

An Experimental Comparison of Two Popular PDA User Interfaces

Teresa Hübscher-Younger, Roland Hübscher, Richard Chapman

Department of Computer Science and Software Engineering
Auburn University

September, 2001

Abstract

In this paper we compare the usability of the user interfaces of the two most widely used operating systems for personal digital assistants: Microsoft Pocket PC and Palm Computing PalmOS. In our experiment, we presented users with a sequence of six tasks that exercise the basic personal information management (PIM) functionality of the built-in applications and the basic operating system. Users were observed and questioned as they performed the tasks, and data about their actions was manually gathered by a human observer. Our results show significantly faster times when users perform the tasks with a PalmOS device versus a Pocket PC device. A majority of our test participants also stated a preference for the PalmOS device when asked about their impressions of the two interfaces afterward.

1 Introduction

Which PDA is the best? This seems to be debated endlessly in on-line forums and consumer electronics periodicals. We certainly don't purport to answer such a question here. The functionality of the various PDA's varies so widely that an overall quality comparison seems futile. However, there is some common ground among the market-leading platforms, and we aim to introduce a bit of objectivity into the debate about PDA usability by actually measuring user performance on some fundamental, common tasks that virtually every PDA user will perform.

Most PDA's provide a set of built-in tools that include a datebook, to-do list, address list, and memo pad (though the names for these tools vary from device to device). This personal information manager (PIM) functionality is implemented on a wide range of hardware devices, but most devices currently run one of two operating systems, and consequently present the user with one of two user interfaces. The two dominant operating systems in the marketplace are Microsoft Corporation's Pocket PC (Windows CE 3.0) operating system and Palm Computing's PalmOS v3.5.

Both OS'es include PIM functionality via a set of included applications that come pre-installed on the PDA devices. It seems that a comparison of the usability of the two platforms should begin with a comparison of the usability of the basic PIM functions. In this paper we identify a number of metrics for evaluating and comparing the two operating systems, we present a methodology for gathering data from a group of users, and we present our results.

2 Related Work

It would appear that very little work has been done in experimental analysis of user interfaces for handheld devices. Several books exist that guide user interface developers, but they base their



Figure 1: A Pocket PC device



Figure 2: A PalmOS device

advice either on general principles or inference from design techniques for interfaces on non-handheld devices [Rhodes99, Muench00, Want95].

The different physical form of PDA's (versus desktops) and differences in the way PDA's are used suggests that inferences made from desktop user interface design might be invalid. Consequently it seemed reasonable to compare the Palm OS interface, developed *ab initio* for handheld devices, with the Pocket PC interface, which attempts to provide users of Windows on desktops with a reasonably familiar, or at least recognizable interface.

While we found that devices running the two operating systems were physically very similar (Figures 1,2) and provided very similar sets of built-in PIM functionality, the design philosophies behind the two interfaces turned out to be completely different. While the Palm OS generally provides a single way, or perhaps two ways to perform a given task, or to navigate from one point in the interface to another, the Pocket PC devices possess many different ways of accomplishing most tasks (see the Methodology Section for an explanation of how this was uncovered during the course of designing the experiment).

The approach taken by the Palm OS is consistent with the quite common approach of underlaying the interaction with one simple and consistent conceptual model [Newman95]. The simplicity of this interaction model will then enable the user to easily predict what outcome each of her actions will have. Thus, the user has little difficulty choosing one of the very few "right" ways of performing a task.

Although this approach does indeed lead to highly usable interfaces if applied as part of an overall design process involving additional elements such as early participation of users and for-

mative evaluation [Dix98], some researchers are somewhat critical of this purist, one-simple-model approach, because it may lead to an overly constraining interface. The contrasting approach, called permissive interfaces, seems to describe the Pocket PC's design philosophy the best.

Permissive interfaces allow a wider range of actions with the idea that "every reasonable action should lead to a reasonable results" [Thimbleby95]. For instance, adding an event to the calendar can be done by either adding the event and then moving it to the right date, or going to the day of the event and adding it there. A permissive interface would necessarily allow either order. Of course, a permissive interface's design is also based on conceptual models because otherwise there can be no "reasonable action." Such an interface may even support several conceptual models, which must, however, be consistent as not to disorient the user.

We do not know if the Pocket PC developers were familiar with the concept of a permissive interface. Nevertheless, the Pocket PC's interface looks, at first sight, like a permissive interface. Many different ways exist to accomplish a task. Almost every action performed by the user will cause some action in the interface. However, our data and observations suggest that the interface is not a permissive interface as described by Thimbleby. First, some reasonable actions lead to surprising results, implying that the underlying model was not obvious to the user. Second, for some participants, simple tasks could not be executed because there was no obvious reasonable (to them) action accomplishing the task.

3 Methodology

3.1 Participants

Twelve members of the Department of Computer Science and Software Engineering at Auburn University (a mix of faculty, graduate students and undergraduate students) volunteered to participate in the study. Two of these volunteers participated in a preliminary run of the tests to help refine our methods. Ten of the volunteers participated in the actual pilot study. Volunteers were solicited from the participant population via class and email announcements during spring semester of 2001. The volunteers were treated in accordance with the *Ethical Principles of Psychologists and Code of Conduct* [APA92]. The participants had a wide range of experience with PDAs, from novice to advanced users.

3.2 Materials

Each experiment had the volunteer use two devices, one with a Palm OS and with the Pocket PC OS. There were four PDAs used during the experiments: a Palm IIIc, Palm IIIx, a Compaq iPaq and an HP Jornada. It was recorded which machines the participant used during an experiment.

Based on our experience with the preliminary two subjects, we found that the six tasks (see Design and Procedures section) took about 90 minutes, and that some sort of a detailed checklist or script for the observer to use made data collection much more reliable.

To aid in gathering the sequence of actions taken by a participant to complete a task, we developed flow charts showing possible routes a user might take through the interface in the performance of a task. An example of such a flow chart is shown in Figure 3, which depicts the possible routes a user of an Compaq iPaq running Pocket PC might take to start the calendar application, starting with the device turned off.

Based on the flow charts, we developed a checklist for each task. An example of the checklist to look up an appointment on a Palm OS device is included in Figure 4. This was one of the shortest checklists. The longest checklist, for entering a repeating appointment, was two pages long.

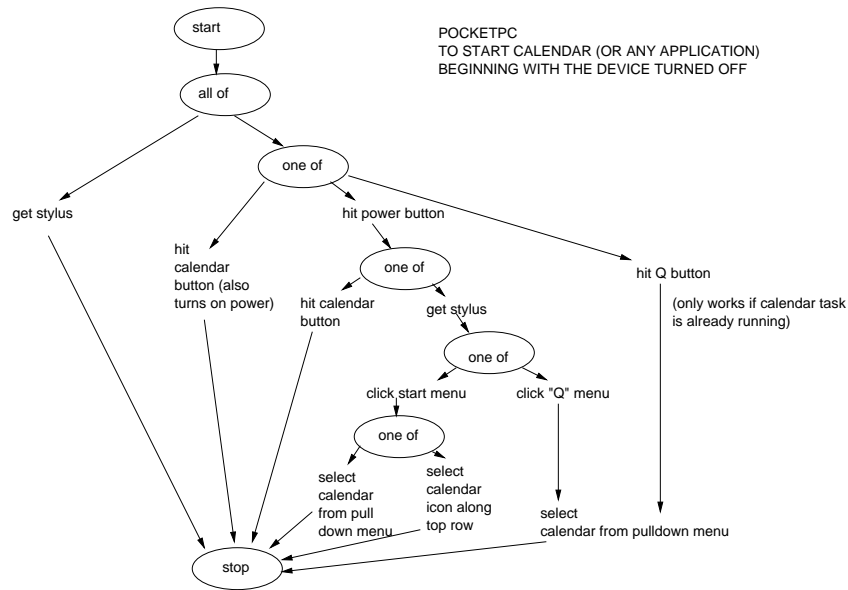


Figure 3: Flowchart for starting an application on a Compaq iPaq

Look up whether there is an appointment on 3/17/2001 at 2:00PM

- Gets to correct day (choose one of)
 - Clicks on “Go To”
 - * Accesses month March
 - * Selects day 17
 - Clicks month icon/view
 - * Scrolls to month March
 - * Selects day 17
 - Presses calendar button until gets to month view
 - * Scrolls to month March
 - * Selects day 17
- Reads what is occurring at 2pm

Figure 4: Checklist for reading a PalmOS memo

3.3 Design and Procedures

The experiment occurred over the duration of a week. The participants were first introduced to the observers and procedure of the experiment. They then filled out a short questionnaire that determined the age, gender, left- or right-handedness, and familiarity with PDA's of the test participants. We believed that we might see some correlation between any of these factors and the facility the participant demonstrated with the PDA's under examination.

From the questionnaire, it was determined which PDA operating system was most familiar to the participant. The participant was then given the PDA with the operating system they were most familiar with first to do the set of tasks. If the participant is either familiar with both or unfamiliar with both interfaces, the order in which the Pocket PC and Palm interfaces were presented is randomized. The participants completed six tasks on one PDA, and then they did the same six tasks on the other PDA.

The six tasks are

- Enter an appointment in the datebook or calendar application
- Look up an appointment
- Look up a text document and read it
- Enter a short text document
- Enter a person's address, phone number, company, and email in the contact list
- Look up a phone number, given a name

Each task was presented in the order above and presented one at a time on index cards. We also developed a set of rules to guide the experimenter in conducting an observation of a participant. These rules included:

1. The experimenter would not provide even absolute beginners with any instruction about how to use the device, even about how to turn it on or find the stylus.
2. The experimenter would not give the participant the answer to a question about how to do something until the participant asked several times.
3. If the experimenter was questioned several times about something, and it appeared the test could not go on without the participant being given more information, then that would be recorded as a failure to complete the task, but the participant's question would be answered and the participant allowed to continue.

For each task, we measured the time to complete the task, logged any errors committed during the performance of the task, logged any comments made by the user during the performance of the task, and recorded the particular sequence of actions the user took to accomplish the task. In addition we measured the time it took a user to initially turn the device on and find the application for the first task.

After the participant completed both sets of tasks, the experimenter asked for feedback about the tasks performed with the two interfaces. All the participants were asked to say which device they would recommend (based on usability) to a friend with only a small amount of experience with PDA's.

	Palm OS device	Pocket PC device
Starting the device task	14	55
Appointment entry task	351	520
Appointment lookup task	31	84
Note entry task	26	115
Note lookup task	244	328
Contact information entry task	204	204
Contact information lookup task	28	27

Table 1: The mean time in seconds for doing each task with the different devices.

4 Results

The overall results imply that users are significantly more efficient with the Palm OS than with the Pocket PC operating system. Multiple regression analysis techniques were used to look at the how the following factors affected the time on tasks: different handheld operating system, level of user experience with handheld devices, the order in which the different handheld machines were used, and manufacturer of Pocket PC device (Compaq iPaq or HP Jornada). The regression model for the time it took to do all the tasks was significant $R^2(20) = .58, p < .05$.

The operating system had a significant effect on the time it took to do all the tasks ($F(1, 20) = 3.55, p = .08$). The participants had a mean time of 14 minutes and 59 seconds when using the Palm handheld machines and mean time of 22 minutes and 13 seconds when using the Pocket PC handheld machines.

For the analysis on the individual tasks, the model for the time it took to do the appointment entry task was significant, $R^2(20) = .60, p < .05$. It was the first task in which the participants had to use an application on the device. While the effect of the operating system is not quite significant for this time ($F(1, 20) = 2.01, p = .18$), the effect of the order in which the operating systems were used to do the task was ($F(1, 20) = 3.79, p = .07$) at 0.1 alpha level. If the participant did this task first with the Palm the mean time to do the task was 3 minutes and 53 seconds. If the participant did this task first with the Pocket PC, the mean time to do the task was 12 minutes and 19 seconds. Since this was the first task they did, the results imply that it took longer to get accustomed to working with the Pocket PC device than it did to get used to working with the Palm device.

The effect of the operating system used in the device was also significant for the tasks for looking up an appointment and looking up a memo or note ($F(1, 20) = 3.90$ and 3.44 , respectively, $ps < .1$). In both cases, it took less time to do this task with the Palm device ($M(10)s = 31$ seconds and 26 seconds) than it took to do with the Pocket PC device ($M(10)s = 1$ minute and 54 seconds and 1 minute and 24 seconds) For all the tasks the mean time for doing the task with the Palm device was smaller than the mean time for doing the task with the Pocket PC device, except in the case of the last two tasks concerning working with contact information (Table 4).

5 Discussion

Novices had difficulty starting to use either device, particularly when it came to text input. Often they believed they had only one option, using either a keyboard or graffiti.

However, users of all experience levels had difficulty understanding the Pocket PC Windows-based interfaces when doing the tasks. The first task, using an application in which the participant enters a repeating appointment, was difficult primarily because of the use of unintuitive menu

choices, such as “Edit Pattern” for repeating a weekly event with an end date. The use of a wizard to actually edit the pattern also caused some difficulty. Surprisingly, entering a repeating appointment on the Pocket PC turned out to be the most difficult of all tasks in terms of steps required and number of possible errors.

Another interesting common problem observed on the Palm PC tasks was difficulty in finding the stylus. Participants would work with their finger or another pen at first. All the participants found the stylus quickly when using the Palm devices.

With the Pocket PC devices, finding the appropriate application to use for a task was often a problem. The number of similar applications for doing a task was higher for the Pocket PC devices than for the Palm devices. To enter or find an appointment, participants frequently went to the “Today” application and clicked on “no upcoming appointments” hoping to find a way into some sort of calendar application. Participants were confused when they had both a Notes and Microsoft Word application. It wasn’t clear why both would be needed and that one could not access the same files through both applications.

We found in general, the Pocket PC devices had several more entry points into applications, as well as possessing a larger number of applications. Sometimes this worked to the advantage of the participant, as in the “Today” example, but often it confused the participant, who wasn’t sure where to start and wasn’t sure they understood all their options.

The use of modes also confused participants with the Pocket PC devices. When the participants started to create a memo, they were in the “Paint” mode, where their pen movements on the screen are recorded in an image, rather than as text. However, often participants assumed that they were not in a special mode and that writing would work as it did in the graffiti writing area. Often, participants were confused when they find themselves already in an edit mode when accessing an application.

Most participants appreciated and used the word completion feature on the Pocket PC devices, and the test administrators thought that this feature would probably help improve the times. Still, the mean times show that on the task which required the most text entry, the note entry task, using the Palm device led to a lower mean time.

The user comments generally reflected the results of the study. Eight of the 10 participants reported that they found the Palm more usable and would recommend it to novices looking for a handheld device.

6 Future Work

The next step is a large scale study with significantly more test participants. One participant made a comment which might lead to an interesting experiment – she said, “I think I could do these tasks faster with a pencil and notepad than anyone could possibly do them with a PDA.” We plan to investigate this claim. Further experiments could test add-on functionality that is not included with the operating system, such as audio and video programs, or tools to access online content.

Acknowledgements

We would like to acknowledge the support of the Auburn Institute for Reconfigurable Smart Components, which funded the purchase of the PDA’s used in our study, and also to acknowledge the assistance of Jean D’Amico and Patricia Lanford in interviewing the test participants. Pocket PC

and Windows CE are trademarks of Microsoft Corporation. Palm OS is a trademark of Palm Computing, Inc. iPaq is a trademark of Compaq. Jornada is a trademark of Hewlett-Packard.

References

- [APA92] American Psychological Association, “Ethical Principles of Psychologists and Code of Conduct”. *American Psychologist*, 47, 1597-1611.
- [Dix98] Dix, Alan, Janet Finlay, Gregory Abowd, and Russell Beale, *Human-Computer Interaction, 2nd ed.*, Prentice Hall Europe, 1998.
- [Muench00] Muench, Chris. “The Windows CE user Interface”, *The Windows CE Technology Tutorial*, Addison-Wesley, 2000, pp.157-262.
- [Newman95] Newman, William M. and Lamming, Michael G., *Interactive System Design*, Addison Wesley, New York, NY, 1995.
- [Rhodes99] Rhodes, Neil and Julie McKeehan. *Palm Programming: The Developer’s Guide*. O’Reilly, 1999.
- [Thimbleby95] Thimbleby, Harold, “Permissive user interfaces” *International Journal of Human-Computer Studies*, Academic Press, 54, 333-350,2001.
- [Want95] Want, Roy, Bill Schilit, Norman Adams, Rich Gold, Karin Petersen, David Goldberg, John R. Ellis, and Mark Weiser. “User Interface Design for Palm Sized Computers.” *The ParcTab Ubiquitous Computing Environment*, Technical Report CSL-95-, Xerox Palo Alto Research Center, 1995.