated the possibility of monitoring students' activities while away from classes, on a deferred mode, with the ability to define their own tools and with the inclusion of behavior and hints on content that was previously passive.

6.3 Mobile Interaction Design Implications

Regarding the design process and the usability questions that were found during the two processes, the framework allowed designers to work closely with the expert teams, easily sharing concepts and their visions through the prototypes. Furthermore, the short prototyping cycles allowed expert-users to quickly assess the feasibility of such systems even on real-case scenarios. These experiences were even more successful since the expert teams and final users were able to use the actual devices, resulting in a much more confident evaluation process, where users were actually involved and on their working environment.

Overall, the framework provided a large step forward during the design process and led to much more efficient results and collaborations. From the usability and design team standpoint, the usage of low-fidelity sketch-based prototypes and high-fidelity prototypes provided interesting results, allowing users to actively prototype their own applications and providing a softer and sounder transition between design fidelities.

Log revision also led to interesting findings. For instance, track-bars, although not requiring text input, raised some difficulties mainly given the small size of the interactive counter. Moreover, when completing a task, if users were seated, they usually used the device's QWERTY keyboard. However, once walking they preferred to use the virtual keyboard, using one hand to hold the device and the other to tap on the virtual keyboard, alternating with any other activity that required their hand. Curiously, once seated again, they would not return to the physical keyboard. Also, while walking, accuracy towards buttons was much lower.

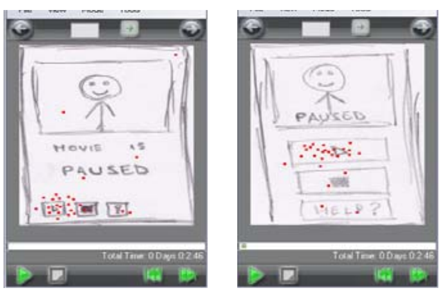

Figure 8. Two different iterations of sketch-based prototypes analyzed on the log player.

Figure 8 shows two screenshots of a low-fi prototype for the movie player being analyzed on the log player. Since all the logs have time-stamps and are cataloged by date, it was simple to correlate the logs and the locations/settings from which they resulted. Moreover, even specific portions of each evaluation session could be identified (e.g., at the beginning of the test, the user was seated; at the end of the evaluation test, the user was walking to another class). These situations were mapped to parts of the log where we noticed different accuracies regarding button selection and interaction, which allowed us to see that most of the missed taps on the screen referred to the situations where users were walking. As expected, while they were seated, accuracy was much higher. However, the log analysis provided a fairly precise idea of the necessary size and location for each button.

On the left side, a first prototype shows that users had some difficulties while using the video controls. This was particularly true when users were walking. On the right side, a second version of the same prototype, with larger buttons, shows that user accuracy, while selecting and using the controls was much higher. Each of the dots marked on the prototype identifies a tap on the screen. These can be viewed simultaneously, as depicted, showing heat zones, or sequentially, based on the actual user behavior.

Other results showed that components placed too close to the edges of the screen also raised some usage difficulties, especially when students used their fingers instead of the device's stylus.

7. RESULTS AND CONCLUSIONS

Throughout the development of the aforementioned case studies, several issues became clear and new goals started to emerge as the prototyping and evaluation sessions took place. Our initial assessment objectives referred to the prototyping framework and to the outcome that it's designing and evaluation features would provide. On the first facet, the tool allowed quick and easy creation of prototypes with different fidelities. End-users were much more satisfied by using actual devices, getting real feedback and actively participated on all stages. The initial probing goals were achieved as users created their own prototypes, generating new ideas and tools while using the framework.

The designers that were involved in both case studies responded to usability questionnaires and were very pleased with the easiness and amount of features available in the framework. This was further validated since other experts (e.g., therapists and teachers), with no particular knowledge in prototyping or programming techniques, were also able to materialize their own visions and needs through the prototyping framework. Here, the usability guidelines played an important role, limiting the amount of components in each screen and automatically docking their location, suiting several devices and screen resolutions. The shorter prototyping periods and intensive participation of non-designers together with the various fidelities and customization possibilities were frequently praised by all the users.

On the evaluation facet, all the involved designers considered the revision of users' behavior, without the need for direct observation, extremely useful. In fact, this allowed the detection of several issues which translated directly into UI improvements. Results were particularly interesting since they focused not only on a wide variety of contexts but also allowed the detection of problems that emerged while transiting between contexts. The logs and respective player provided insight on navigation patterns, size and location of components, amount of text, font size, among others. The different fidelities in concert with the realistic usage experience, since users roamed through different contexts with actual devices, allowed the evaluation of UI layouts, color arrangements, components, even detecting what colors were more adequate to certain lighting conditions and in which locations the user interfaces needed more contrast. This information was complemented by the questionnaires that were prompted during their utilization, capturing contextual information on-the-spot.

Moreover, since users interacted with the prototypes without direct observation and on familiar settings, their behavior was more natural and allowed us to see a set of interesting behavior patterns (e.g., keyboard usage, application exchange). Once again, the utilization of actual devices played an important role since it allowed the usage of the prototypes on devices with different screen resolutions, weight, size and interaction characteristics.

These results were even more interesting when used in conjunction with video equipment. The video capturing kit that we used was composed by inexpensive and common material available in our lab (e.g., backpack, webcam, laptop, and hat). Still, it provided very useful footage of users interacting with the prototypes and with the contexts through which they passed. By correlating the time-stamped logs and videos it was possible to detect, hesitations, reactions and usability problems as well.

The two case studies validated the positive influence of the prototyping and evaluation framework on the design process. Some of the findings resulted in modifications that were specific to the domains of each case study while others can be translated into generic guidelines that can apply to most mobile devices when used ubiquitously. The prototypes and evaluation sessions gave designers and other researchers the opportunity to assess the feasibility and adequacy of the envisioned applications on real-life scenarios. Moreover, the analysis of the evaluation data played an important role on research fields such as psychotherapy and education. In fact, a conclusion that was drawn from these experiences points the possibility of using the framework to create fully functional applications to support paper-based activities.

Given the positive results from these tests and experiences, we have integrated the framework into a new group version. Although beyond the scope of this paper, the team prototyping framework is worth mentioning and has benefited from the developments and results achieved through the experiences that we have presented. It introduces a set of features that, using the mobile prototyping framework, allow designers to cooperate in the creation and adjustment of designs, sketches and prototypes. The tool includes a large screen display module where several prototypes can be seen simultaneously. Moreover, it contains communication tools that provide means to visualize the evaluation sessions in real-time. Used in concert with the log player, it enables teams to review several logs simultaneously, comparing a user or a prototype’s performance in various settings.

Finally, this work is part of and based on a complete methodology that was developed and aims at supporting the design of mobile applications through a user-centered design approach. Following a parallel research direction and acting as a complement for the prototyping framework, it compiles a set of guidelines that suggest the generation of scenarios and selection of appropriate contexts and techniques for the evaluation of mobile applications.

8. ACKNOWLEDGEMENTS

This work was supported by LaSIGE and FCT, through project JoinTS, through the Multiannual Funding Programme and scholarship SFRH/BD/28165/2006.

9. REFERENCES

[1] Barnard, et al. Capturing the effects of context on human performance in mobile computing systems. Personal and Ubiquitous Computing, Vol. 11, No.46, pp.81-96.

[2] Brewster, S. (2002). Overcoming the lack of screen space on mobile computers. Personal and Ubiquitous Computing, 6.

[3] Carriço, L., Sá, M. Hand-held Psychotherapy Artifacts. In Procs. of HCII 2005.

[4] Carter, S., et al. Momento: Support for Situated Ubicomp Experimentation. CHI’07, pp. 125-134, ACM.

[5] Consolvo, S. and M. Walker, Using the Experience Sampling Method to Evaluate Ubicomp Applications. IEEE Pervasive Computing, 2003. 2(2): pp. 24-31, IEEE.

[6] Davis, R., et al. SketchWizard: Wizard of Oz Prototyping of Pen-Based User Interfaces. UIST’07, pp. 119-128, ACM.

[7] Duh, H.B.-L., G.C.B. Tan, and V.H.-h. Chen. Usability Evaluation for Mobile Device: A Comparison of Laboratory and Field Tests. Mobile HCI'06. ACM.

[8] Froehlich, J., et al. MyExperience: A System for In Situ Tracing and Capturing of User Feedback on Mobile Phones. MobiSys’07, pp. 57-70, ACM.

[9] Hartmann, B., et al. Wizard of Oz Sketch Animation for Experience Prototyping. Adjunct Procs. of Ubicomp 2006.

[10] Holmquist, L., Prototyping: generating ideas or cargo cult designs? Interactions, 2005. 12(2) p. 48-54, ACM.

[11] Hulkko, S., et al. Mobile Probes. NordiCHI'04. ACM.

[12] Iachello, G., et al. Prototyping and Sampling Experience to Evaluate Ubiquitous Computing Privacy in the Real World. CHI’06, pp. 1009-1018, ACM.

[13] Jones, M. and G. Marsden, Mobile Interaction Design. 2006, John Wiley & Sons, England.

[14] Kjeldskov, J. and C. Graham. A Review of Mobile HCI Research Methods. Mobile HCI'03, pp. 317-335, Springer.

[15] Kjeldskov, J. and J. Stage, New Techniques for Usability Evaluation of Mobile Systems. International Journal of Human Computer Studies, Elsevier, 2003.

[16] Klemmer, S. R., et al, SUEDE: A Wizard of Oz Prototyping Tool for Speech User Interfaces. In Proc. of UIST’00, San Diego, California, USA, pp. 1-10, ACM.

[17] Landay, J. SILK: Sketching Interfaces Like Krazy. CHI’96, pp.398-399, ACM.

[18] Li, Y., Welbourne, E., Landay, J. Design and Experimental Analysis of Continuous Location Tracking Techniques for Wizard of Oz Testing. CHI’06, pp. 1019-1022, ACM

[19] Lin, J. et al. Denim: Finding a Tighter Fit Between Tools and Practice for Web Site Design. CHI’00, pp.510-517, ACM.

[20] McCurdy, M., et al. Breaking the Fidelity Barrier: An Examination of our Current Characterization of Prototypes and an Example of a Mixed-Fidelity Success, CHI’06, ACM.

[21] Nielsen, C.M., et al. It's Worth the Hassle! The Added Value of Evaluating the Usability of Mobile Systems in the Field. NordiCHI’06. ACM.

[22] Sá, M., Carriço, L. Handheld Devices for Cooperative Educational Activities. *SAC’06*, pp. 1145-1149, ACM.

[23] Sá, M.d. and L. Carriço. Low-fi Prototyping for Mobile Devices. CHI'06 (extended abstracts), pp 694-699, ACM.

[24] Liu, L. A., Li, Y. BrickRoad: A Light-Weight Tool for Spontaneous Design of Locaton-Enhanced Applications, CHI’07, pp 295-298, ACM.