Designing Interfaces for Small Screens

Author: Martha Lyngnes
Department of Product Design
Norwegian University of Science and Technology

ABSTRACT

This paper deals with issues relating to information presentation and interaction with small electronical devices, such as personal digital assistants (PDA). The main challenge with these devices is the limited size available, and the constraints it may inflict on information display and interaction. With the introduction of PDAs within the field of healthcare, this issue becomes even more pressing. This paper will extract knowledge about Graphical User Interface (GUI) design from the fields of information organization, interface design, interaction, and information perception. Selecting the information and elements to present on the screen is discussed as a substantial part of interface design, and in particular to small GUIs. The different sections result in guidelines and principles for the design for screen. Central factors in designing for small GUIs are highlighted and exemplified.

KEYWORDS

Information organization, interface design, interaction, information perception, small screen display

1.0 INTRODUCTION

This paper discusses different aspects of designing small screen displays. A typical screen size for PDAs is 240x320 pixels, or 3,8” (5,76 cm x 7,68 cm)(HP iPAQ h5550) in comparison a computer-screen has a screen area about 15 times that size. The background for exploring this field is the introduction of PDAs within the field of healthcare. Especially in the setting of doctors in a hospital, it is essential that the design of the interface is usable, that is, effective to use, efficient in use and satisfying to use. The environment in hospitals is often stressful, and the interface must be easy to operate under these conditions. The goal of the article is to come up with a collection of guidelines and principles for designing interfaces for small screens. An important part of designing interfaces is the issue of interaction, - the visualization is not successful if the interaction part fails.

There is limited literature available primarily concerning design for small screens within the field of interface design. However, within the field of general interface design there is several literature sources, and from these knowledge can be extracted and applied to use on small GUIs. Several books by Edward R. Tufte deal with information visualization in general; “Envisioning Information”(1990) is about representing the visual world of experience and measurement on paper, and “Visual Explanations” (1997) describes design strategies for presenting information. About designing for the screen, “Designing Visual Interfaces – Communication Oriented Techniques” by Kevin Mullet and Darrell Sano (1995), defines a set of fundamental techniques for the designing of visual interfaces. Furthermore, “Interaction Design – Beyond human computer interaction” by Jennifer Preece, Yvonne Rogers and Helen Sharp (2002) introduces interaction design. In “Usability Engineering” by Jakob Nielsen (1993) the usability perspective is explored. “The Design of Everyday Things” by Donald A. Norman (1988) also investigates this subject. The way in which humans perceive information from electronic devices is dealt with in "Things that make us smart" by Donald A. Norman (1993).

The following parts will deal with different aspects of designing interfaces. How should the interface be organized to present the information clearly, what information should be presented, what guidelines exist for designing interfaces, how to design for usability and how the users perceive the information presented. In some areas, the guidelines and principles presented applies as much to general computer screens as to small screens. However, aspects relating especially to design for small screens will be enunciated. The advantage provided by small devices is their mobility, and this advantage is considered superior enough for people to accept the drawback of the limited area for visualization. The article will try to reveal ways to minimise this drawback. By studying the available literature on designing for screen and interaction, the main issues will be presented in the pursuing sections. Issues concerning information organization, interface design, interaction and information perception will be discussed.

2.0 INFORMATION ORGANIZATION

The information presented needs to be organized and structured to communicate the meaning in a clear way. At the same time it needs to be able to show a dynamic content and adjust to different needs, and therefore have the possibility to adjust to various situations. As Tufte (1990: 51) states: "Clutter and confusions are failures of design, not attributes of information". According to Mullet and Sano (1995), structure affects the visual experience at its most primitive level because it is the first aspect of the display to be perceived as information. This information is extracted and used to guide subsequent interaction. Furthermore, they state the key benefits of structure to be:

Unity: Visual structure ties even highly disparate design elements together and allows them to work altogether towards a common communication goal

Integrity: A strong and coherent structure keeps the design focused on the communication goal by creating an emergent form that contributes to the meaning of the composition.

Readability: Structure enhances readability by dividing the information content of the entire display into manageable subsets that can be processed separately or in parallel, according to the designer's wishes.

Control: Structure allows users to predict areas of interest and eases their navigation through the composition.

Structuring the information is substantial for small screens because the display will more easily become cluttered and chaotic when several elements are presented without an overall agenda. The structuring of elements can prevent this undesired effect.

2.1 The Gestalt Principles

The main concern in structuring elements on a screen is not how it is intended from the designer, but how the user perceives them. An important background for perceptual organization is the Gestalt principles, described by the psychologists of the Gestalt school in the 1920´s. Wertheimer (1958) defines the principles describing the processes by which individual elements are grouped into gestalts (wholes) in "Principles of Perceptual Organization". The principle of **proximity** describes the tendency of individual elements to be associated more strongly with nearby elements than with those that are farther away. Figure 2.1.1a illustrates this phenomenon. The eye organizes the dots first into four vertical columns because the horizontal separation is much greater that the vertical separation. Then, because the separation between the middle two columns is greater that the outer gaps, the whole figure is seen as two groups of two columns each. The principle of **similarity** observes that elements will be associated more strongly when they share basic visual characteristics, such as shape, size, colour, texture, value, and orientation, than when they differ along these dimensions. This can be observed in figure 2.1.1.b, where it is perceived as two groups of two columns each, despite the fact that the spacing between elements and columns are equal.

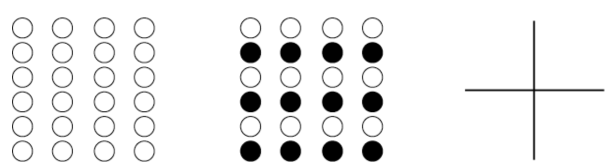

Figure 2.1.1: Gestalt grouping phenomena: a. proximity, b. similarity, c. continuity (Mullet and Sano, 1995)

The principle of **continuity** describes the preference for continuous, unbroken contours with the simplest possible physical explanation, rather than more complex but equally plausible combinations of more irregular figures. The form in figure 2.1.1c is thus perceived as two crossing lines rather than four attributing lines or two (or even four) opposing angles. The related principle of **closure** describes the powerful human tendency to interpret visual stimuli as complete, closed figures, even when some necessary contour information is absent. Figure 2.1.2a is easily seen as a triangle superimposed on three complete circles even neither of these forms is technically present. Furthermore, Mullet and Sano (1995) have described two principles adding to the gestalt theory, area and symmetry. The principle of **area** states that the smaller of two overlapping figures will tend to be interpreted as a figure while the larger is to be interpreted as ground. In figure 2.1.2b, the inner square is perceived as a distinct form in front of a larger square instead of a hole in the larger form. Finally, the principle of **symmetry** describes grouping based on the emergent properties of the form instead of the characteristics of its constituent parts. Figure 2.1.2c, is thus seen as two overlapping rather than three objects.

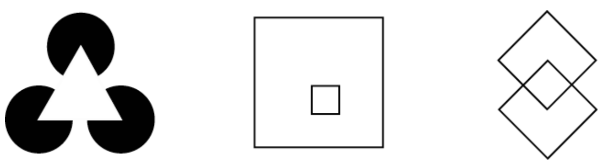

Figure 2.1.2: Gestalt phenomena in form perception: a. closure, b. area, c. symmetry (Mullet and Sano, 1995)

The gestalt principles can be used to guide the attention of the user, by grouping and arranging elements in the manner described. In this way the designer can establish a hierarchy of importance for the elements, which will be described in the next section.

2.2 Achieving organization and structure

To achieve organization and structure, Mullet and Sano (1995) suggests four principles: grouping, hierarchy, relationship and balance. Organization begins with classification, which involves grouping related elements and establishing a hierarchy of importance for elements and groups. Then the display itself can be structured to reflect the relationship between the elements while maintaining a pleasing balance in the resulting composition. To achieve this they suggest four techniques for structuring a display:

Using symmetry to ensure balance: Symmetry provides a clear organization, however sometimes at the expense of visual interest. The design of effective interface design nevertheless is to present information efficiently and nonintrusively, and not to entertain or excite. Thus the restful character of the symmetrical layout is often perfectly appropriate.

Using alignment to establish visual relationships: Positional alignment of elements reduces the complexity of a display by making the global form cleaner and more understandable. By limiting exceptions to the positioning rules obeyed by the elements within the composition, alignment makes intentional deviations more salient. Used in conjunction with negative space, alignment is an important tool for constructing visual hierarchies.

Optical adjustment for human vision: Visual design is grounded in perceptual, rather than physical phenomena, so compensation for the peculiarities of human vision is often required. In general, the more acute the angle of the object, more compensation is needed to maintain the proper alignment. In optical spacing, the rule of thumb is to equalize the area rather than the distance between elements, as shown in figure 2.2.1a and b.

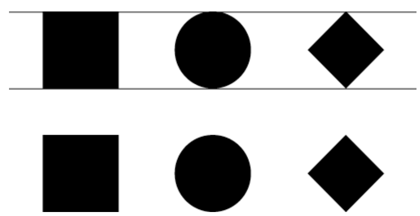

Figure 2.2.1a: With equivalent scaling, rounded or acute forms appear too small relative to rectangular elements

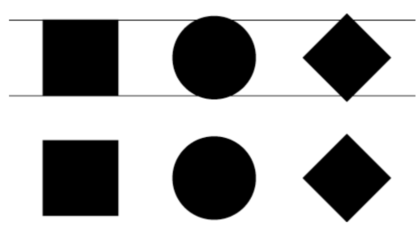

Figure 2.2.1b: Extending rounded or slightly beyond the target dimension produces optically equivalent scaling

Shaping the display with negative space: Empty regions are utilized in a well-organized display. Negative space plays the crucial role of directing the viewer’s attention to the regions where important information is provided and allowing the global structure of the composition to assume a meaningful configuration.

If organization and structure are applied the user will become familiar with the appearance of the application, and it can be easier to quickly gain an overview of the display and complete the desired task. In achieving this, the subject of consistency stands out. To provide consistency within an information presentation consisting of several displays, a unity in the design approach is essential. In Mullet and Sano (1995) this comprehensive system of organization is defined as a program.

2.3 Design Programs

Programs are based upon repeated sizes and proportions (modules) or upon forms and ideas (themes) that bring regularity and structure to the user experience. Subsequently, humans are pattern-recognition animals. Moreover, they state the benefits of a systematic approach to be:

Structure: Module is intimately related to structure, and the module reflects and draws justification from structural requirements while the structure is revealed and reinforced through consistent application of the module.

Predictability: Programmatic design simplifies the communication task by preparing the user to respond to a small number of familiar patterns in a predictable way.

Efficiency: Modular design permits great economies of production once the general scheme has been extended to cover the entire problem space.

Additionally, they specify the grid as the central element of any successful program. If a grid is present, it allows the static layout principles to be codified and propagated consistently across a series of displays. By structuring each screen along similar lines, the grid ensures that users will benefit from experience with the system as they learn to predict where a particular piece of information will be found. For instance, that the

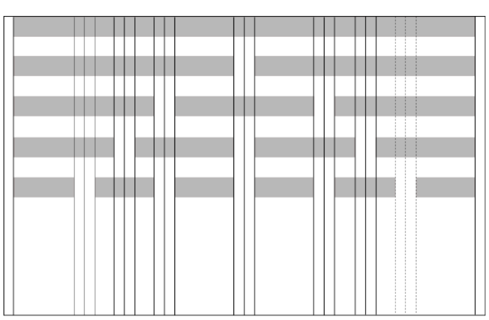

Figure 2.3.1a: A canonical grid (Mullet and Sano, 1995)

OK command always is at the lower right of a dialog box. In figure 2.3.1a a canonical grid is displayed. This grid supports two-, three-, four-, and six-column layouts in any graphical user interface (the 1/6 and 5/6 divisions are implicit). In the screen shot in figure 2.3.1b, is based on the canonical grid. To visualize the grid, all but the middle three lines of the grid in figure 2.3.1a can be ignored.

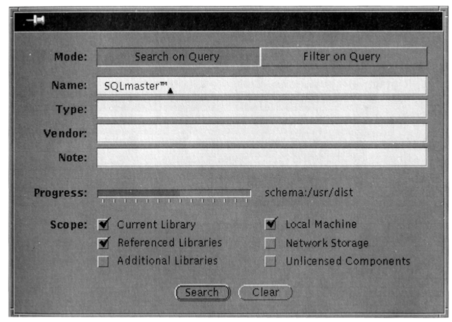

Figure 2.3.1b: The use of grid structure (Mullet and Sano, 1995)

For the use of a design program, Mullet and Sano (1995) presents three principles: **focus, flexibility** and **consistent application**. The focus within a program should rely on a small number of modular units that reveal the underlying spatial logic, the grid, of the program. A clear focus enhances the readability of a display by introducing a rhythm and regularity that makes the structure predictable and explicit. These qualities simplify the movement of visual attention across the display by allowing the viewer to unconsciously estimate the distance between resting points and to skip over uninteresting portions when necessary. Flexibility becomes an issue when unanticipated situations occur. The best programs are designed to encompass boundary conditions, since this is where a design program most often fails. For the design program to be effective, it must be used consistently wherever it appears to ensure that its programmatic aspects will become visually apparent after even minimal exposure.

The small GUI has limited space, nevertheless the use of negative space as a design element is important in small displays, both to ease the perception and provide an overview. The importance of consistency throughout the application is enhanced with the use of small displays. When using the application, users will get accustomed to the organization, and this will allow users to orientate themselves and respond to the tasks more quickly once they have reached the level of experienced users. By making and applying a design program for the various displays in the medication interface, the health personnel will swiftly become familiar with the application and how to interact with it. This ability can prove useful when using the medication system on a different computational device. When a doctor is prescribing a medicine, and for example is doing this from his desktop-computer in his office: using familiar structure and elements, and applying the same design program to both variations of the system, can lower the threshold between using the system on different computers.

3.0 Interface Design

The medication interface today, is in the form of a chart on an A4-sized sheet of paper and is to be filled in manually, as shown in figure 3.0.1.

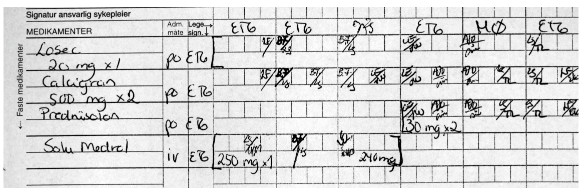

Figure 3.0.1 Excerpt from the medication chart (Fagerholt, 2003)

Through the remediation of the medication process from paper to computer, it is important to take advantage of the new medium, not just remediate another solution adjusted to a different format and medium (Bolter and Grusin, 1990). By turning the medication process electronic, one of the advantages is that the record of medication can go directly into the Electronic Patient Journal (EPJ) thus one work operation can be removed. Before, the manually filled in form had to be scanned in subsequently to be a part of the journal. The alteration of size from A4 to a small screen necessarily imposes some changes. It is important that the medium does not stand in the way of the information it tries to communicate (Bolter and Grusin, 1990). Completing the task of medication should not be any more complicated on a PDA than on paper, thus the computer should be "transparent" for the user. The information needed to complete the task will be the same after the remediation of the medication system as before. Which leads to the next section, selecting the relevant information to be displayed.

3.1 Information Selection

The information communicated through the screen is the utmost important element. Users need this information to respond to, and to complete their task. For information content on screens, Nielsen (1993) states, "Less is more"; user interfaces should be simplified as much as possible, since every additional feature or item of information on a screen is one more thing to learn, one more thing to possibly misunderstand, and one more thing to search through when looking for the relevant piece of information. Adding information and data fields to a user interface can distract the user from the primary information. The ideal is to present exactly the information the user needs at exactly the time and place where it is needed. Another ideal presented by Nielsen (1993) is that information that will be used together should be displayed close together, and at minimum on the same screen. Deciding the essential information to display in designing for small screens, is a difficult, nevertheless, highly relevant task.

Performing a task analysis based on the existing system of medication can be guidance in the process of choosing which information to present and how to present it in the new system. The task analysis searches to answer three questions: what are the users trying to achieve, why are they trying to achieve it, and how are they going about it (Preece, Rogers and Sharp, 2002). Based on a proper task analysis, it is often possible to identify the information that is truly important to users, and which will enable them to perform almost all of their tasks (Nielsen, 1993). In Preece, Rogers and Sharp (2002) a technique for performing a task analysis in practice is presented: Hierarchical Task Analysis (HTA). HTA involves breaking a task into sub tasks and then further into sub tasks and so on, as illustrated in figure 3.1.1. These are in turn grouped together as plans that specify how the tasks might be performed in an actual situation. HTA focuses on the physical and observable actions that are performed, and includes looking at actions not related to software or an interaction device at all. The starting point is a user goal. This is then examined and the main tasks associated with achieving that goal are identified, and if appropriate, further subdivided into sub tasks.

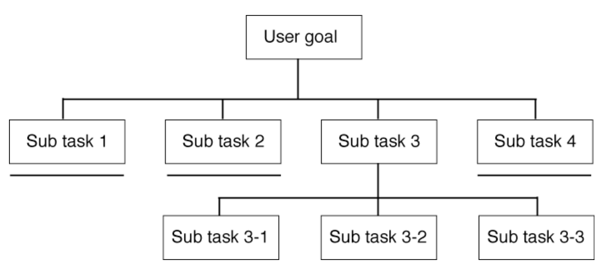

Figure 3.1.1: A graphical representation of Hierarchical Task Analysis

When the relevant information is defined, it will normally be better to design a single screen with this information and relegate less important information to auxiliary screens than to cram all the information that might be useful into a set of screens that will require the user to switch screens for even the simplest task (Nielsen, 1993) Or as emphasized by Rice Skogen (2004); instead of presenting many simplified menus, one should rather provide a few dense screens of choices.

Not only the information must be carefully selected, but also every element to put on the screen. This task is utterly pressing on a small screen where every pixel is valuable. In Nielsen (1993) avoiding unnecessary information in interface design is discussed. Some information may not be necessary, for example, many programs dedicate large amounts of screen space to display the name of the program, the vendor's logo, the version number, and other similar information. This information is often placed in the top right area of the screen, which due to our reading direction is the prime area of the display (Rice Skogen, 2004). Even though this information is potentially important and should be available for users making bug reports, it normally takes up screen space that could have been used for other purposes (maybe even as "white space" to make a better layout). And sequentially, any piece of information is something users will have to look at when they search the screen, and it will therefore slow down their performance by some fraction of a second (Nielsen, 1993). For an application in frequent use, such as the medication interface, effective performance is significant - time is money.

Especially for small screens Rice Skogen (2004) recommends short navigational routes. Scrolling (both horizontal and vertical) in a screen to get an overview of the information should be avoided. The reason for this is that the user will be occupied with keeping track of the changes rather than paying attention to what they are trying to achieve. However, Rice Skogen (2004) states that when unavoidable, deeper hierarchies are to prefer in front of long scrolling pages and to shorten hierarchies, indexes can be used to provide direct access to content. In the often-stressful situation in a hospital, time is precious, and the reason for introducing information technology into the healthcare sector is to free some of the valuable time the employees uses for administration duties, so that this time can be used to aid their primary goal: to take care of the patients. Hence, the information should be accessed efficiently, and with as few steps as possible.

3.2 Colour for the Screen

With the communication of information through the screen, the problem can be described by Tufte (1990: 89) "as the essential dilemma of a computer display: at every screen are two powerful information-processing capabilities, human and computer. Yet all communication between the two must pass through the low-resolution, narrow-band video display terminal, which chokes off fast, precise, and complex communication." To make this communication as smooth as possible some special guidelines for the design for the screen can be established.

The use of colour in interface design can improve the communication through the screen, and upgrade the user experience. In Nielsen (1993: 119, with reference to Rice, 1991 and Travis, 1991) the three most important guidelines with respect to the use of colour in screen design are presented:

The design ought to be limited to a small number of consistently applied colours. Colour-coding should be limited to no more than 5 to 7 different colours since it is difficult to remember and distinguish larger numbers. Light greys and muted pastel colours are often better as background colours than screaming colours are.

The interface should be able to use without the colours, owing to the fact that about 8% of males (and about 0,5% of women) are colour-blind. Hence any colour coding of information should be supplemented by redundant cues that make it possible to interpret the screens even without being able to differentiate the colours.

Colour should only be used to categorize, differentiate, and highlight, not to give information, especially quantitative information.

These guidelines for the use of colour in screen design should be followed in the medication interface. Many different users will use this application frequently, and making an interface with neutral, pleasing colours can make the interaction with the system, at least, less laborious for the vision. Toleration for colour-blindness should be incorporated, so that in the case of a colour-blind user, no essential information can be misinterpreted.

3.3 Type for the Screen

When using text on the screen, the legibility is utmost important. For the use on screen there is developed a set of bitmapped fonts. In Götz (1998) the sans serif typefaces is presented as the better to use on screen, this due to the serif typefaces having varying thickness of the strokes that make up the letter, and the rounded nature of the serifs themselves, - these characteristics together with the modular character of pixels do not add up to legibility. The modular structure of the pixels gives them great flexibility, but they also have the disadvantage of being tied to a grid structure. To avoid distortion Götz (1998) states that the screen resolution of 72 pixels per inch requires special processing for small type sizes up to 20 pt, depending on the bitmap. The resolution of a computer screen compared to printed material is greatly inferior in quality. Small type sizes (6 to 9 pt) that are perfectly legible on paper can hardly be read on a screen. Thus, Götz (1998) suggests that text on the screen should be at least 10 point, but at best between 11 and 14 point. And the corresponding title type size should be between 14 and 20 point. However, in choosing the size, the individual typeface needs to be taken into account. Furthermore, it is important that the letters on the screen are properly spaced. Götz (1998) recommends a tracking of 5 to 10 units for improved legibility. Interlinear spacing is an important element in making text easy to read, and it should be set to a more generous value on screen than for text on paper, and recommended in Götz (1998) is that the line spacing is set to about 150% or more.

These guidelines for use of text on screen must be emphasized in the design for small screens. The temptation for the designer to reduce the text-size to cram more information into a display should not be given into. The sight of the text should be clear even for persons with slightly imperfect vision. This is highly relevant in medication to avoid misreading names of medicine, and thus imposing a danger to the patients.

4.0 INTERACTION

In Dourish (2001) the system is described as the medium through which a designer and a user communicate. To get this communication to run smoothly, it is important that the usability of the interactive product is well thought-through. In the ISO 9241 standard (Preece, Rogers and Sharp, 2002) the usability term is defined as: the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use.

Nielsen (1993) has defined five usability attributes, regarding learnability, efficiency, memorability, errors and satisfaction. Learnability implies that the system should be easy to learn so the user rapidly can start to get work done with it. The medication system should have a shallow learning curve. If it is too difficult to start using, the health personnel will not use the system. They will then prefer the old system in which they are experts. Efficiency means that the user, once having learned the system, can reach a high level of productivity. As mentioned before, in section 3.1, the reason for introducing information technology into health care is to free some of the valuable time of the employees, by removing additional elements of administration. Memorability concerns that the system should be easy to remember, so that the casual user is able to return to the system after some period, without having to learn everything all over again. The need for a rememberable system arises due to the fact that health personnel often work on shift, and therefore the importance of not starting from scratch after every week off is pronounced. Errors should be obliterated and the system should have a low error rate, meaning that users make few errors during the use of the system, and that errors they can easily be recovered from. Furthermore, catastrophic errors must not occur. The absence of errors is the most important attribute to a medication system, - if errors do occur, the life of the patient may be at stake. Satisfaction means that the system should be pleasant to use, so that users are subjectively satisfied when using it; they like it. And of course, if the system is pleasurable to use for the health personnel, this may improve the quality of their workday.

To compare the two definitions, the ISO standard gives a more overall picture of what usability is, whereas Nielsen's definition associates usability with more specific attributes of a system. Nevertheless they transmit the same core values. How to achieve usability will be discussed in the next section.

4.1 Usability Heuristics

For achieving usability there are both design principles and usability principles available. The different applications of the two are explained in Preece, Rogers and Sharp (2002): the design principles are generalizable abstractions intended to orient designers towards thinking about different aspects of their design, to explain and

improve the design. In comparison, the usability principles tend to be more prescriptive, and are mostly used as the basis for evaluating prototypes and existing systems. When the usability principles are used in practice they are commonly referred to as heuristics. However overlapping occurs between the two principles.
The most referred to usability heuristics are these ten established by Nielsen (2001):

Visibility of system status. The system should always keep users informed about what is going on, through appropriate feedback within reasonable time.

Match between system and the real world. The system should speak the users' language, with words, phrases and concepts familiar to the user, rather than system-oriented terms. Follow real-world conventions, making information appear in a natural and logical order.

User control and freedom. Users often choose system functions by mistake and will need a clearly marked "emergency exit" to leave the unwanted state without having to go through an extended dialogue. Support undo and redo.

Consistency and standards. Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform conventions.

Error prevention. Even better than good error messages is a careful design, which prevents a problem from occurring in the first place.

Recognition rather than recall. Make objects, actions, and options visible. The user should not have to remember information from one part of the dialogue to another. Instructions for use of the system should be visible or easily retrievable whenever appropriate.

Flexibility and efficiency of use. Accelerators - unseen by the novice user -may often speed up the interaction for the expert user such that the system can cater to both inexperienced and experienced users. Allow users to tailor frequent actions.

Aesthetic and minimalist design. Dialogues should not contain information that is irrelevant or rarely needed. Every extra unit of information in a dialogue competes with the relevant units of information and diminishes their relative visibility.

Help users recognize, diagnose, and recover from errors. Error messages should be expressed in plain language (no codes), precisely indicate the problem, and constructively suggest a solution.

Help and documentation. Even though it is better if the system can be used without documentation, it may be necessary to provide help and documentation. Any such information should be easy to search, focused on the user's task, list concrete steps to be carried out, and not be too large.

The advantage of using usability principles for heuristic evaluation is described by Nielsen (1993) to be that it provides a list of usability problems in the interface, with respect to the usability principle that where violated by the design. However, heuristic evaluation does not provide a systematic way to generate fixes to the usability problems or a way to assess the probable quality of any redesigns. In this regard the design principles are available for guidance.

4.2 Design principles

The design principles can be a part all through the design process, and lead the designers mind towards a design where the usability is integrated. These six principles regarding design for usability, is established by Norman (1988):

Visibility. Visibility of the system means that the product should reveal by itself what elements that can be manipulated.

Feedback. Feedback is related to the concept of visibility. Feedback is about sending back information about what action has been done and what has been accomplished, allowing the user to continue with the activity.

Constraints. Constraining refers to determining ways of restricting the kind of user interaction that can take place at a given moment.

Mapping. Mapping refers to the relationship between controls and their effects in the world.

Consistency. Consistency refers to designing interfaces to have similar operations and similar elements for achieving similar tasks. One of the benefits of consistent user interfaces is that they are easier to learn and use.

Affordances. Affordance is a term used to refer to an attribute of an object that allows people to know how to use it.

Preece, Rogers and Sharp (2002) suggests following implications for the design principles:

The more visible functions are, the more likely users will be able to know what to do next. In contrast, when functions are "out of sight", it reduces learnability and increases memory and cognitive load. The relationship between the way functions is positioned and what they do, will make it easier for the user to find the appropriate function for the task at hand. The principle of visibility can be complicated to achieve on a small screen, where the place is very limited. Various kinds of feedback are available for interaction design, - audio, tactile, verbal, visual, and combinations of these. Deciding which combinations are appropriate for the particular system is central. Using feedback in the right way can also provide the necessary visibility for user interaction. In the medication system both visual and audio or verbal feedback can be useful, the visual feedback by directly showing the user what operation is completed, and the verbal or audio feedback for attracting attention to the system completing a task. A common way of constraining in graphical user interfaces is to deactivate certain menu options by shading them, thereby restricting the user only to actions permissible at that stage of the activity. In the medication system, this can be utilized by deactivating the choice of signing for a medicine, when no prescription is given. An advantage of this form of constraining is that it prevents the user from selecting incorrect options and thereby reducing the chance of making a mistake. A disadvantage is that there is a hidden causality between functions that the user need to know, or the designer must illustrate through feedback. By applying mapping, the user interface will become more logical. An example of a good mapping between control and effect is the up and down arrows used to represent the up and down movement of the cursor, respectively, on a computer keyboard. Consistency for an interface can be achieved by using the same operator to achieve similar tasks, or in more complex interfaces to create categories of commands that can be mapped into subsets of operations. Hence, all operations connected to managing the document itself is concealed under the label "File". At a very simple level, to afford means, "to give a clue" (Norman, 1988). When the affordances of a physical object are perceptually obvious it is easy to know how to interact with it. Since the establishing of this term, Norman has tried to clarify that there are two kinds of affordance: perceived and real. Physical objects have real affordances, that are perceptually obvious and do not have to be learned. In contrast, screen-based user interfaces are virtual and do not have these kinds of real affordances. Screen-based interfaces are better conceptualized as perceived affordances, which are essentially learned conventions. A button on a display affords clicking only because it is a learned convention. The perception of information will be explored further in the next section.

5.0 Information perception

The way the system is perceived by the user can affect its ability to help the user achieve their goal. According to Norman (1993) there are two major tasks for the user of an information display, - the first, finding the relevant information, and the second, computing the desired conclusion. For the users of a medication interface, this relates to a process of scanning the display to find the desired elements, for instance a text field, and then correctly fill in the information about the medication. Norman (1993) illustrates that many situations seem designed as if to deliberately form a mismatch with human capabilities. For instance, human memory is well tuned to remember the substance and meaning of events, not the details. Humans can essentialy attend to only one conscious task at a time and cannot maintain attention on a task for extended periods. Basically, humans are sensitive to changes in the environment and attend to changing events, not to continual ongoing ones. The same is true for memory: the tendency to remember novel and unexpected events better than regular, recurring ones. Humans are pattern-recognition animals, matching things that appear similar to past events.

Norman (1993) describes the human as a distractible, ungrammatical, illogical, and errorful creature. From a machine-centred point of view, these are all errorful behaviours, however from a human-centred point of view, these are all reasonable, sensible ways of acting. Norman (1993) emphasizes that when technology is not designed from a human-centred point of view, it does not reduce the incidence of human error nor minimize the impact when errors do occur. Due to this the technology should be designed to take this fact into account. Instead, the tendency is to blame the person who errs, even though the fault might lie with the technology, even though to err is very human.

5.1 Experiential and Reflective Cognition

Norman (1993) distinguishes between two kinds of cognition relevant in this concern: experiential cognition and reflective cognition. The experiential mode leads to a state in which events surrounding one can be perceived and reacted upon, efficiently and effortlessly. This is the mode of expert behaviour, and it is a key component of efficient performance. The reflective mode is that of comparison and contrast, of thought and of decision-making. This is the mode that leads to new ideas and novel responses. A system needs to support a combination of these two cognition types. Experiential thought is essential to skilled performance: It comes rapidly, effortlessly, without the need for planning or problem solving: look, see and respond. The patterns of information are perceived and assimilated and the appropriate responses generated without apparent effort or delay. Examples on experiential cognition include riding a bicycle and reading. Reflective reasoning does not have the same kind of limits on the depth of reasoning that apply to experiential cognition, but the price one pays is that the process is slow and laborious. Reflective thought requires the ability to store temporary results, to make inferences from stored knowledge, and to follow chains of reasoning backward and forward, sometimes backtracking when a promising line of thought proves to be unfruitful. The process takes time. Examples on reflective cognition include designing and learning.

It is of importance that experiential and reflective thought is used in the right places of an appliance. Pitfalls may occur if this is not accomplished, as listed in Norman (1993):

Tools for experiential mode behaviour that requires reflection. These tools turn simple tasks into problem-solving exercises, causing needless mental effort and taking needless time.

Tools for reflection that does not support comparisons, exploration, and problem solving. In many cases we need to be able to overlook the situation and compare alternative courses of action, or perhaps just ponder and reflect upon the variables involved.

Experiencing when one should be reflecting. The experiential mode leads to responses without thought, without contemplation. This is essential when events move rapidly, but if the situation changes, experiential cognition may not be flexible enough to change appropriately.

Reflecting when one should be experiencing. Too much reflection and the world will pass one by.

To get the right balance between experiential and reflective cognition in the medication system is important, so that simple tasks do not take excessive amounts of time, whereas tasks requiring use of knowledge and skills should encourage reflection. In the process of evaluating a patients symptoms and prescribing the right medicine, the system should provide the use of reflective cognition. If this process is entirely experiential and the doctor has not had any experience with a certain disease, he of she can fail to interpret symptoms correctly and give the medicine simply on experiential background. In contrast, the routine operation of signing that a medicine has been given should be performed with the skill of experiential cognition. As mentioned in section 3.1, the time of the health personnel is costly.

5.2 Cognition Processes

Cognition can also be described in terms of specific kinds of processes (Preece, Rogers and Sharp, 2002). These include:

Attention. Attention is the process of selecting things to concentrate on, at a point in time, from the range of possibilities available. The extents to which this process is easy or difficult depends on whether clear goals are present, and whether the information needed is salient in the environment.

Perception and recognition. Perception refers to how information is acquired from the environment, via the different sense organs and transformed into experiences. In general, information needs to be represented in an appropriate form to facilitate the perception and recognition of its underlying meaning.

Memory. Memory involves recalling various kinds of knowledge that allows us to act appropriately. How information is interpreted when it is encountered greatly affects how it is represented in memory and how it is used later. Other factors that affects the extent to which information can be subsequently retrieved is the context in which it is encoded, and that the human mind is better at recognition rather than recall.

Learning. According to Norman and Rumelhart in Norman (1993) there are three kinds of learning: accretion, tuning and restructuring. Accretion is the accumulation of facts, that add to our knowledge. Tuning is the practice between the initial stages of novice performance and the skilled performance of the expert. Restructuring is the difficult part of learning. In this stage the goal is forming the right conceptual structure. Accretion and tuning are primarily experiential modes, whereas restructuring is reflective.

Reading, speaking and listening. These three expressions refer to the language processing capabilities. Similar for these are that the meaning of a sentence or phrase is the same regardless of the mode in which it is conveyed. However several differences between these modes exist, for instance, voice is serial whereas vision is parallel, voice is transient whereas printed or displayed information is relatively permanent. And uttermost, people differ in their ability to use and understand language, both in reading, speaking and hearing.

Problem solving, planning, reasoning and decision-making. These are all cognitive processes involving reflective cognition. They include thinking about what to do, what the options are, and what the consequences might be of carrying out a given action. They often involve conscious processes, discussion with others, and the use of various kinds of artefacts.

Furthermore, Preece, Rogers and Sharp (2002) have developed a set of design implications for these different processes, which also can bee applied to small GUIs. The attention process requires that information is made salient when it needs attending to a given stage of a task. Different techniques like animation, colour, underlining, ordering, sequencing and spacing, can be used to achieve this. Nevertheless, these techniques must be used carefully to avoid cluttering the interface, too much of this results in distraction and annoyance. Representations of information need to be designed to be perceptible and recognizable across different media. Icons, and other graphical representations, sounds, speech output, text and tactile feedback should all be distinguishable from the environment to present their meaning clearly. Implications for the memory process includes not overloading users memory with complicated procedures for carrying out a task and promoting recognition rather than recall by applying elements consistently throughout the layout of the system. Interfaces should be designed to encourage exploration to improve the learning, to achieve this constraints can be applied and suggestions put forward to prevent users from making errors while "taking a tour" through the system. The processes of reading, speaking and listening imply that the length of speech-based menus and instructions should be kept to a minimum. Research has shown that people find it hard to follow spoken menus with more than three or four options. Additionally, the intonation of artificially speech voices should be accentuated, as they are harder to understand than human voices. The opportunity for making text large on a screen, without affecting formatting should also be provided for people who find it hard to read small text. Eventually, design implications for the problem solving, planning, reasoning and decision-making processes, can be to provide additional hidden but easy accessed information for users who wish to understand more about how to carry out an activity more effectively.

5.3 The Mind as an Information Processor

An approach from cognitive psychology is the idea that the mind is an information processor (Preece, Rogers and Sharp, 2002). Information is thought to enter and exit the mind through a series of ordered processing stages as illustrated in figure 5.3.1.

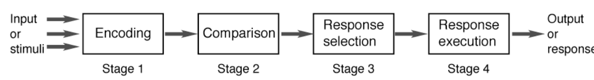

Figure 5.3.1: The human information-processing model (Preece, Rogers and Sharp, 2002)

Within these stages, various processes are assumed to act upon mental representations. Processes include comparing and matching. Mental representations are assumed to comprise images, mental models, rules, and other forms of knowledge. The information processing approach is based on modelling mental activities that happen exclusively inside the head. However, most cognitive activities involve people interacting with external kinds of representations, for instance computers. It is therefore of importance to study in which context the interaction takes place, and how structures in the environment can both aid human cognition and reduce cognitive load. In small GUIs this is a substantial point because of their portability, and the opportunity for using physical space and proximity as context indicators (proximity to patient displays the patient data). A major issue in using context indicators, is ensuring that the data for the patient in question is the data displayed, and not the data for the patient in the next bed.

The perception of information is only a part in the successful design of small GUIs. In the following, the extraction of useful knowledge will be presented.

6.0 CONCLUSION

In the first section of the paper, five questions were formulated:

- How should the interface be organized to present the information clearly?
- What information should be presented?
- What guidelines exist for designing interfaces?
- How to design for usability?
- How do users perceive the information presented?

In this paper, these questions have been discussed and answered. The first question is answered with respect to structure, organization and consistency. The overall structure of the screen is the first visual clue presented to the user. To make this first visual clue more eloquent, the gestalt principles can be applied to elements to indicate their place in the hierarchy of information. Structuring the entire display according to the techniques of using symmetry to ensure balance, using alignment to establish visual relationships, make use of optical adjustment for the human vision, and shaping the display with negative space (Mullet and Sano, 1995) can generate a clear organization. In using a design program for all the different screens in an application, consistency can be achieved. Structure, organization and consistency within an application, will provide users with knowledge of the interface, and allow them to orientate themselves and respond to the tasks more quickly once they have reached the level of experienced users. In a small GUI, a good organization is essential to avoid clutter.

The second question, what information to present, has the short answer "Less is more"; user interfaces should be simplified as much as possible (Nielsen, 1993). Nothing but exactly the information the user needs to perform a specific task should be presented in the specific screen. In finding the information highly relevant for the task, performing a HTA can be a useful process. In addition, all elements to put in the interface must be evaluated critically, since every pixel is precious in a small screen, and every "extra" element is a piece of information the user will have to look at when searching the screen for the relevant information.

General guidelines for interface design concern the use of colour and type. The overall issue with use of colour in screen is to not exaggerate. The background should be calm and pleasing to look at, and strong colours should be used with care. Due to the percentage of people being colour-blind, the interface should be able to use without the colour coding. For the use of type on the screen, bitmapped sans serif typefaces should be used to improve the legibility, and the type size set to at least 10 points. Legibility is highly relevant in medication to avoid misreading names of medicine, and thus imposing a danger to the patients. Maintaining the minimal type size even though the display decreases in size, is essential for the legibility.

To design for usability the keywords are visibility, feedback, constraints, mapping, consistency and affordances (Norman, 1988). These attributes should be a part of the interface revealing what elements can be manipulated, providing feedback to keep track of what action has been done, restricting user activity that should not be performed at a certain stage, relating controls and their effects in the world, use similar elements and have similar operations to achieve similar tasks, and to give a clue about how to use an object. In the evaluation of a system, its attributes can be compared to the desired qualities of the ten usability heuristics (Nielsen, 2001).

Two terms in understanding how the user perceives the information displayed are experiential and reflective cognition (Norman, 1993). The mode of experiential cognition is one of expert behaviour, the patterns of information are perceived and assimilated and the appropriate responses generated without apparent effort or delay. The mode of reflective cognition is that of comparison and contrast, of thought and of decision making. The balance between experiential and reflective cognition in the medication system is vital. Simple tasks should not take amounts of time, whereas tasks requiring use of knowledge and skills should provide the opportunity to use reflection instead of simply filling in a box.

All the subjects discussed convey different aspects of designing for screen. These aspects become even more pronounced when designing for small screen devices, after all a typical screen size for a PDA is limited to 240x320 pixels, or 3,8" (5,76 cm x 7,68 cm)(HP iPAQ h5550).

REFERENCES

Literature

Bolter, Jay David and Grusin, Richard, ”Remediation. Understanding New Media”, MIT Press, 2000

Dourish, Paul, “Where the action is – The foundations of embodied interaction”, MIT Press, London, 2001

Götz, Veruschka, “Color & Type for the Screen”, chapter 4, page 68-99, RotoVision, 1998

Mullet, Kevin and Sano, Darrell, “Designing Visual Interfaces – Communication Oriented Techniques”, Sun Soft Press, California, 1995

Nielsen, Jakob, “Usability Engineering”, chapter 2 & 5, Academic Press, Inc., 1993

Norman, Donald A., “The design of everyday things”, Doubleday, 1988

Norman, Donald A., “Things that make us smart”, Addison-Wesley, USA, 1993

Preece, Jennifer, Rogers, Yvonne and Sharp, Helen, “Interaction Design – Beyond human computer interaction”, John Wiley & Sons, Inc., USA, 2002

Tufte, Edward R., “Visual Explanations”, Graphics Press, Cheshire, Connecticut, 1997

Tufte, Edward R., “Envisioning Information”, Graphics Press, Cheshire, Connecticut, 1990

Wertheimer, Max, “Principles of Perceptual Organization”, p.115-135 in Readings in Perception, Selected and Edited by Beardslee, David C. and Wertheimer Michael, D. Van Nostrand Company, Inc., 1958

Internet Resources

Nielsen, Jakob, ”Ten Usability Heuristics”, available at www.useit.com/papers/heuristic, 2001

Lectures

Skogen, Martha Rice, ” Design Principles for Electronic User Interfaces - Especially in small-screen format”, 06-10-2004

Personal Communication

Skogen, Martha Rice

Øritsland, Trond Are

Unpublished material

Fagerholt, Henrik, Natural Language Generation of patient’s medical status in an electronic healthcare record, 2003

Figures

Figure 2.1.1: Mullet, Kevin and Sano, Darrell, "Designing Visual Interfaces – Communication Oriented Techniques", Sun Soft Press, California, 1995

Figure 2.2.1: Mullet, Kevin and Sano, Darrell, "Designing Visual Interfaces – Communication Oriented Techniques", Sun Soft Press, California, 1995

Figure 2.1.2: Mullet, Kevin and Sano, Darrell, "Designing Visual Interfaces – Communication Oriented Techniques", Sun Soft Press, California, 1995

Figure 2.3.1: Mullet, Kevin and Sano, Darrell, "Designing Visual Interfaces – Communication Oriented Techniques", Sun Soft Press, California, 1995

Figure 3.0.1: Fagerholt, Henrik, "Natural Language Generation of patient's medical status in an electronic healthcare record", Hovedoppgave NTNU, 2003

Figure 3.4.1: Preece, Jennifer, Rogers, Yvonne and Sharp, Helen, "Interaction Design – Beyond human computer interaction", John Wiley & Sons, Inc., USA, 2002

Figure 5.3.1: Preece, Jennifer, Rogers, Yvonne and Sharp, Helen, "Interaction Design – Beyond human computer interaction", John Wiley & Sons, Inc., USA, 2002