Available online at www.sciencedirect.com
ScienceDirect
Computers in Human Behavior 24 (2008) 1246–1274
Computers in Human Behavior
www.elsevier.com/locate/comphumbeh

Landmarks or surveys? The impact of different instructions on children's performance in hierarchical menu structures

Susanne Bay, Martina Ziefle *

RWTH Aachen University, Jaegerstr. 17-19, 52064 Aachen, Germany

Available online 27 June 2007

Abstract

Which kind of instruction helps children aged 9–14 years interact efficiently with a mobile phone? Due to analogies between navigation in menu structures and the natural environment, three instructions providing different forms of spatial knowledge were under study: A step-by-step instruction featuring landmark knowledge of the menu functions to be selected, a diagram of the menu structure providing survey knowledge and a free exploration of the menu, also giving the children the opportunity to develop survey knowledge. Results show a superiority of the two instructions that provide survey knowledge, except for the youngest children aged 9–10 years. This group showed to have lower spatial abilities and is therefore presumably not able to understand and integrate this type of knowledge. For those very young children, the landmark information given in traditional step-by-step instructions is more helpful. It is concluded, that simple diagrams of the menu structure can help children from 11 years on to significantly ease their interaction with small menu driven devices.

Keywords: Children; Instructions; Manual; Mobile phone; Spatial orientation; Mental representation

* Corresponding author. Tel.: +49 241 8096130; fax: +49 241 8092318.
E-mail addresses: Susanne.Bay@psych.rwth-aachen.de (S. Bay), Martina.Ziefle@psych.rwth-aachen.de (M. Ziefle).

0747-5632/$ - see front matter
doi:10.1016/j.chb.2007.05.003

1. Introduction

In spite of the fact that more and more children all over the world possess mobile phones, little is known about the difficulties they experience and the factors hampering or benefiting children's purposeful interaction with these devices. In Germany nearly 50% of the children aged 10–13 (Informationszentrum Mobilfunk, 2004) and in the UK 5.5 million school children have their own mobile phone (Mobileyouth.org, 2005). Mobile phones, very prototypical devices standing for a whole class of devices with small screens but a high complexity of functions implemented, can impose considerable cognitive load on users. As was found, users often delve into distraction having difficulties finding the targeted function and losing their bearings in the phones' menu. This was observed in users of all ages, but was disproportionately more pronounced in older adults (Ziefle & Bay, 2004, 2005, 2006) as well as in children (Bay & Ziefle, 2005; Ziefle, Bay, & Schwade, 2006). In very recent studies focusing on children's interaction with mobile phones, it was shown that children aged 9–14 years who do not have much experience using mobile phones have major difficulties carrying out simple and very basic phone operations (Bay & Ziefle, 2005; Ziefle et al., 2006). They needed, for example, on average more than 2 min to effect a call to a number stored in the internal phone directory and more than 4 min to change a setting of the phone in order to hide their own number when calling someone.

Naturally, manufacturers should aim at designing mobile phones that are easy to use by all age groups. Yet, as so many different factors are involved in the ease of use of mobile phones (i.e. naming and categorization of functions, design and functionalities of navigation keys, number of functions implemented, depth and breadth of the menu structure, information presentation on the screen, etc.), one easier and faster option to approach this problem is to design simple and understandable instructions. There are three questions to be answered: First, how should such an instruction be designed in order to be easily understandable? Secondly, can any instruction be as or even more helpful than an active exploration of the functionalities which is the most common way for children to interact with a new technological device (e.g. Bay & Ziefle, 2005; Beckwith et al., 2006)? Third, do mobile phone instructions for children have to be specifically designed taking into account children's less sophisticated cognitive abilities due to their ascending developmental status? This is of major concern as, to our knowledge, no study so far focused on this question even though many studies have dealt with the utility of different instruction forms and training tutors, e.g. for word processing programs, for adults (e.g. Carroll & Mack, 1985; Frese et al., 1988; Frese, Brodbeck, Heinbokel, & Mooser, 1991; Ziefle, Bodendieck, & Kuenzer, 2004).

To create instructions that are easily understandable by children, long texts seem not to be appropriate as reading abilities in children are not only less sophisticated compared to adults, but also differ considerably among children. Furthermore, it was shown that adults do not spend much time reading manuals, either (e.g. Carroll, 1990), presumably because the study of text heavy instructions may be conceived as time-consuming and bothersome. Addressing this problem, Carroll (1990) proposed an approach of "minimalist instructions" that is characterized by three aspects: (1) learners should be allowed to start immediately on meaningfully realistic tasks, (2) the amount of reading should be reduced and (3) errors and error recovery should be made less traumatic and more pedagogically productive. Therefore, graphical illustrations providing the information's inherent structure with a very reduced amount of text to be studied should be very valuable. Yet, there is an abundance of research describing the ineffectiveness of detailed textual instructions but there has not been a test on diagrammatic instructions (Rodriguez & Polson, 2004). A study by Rodriguez and Polson (2004) showed that the most severe failure in designing diagrammatic instructions is the lack of unambiguously picturing the initial step of a subtask including the location and object of a required action. Transferring this concept to the design of appropriate instructions for menu-driven interfaces, as the mobile phone, it must be crucial to tell the user *where within the menu* a certain task has to be accomplished.

To get an idea of how to sketch the information that is relevant for the user, an insight into the cognitive processes involved when interacting with a mobile phone may be helpful. In a number of studies it is assumed that the orientation in the menu of technical devices resembles the orientating behavior in the real world (e.g. Bay & Ziefle, 2003b; Downing, Moore, & Brown, 2005; Kim & Hirtle, 1995; Norman, 1994; Vicente, Hayes, & Williges, 1987; Ziefle & Bay, 2004, 2006). Though, this assumption needs further evidence, as a completely two-dimensional menu of a technical device is qualitatively different from the three-dimensional information given in the real world. From studies focusing on spatial orientation in the natural environment (e.g. Allen, 2004; Taylor & Tversky, 1992, 1996; Thorndyke & Goldin, 1983), it is known that users adopt different perspectives when orientating in space. Thorndyke and Goldin (1983), for example, proposed three spatial perspectives as crucial for proper navigation: Landmarks (outstanding points in the environment), routes (the paths connecting different landmarks) and survey knowledge (a graphical outline of a terrain, i.e. a bird's eye view of the environment). As the orientation in space combines all types of perspectives, users may take up a supra-position, where a fixed point outside the environment is adopted and structural and survey information is retrieved. The locations of landmarks and buildings are learned relatively to one another in terms of cardinal points. But also route knowledge is essential. In combination with landmarks, routes serve as important knowledge base for paths that are repeatedly followed by a person.

Transferring this concept to menu navigation in technical devices it is argued that interaction with menus requires the users to take up similar perspectives. Interacting with hypertext or hierarchical menus requests users to navigate through different nodes and links, also to decide at which "crossings" to "take left or right" and to remember specific function names as spatial cues for knowing where the current position is located relatively to where users came from and relatively to the point they are going to. Thus, when interacting with complex menu structures, it is assumed that users need to build up a mental representation of the "spatial" structure of that menu in order to orient themselves and interact with the device in a purposeful way. The knowledge of how a menu is spatially structured guides users in their search through different levels in different menu depths. The general usefulness of spatial knowledge for menu navigation performance was corroborated by empirical evidence stemming from two major sources. (1) Consistently, it was found that persons having high spatial abilities showed a distinctly higher performance interacting with technical menus than persons with lower spatial abilities (e.g. Bay & Ziefle, 2003b, 2004; Downing et al., 2005; Egan, 1988; Kim & Hirtle, 1995; Vicente et al., 1987; Westermann, 1997; Ziefle & Bay, 2006). In small screen devices, where the screen size is very limited and the overall structure of the menu is not transparent, spatial abilities may be even more crucial, because the orienting process is not visually supported and most of the menu functions are most of the time hidden from sight. Thus, it is assumed that spatial abilities specifically advantage navigation performance by supporting users in constructing a proper mental representation of the systems' structure (e.g. Mayer, 1986; Sein, Olfman, Bostrom, & Davies, 1993; Ziefle, Arning, & Bay, 2006). (2) It was shown that users, who had a correct representation of the overall structure of the phone menu (survey knowledge) were specifically advantaged over those who had an inappropriate representation. Furthermore, users, who had a correct representation of the correct path to be taken (route knowledge), and also knew salient features of the menu branch to be taken (landmark knowledge) showed higher effectiveness and efficiency of menu navigation in mobile phones (Bay & Ziefle, 2003a, 2003b; Ziefle et al., 2006; Ziefle & Bay, 2004).

If when navigating through the hierarchical menu structure of a mobile phone comparable processes to the ones involved in spatial orientation in the natural environment take place, the development of the different kinds of spatial knowledge should be supported by appropriate instructions. In order to build up survey knowledge, a map of the menu structure could be provided. For large computer displays and adult users the efficiency of such a map has already been demonstrated (e.g. Billingsley, 1982) and recently also for small screen devices such as the mobile phone (Bay, 2003; Bay & Ziefle, 2004). The benefit of a navigation aid based on survey knowledge was found to be still more useful for persons having high spatial abilities (Bay & Ziefle, 2004; Ziefle & Bay, 2006). Conceptualizing this specific advantage it is assumed that persons with high spatial abilities are able to construct a mental representation of the systems' structure during navigation (Sein et al., 1993) and to incorporate the survey knowledge provided by the map into their mental representation. In contrast, persons with low-spatial abilities have much more difficulties building up an appropriate mental representation of the system's structure and, therefore, do not or to a much lesser extent take advantage of the map's survey knowledge.

But can children also profit from a map of the menu containing survey knowledge? Presumably, maps can only be understood as soon as the child has developed a concept of space. Siegel and White (1975) describe children's development of spatial representations of large-scale environments as follows: First landmarks are noticed and remembered. Once landmarks are established, the child's acts are registered and accessed with reference to them. Then, clusters of landmarks are formed to "mini-maps", which include routes between landmarks. A key issue in the development of spatial representations is the child's formation of some kind of objective frame of reference. Finally, survey maps appear as coordinates of routes within an objective frame of reference. According to a study by Shemyakin (1962) this last stage of development is attained at about 12 years of age.

Taking theses findings into account, it may be hypothesized that an instruction containing survey knowledge may help children in their purposeful interaction with a mobile phone menu, if spatial abilities are already developed properly. However, children younger than 12 years of age may have difficulties understanding and integrating the survey knowledge into an internal frame of reference, as they are not yet able to build up a mental representation and should therefore not profit from such an instruction. Presumably, a sequential aid (as mostly present in conventional technical leaflets of devices) in form of a step-by-step instruction of the different actions to be carried out one after another, referring to landmarks that have to be found, is more helpful for them. Though, it should be considered that the memorization of the single steps to be executed one after the other imposes considerable memory load which is also not sure to be handled properly by children or at least should be more difficult for the younger children.

On the other hand, an explorative interaction with a new device may be much more appropriate for children, since this conforms more to their usual behavior, preferring a natural and playful trial and error interaction style (e.g. Bay & Ziefle, 2005; Beckwith et al., 2006). Playful experimentation, also known as tinkering, is assumed to yield educational benefits because the users may incidentally gain knowledge of the system by exploring its structure. In so doing children may get to know which features and functions are available, where in the menu these functions are located and also how to activate the functions (Beckwith et al., 2006). Interestingly, tinkering is a behavior that is more often observed in male than female users (e.g. Beckwith et al., 2006; Van den Heuvel-Panheizen, 1999). Studies dealing with instructional learning in adults (e.g. Carroll & Mack, 1985; Frese et al., 1988, 1991; Ziefle et al., 2004) also strongly favor the explorative learning over different kinds of tutors. The specific advantage of exploratory learning is explained by the possibility of constantly and continuously shaping a mental representation of the systems' structure while interacting with the device. A cautionary note, though, is to be considered because the building up of a mental representation requires users to have the ability to establish a (spatial) frame of reference into which the knowledge gained by exploration can be integrated, and this again may not be yet fully developed in children, at least not in the younger ones. In terms of the development of spatial knowledge in children exploration may though be useful since all types of knowledge of the natural environment can be acquired through active navigation through the environment (Thorndyke & Hayes-Roth, 1982).

The present study attempts to answer the question which form of instruction is most helpful for children. Three different forms of instructions were examined: One type provided solely landmark knowledge (step-by-step instruction), the second type contained survey as well as landmark knowledge (map of the menu structure) and the third represented an exploration condition in which landmark, survey as well as route knowledge may be gained. Furthermore, a developmental perspective was taken into consideration by analyzing the performance of different age groups (9–10 year olds, 11–12 year olds, 13–14 year olds). In order to estimate at about which age children may take advantage of instructions and are able to use the mobile phones purposefully, it may be helpful to categorize the children according to their age and their status of cognitive development. Following Piaget (1983), the 9–10 year olds are assumed to be in the stage of concrete operations. They start representing operations mentally, thus being able to understand reversibility and that two dimensions sometimes have to be considered simultaneously as well as to begin understanding and solving hierarchical classification tasks. At the age of about 12 years – according to Piaget – children enter the stadium of formal operations. Cognitions become more systematic and children start isolating different variables when dealing with complex problems developing hypothetic-deductive strategies. Thus, it is assumed that the children in the stage of formal operations (about 12–14 years) should clearly outperform the children in the stadium of concrete operations (about 9–10 years) in the general competence of using the mobile phones. However, Piaget's clear-cut age limits were criticized in its rigid form and found to be modulated by cognitive acceleration processes (e.g. Rose, 1985; Shayer & Wylam, 1978), thus are to be treated with caution. Anyway, the mere chronological age classification has only a limited utility as age has to be regarded as a carrier of cognitive development. Therefore, in the present study, two major cognitive abilities were surveyed, which are expected to be of central importance for the usage of hierarchical menus in technical devices (Bay & Ziefle, 2003a; Ziefle & Bay, 2006): children's short-term (verbal) memory and spatial ability. In this context, it was of interest whether a certain level of spatial and memory abilities is necessary to bring the benefit of an instruction type into fore and if the cognitive abilities and the benefit by instruction type reflect the developmental status. Another point at issue was the sustainability of the different instructions or navigational aids. In order to determine learnability effects, the children had to process the same four mobile phone tasks twice consecutively. The instructions were presented only prior to the first trial, but children had to solve the second trial without being supported by an instruction. Thus, the differential effects of the instruction types and the interaction with the cognitive developmental status were aimed at.

2. Method

The study's central focus was the utility of different instruction types; therefore instruction types and their operationalization are described first. Then methodological details with respect to the independent and dependent variables, the sample studied and the experimental procedures are given.

2.1. Instruction types

1. The first type of instruction was a *step-by-step instruction* containing the different menu functions that had to be selected one after another in order to solve a task. The knowledge provided by this instruction is landmark knowledge as well as route knowledge (names of the functions to be visited one after another), without revealing the menu structure. Fig. 1 shows one example of such an instruction. Children were told that the functions shown in the step-by-step diagram are the key terms they had to memorize, as they are the cardinal points in the menu they should follow. This instruction was presented to the children for a period of 20 s prior to each task in the first trial.
2. The second form of instruction was a *menu structure* visualized in a graphical tree where the name and location of the functions that have to be selected to solve a task are given. This instruction type provides primarily survey knowledge, but contains also information on landmarks (the functions' names). Fig. 2 shows an extract of a menu tree containing the path and the menu functions (original size). In order to illustrate the whole information, which was presented, Fig. 3 shows the path and menu information for the four tasks the children received in this condition. Note that the information is greatly scaled down for demonstration purposes. In the experiment, this information was plotted on a sheet of paper, separately for each task, and pinned on the wall. The children studied the respective menu structure and the path that they had to go for 20 s before they started the completion of the respective task.The children were instructed that this structural map informs them about all the different paths and branches of the phones' menu as well as their interconnections. Children were told that the functions shown in the map were the key terms they had to memorize, as they are the cardinal points in the menu they should follow.

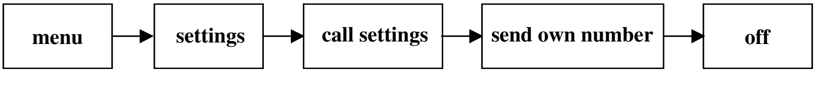

Fig. 1. Example of a step-by-step instruction.

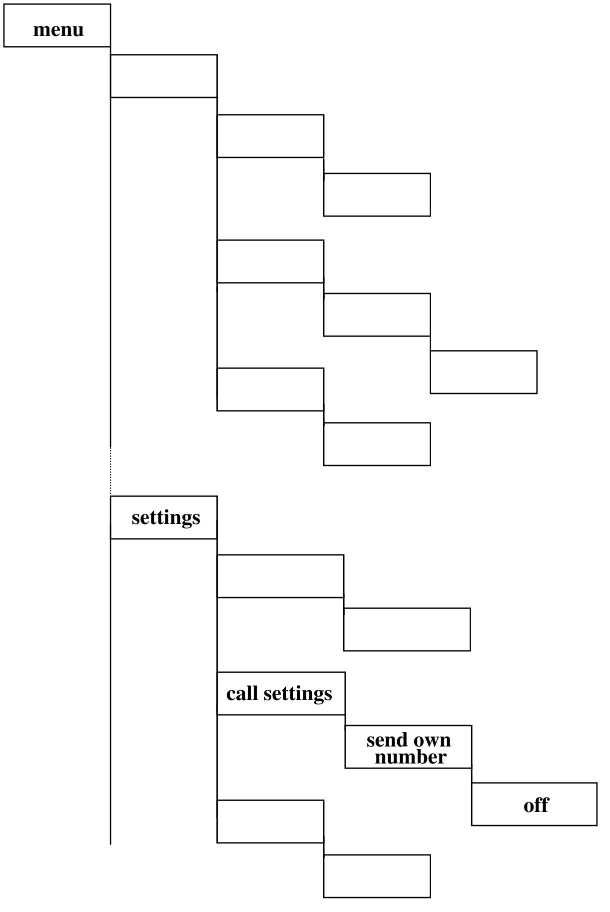

Fig. 2. Example of a menu structure instruction (original size).

3. A third type of instruction was adopted by giving the children the opportunity to *explore* the functionalities of the menu of the mobile phone. They were invited to explore the phone's functionalities by trying the different keys and navigating through the functions for a period of 5 min prior to solving the experimental tasks. The children did not receive further guidance, but were generally encouraged to capitalize on their possibility to get to know the mobile phone in advance. This instruction is assumed to provide all types of spatial knowledge, that is, survey knowledge, route knowledge as well as landmark knowledge.

2.2. Independent and dependent variables

Four independent variables were under study. The first independent variable was the type of instruction the children received. The second independent variable was users' age. Three age groups were examined: 9–10 year olds, 11–12 year olds and 13–14 year olds.

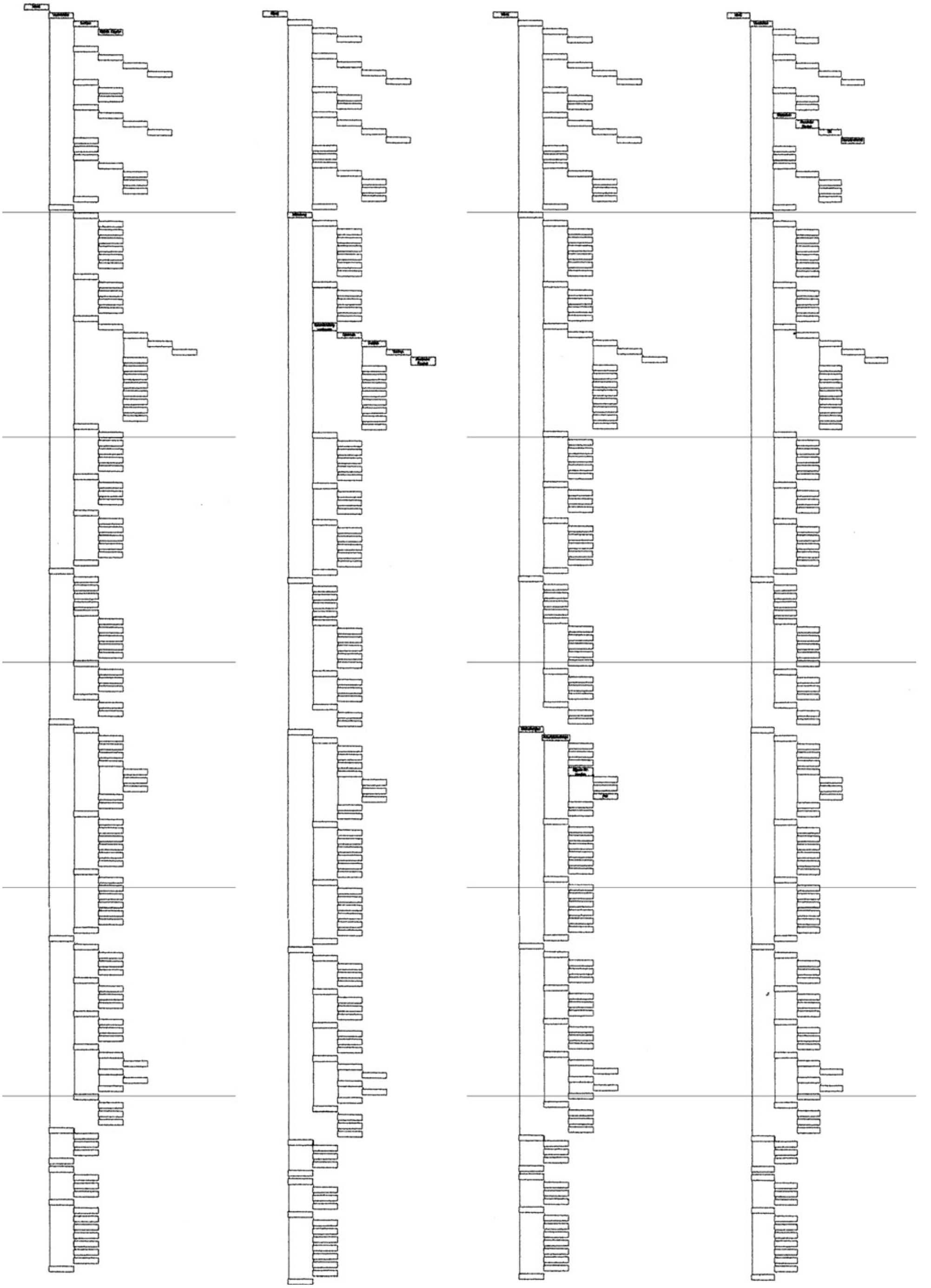

Fig. 3. Menu structures for the four different tasks as presented to the children (largely scaled down for demonstration purposes).

Furthermore, the children's spatial ability (third independent variable) and their verbal short-term memory capacity (fourth independent variable) were assessed. They were treated as between subject variables possibly affecting performance using the mobile phone and interacting with the instructions.

As dependent variables four different performance measures were analyzed. As effectiveness measure, (1) the number of tasks successfully solved was determined. Furthermore, measuring efficiency of menu navigation, (2) the time needed to process the tasks, (3) children's detour steps (number of keystrokes carried out that were not necessary when solving the task the shortest way possible) and finally, (4) hierarchical returns to higher levels within the menu indicating that children felt they had taken the wrong path in the menu hierarchy and needed to go back to a known position.

The number of tasks successfully solved and the time the children needed to process the tasks represent the two standard measures (EN ISO 9241-11, 1997). However, taking into account that children might loose their motivation to bring a difficult task to a successful end, effectiveness might not be an appropriate measure to sensitively reflect the different impact of the instruction type. The same might be true for the time on task, considering that younger children may act more cautionary and less snappy compared to older children. Therefore, the measuring scope was broadened: The number of detour steps, thus, the detouring of users within the shortest menu path possible and the frequency with which the children have to go back to higher menu levels in order to re-orientate, seem to have a high sensitivity to reflect the effects of the different navigation aids and their stability over time.

2.3. Tasks

Four basic and very common mobile phone tasks had to be solved by the participants. In total, 36 steps were necessary to solve the four tasks in the shortest way possible. To ensure that the children understood intellectually and semantically what they had to do in the different tasks, the tasks' instructions were tailored with respect to a child-friendly setting with examples taken from the children's experience:

(1) Calling a number stored in the phone directory (4 key strokes were needed to solve the task).
(2) Sending a text message. In order to control the differences in the speed of typing, the message was already provided and only had to be sent when the children reached the adequate point in the menu hierarchy (10 key strokes were necessary to process the task).
(3) Changing the setting of the mobile phone with the goal of hiding ones own number when calling someone (12 key strokes were necessary).
(4) Making a call divert to the mailbox (10 key strokes were necessary).

In order to measure learnability, all four tasks were presented twice. In the second run, the children were not given any instructional help. This was of specific significance as it was a central interest to analyze the sustainability of the respective instruction type. After a short break, each task was given a second time with slightly modified instruction (e.g. in the first trial, children had to deactivate the transmission of their phone number and in the second run they were requested to re-activate it).

The presentation order of tasks was held constant, from one to four in the first as well as in the second trial. It was of crucial importance not to ask too much of the children with respect to the tasks' difficulty. In order to ensure a sense of achievement, the more easy tasks were to be solved first, followed by more difficult tasks (as taken from previous findings, Bay & Ziefle, 2005). As it was found out in previous experiments, learnability effects from the first to the second run were not affected by the tasks' order. Therefore, it seemed appropriate to keep the order of tasks constant. Furthermore, the length of the experiment was tested prior to the experiment. Children's concentration should not be encumbered, especially as they did not only have to accomplish the phone tasks, but also the preliminary survey assessing the children's experience with technical devices and afterwards, the psychometric testing. Therefore, a pre-experiment with one 9-year-old child was carried out. The four tasks processed twice, plus the completion of the tests lasted about 80 min at the most. As the child was not overstressed but rather enjoyed the procedure, the experiment was assumed to have a reasonable length and difficulty.

2.4. Assessing user characteristics (spatial ability, memory and experience with technical devices)

Determining spatial and memory abilities, two subtests of the HAWIK-R (a standardized psychometric test for determining children's intelligence) were carried out: one test on spatial visualization ability (Fig. 4) and one test on short-term memory (Fig. 5).

In the test on spatial ability ("Mosaic Test") the experimenter showed the child a picture (cf. Fig. 4, right) and the child's task was to reproduce the picture using cubes with different patterns on each of the sides (shown in Fig. 4, left). Eight items from the Mosaic Test were selected comprising more easy spatial tasks where the children had to reproduce the respective form with four cubes, as e.g. the forms a and b in the upper row of Fig. 4, but also more difficult tasks with 12 cubes (forms c and d) and even forms with 16 cubes (not pictured here) were to be solved by the children. A maximum of 56 points could be attained.

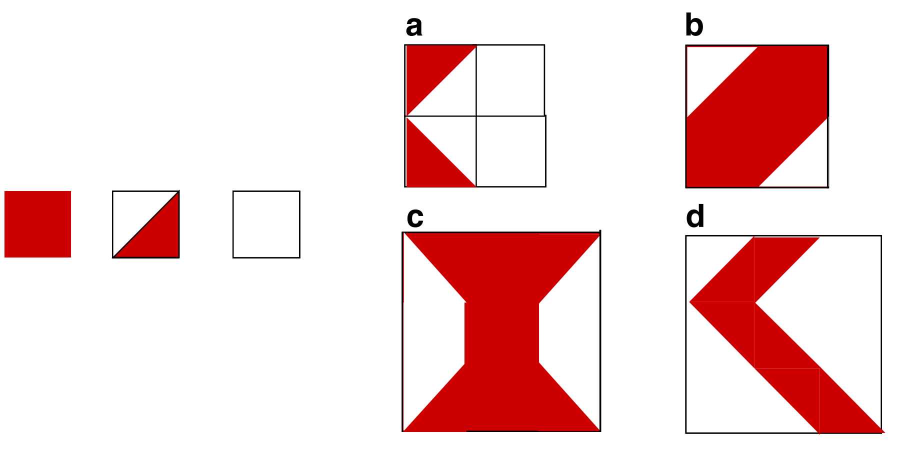

Fig. 4. The Mosaic Test to assess mental rotation ability (Tewes, 1983).

1. Trial	2. Trial
1. 3-8-6	1. 6-1-2
2. 3-4-1-7	2. 6-1-5-8
3. 8-4-2-3-9	3. 5-2-1-8-6
4. 3-8-9-1-7-4	4. 7-9-6-4-8-3
5. 5-1-7-4-2-3-8	5. 9-8-5-2-1-6-3
6. 1-6-4-5-9-7-6-3	6. 2-9-7-6-3-1-5-4
7. 5-3-8-7-1-2-4-6-9	7. 4-2-6-9-1-7-8-3-5

Fig. 5. Examples from the short-term memory test (recall of numbers) (Tewes, 1983).

The test on short-term memory required the children to verbally repeat a row of numbers read aloud by the experimenter. The test consisted of seven rows of three to nine numbers, which had to be reproduced by the children directly after the experimenter had read them aloud. Examples for this task are given in Fig. 5. The children were given two trials to correctly reproduce each row. The number of correctly reproduced rows was added to a final score. The maximum score to be reached was 14.

Furthermore, the children's previous experience with technical devices was determined. First, it was important to learn if age differences in performance are due to the age-correlated experiences with technical devices. Second, it was methodologically important to ensure that results can be referred to the effects of the three different instruction types. Therefore, a detailed pre-screening of participants' experience with mobile phones and other technical devices as well as the reported ease of using them was carried out. The children were to state if and how often they use technical products (mobile phone, PC, video cassette recorder (VCR) or a DVD player), using a five-point scale (1 = several times per day, 2 = once per day, 3 = once or twice a week, 4 = once or twice per month and 5 = less than once or twice a month). Further, the estimated ease of use when using different technical devices had to be stated again on a scale with four answering modes (1 = the usage is easy, 2 = the usage is rather easy, 3 = the usage is rather difficult and 4 = the usage is difficult). Moreover, all participants rated their general interest in technology, using a four-point scale (1 = low interest; 2 = rather low interest, 3 = rather high interest, 4 = high interest).

2.5. Apparatus

A mobile phone corresponding to the Nokia 3210 model regarding menu and keys was simulated as software solution, run on a PC and displayed on a touch screen (TFT-LCD Iiyama TXA3841, TN, 15′ with a touch logic by ELO RS232C). It was a general question which phone model was appropriate for experimental purposes. First, one aim was to take a real model in order to meet demands of ecological validity. Second, as our youngest children were just 9 years of age, it was of central impact that the complexity of the phone was not too high and the interface, from an ergonomic point of view, comparably easy to use. Otherwise, the utility of the different instruction types would not come into fore, being possibly veiled by a high difficulty when using the phone. Therefore, we chose the Nokia 3210 model, a mass model with a low complexity in both its menu structure and navigation keys, which could also be used properly by children, as shown in a previous study (Bay & Ziefle, 2005). The software prototype mirrored exactly the menu structure and the navigation keys given in the real phone. However, the keys and the display of the mobile phone were enlarged compared to the original model in order to exclude motor and visual difficulties of the children. To avoid biases and conceal the real brand the simulated phone had a very neutral and unobtrusive design. Three menu functions were presented on the screen at a time. In order to measure effectiveness and efficiency in detail and without disturbing the children while working, user actions were recorded online on the keystroke level. Thus, frequency and type of keys used, time spent on tasks and the navigational route could be reconstructed very precisely. In order to assure a comfortable and relaxed body posture while completing the tasks, the children sat at a table and worked on the touch screen, which was fixed to the edge of the table (the mounting equipment for fixating the screen was specifically constructed for experimental purposes). Fig. 6 shows a child using a simulated phone displayed on the touch screen (taken from a similar experiment).

As was observed, the children in fact had no difficulties with the simulated phone and the usage of the touch logic with their fingers. This was accomplished by giving the children the opportunity to work with the touch screen before they started to process the mobile phone tasks. The questionnaire assessing demographic data of the children prior to the experiment was presented on the computer too, and the children operated the touch screen while answering the initial questions thereby getting accustomed to the system.

2.6. Participants

Thirty children aged 9–14 years took part in the experiment. They were distributed equally to the three experimental conditions holding constant users' age. Thus, in each condition there were three children aged 9–10 years, three children aged 11–12 and four children aged 13–14 years. Nine children reported never to have used a mobile phone

Fig. 6. Photo of a child when working with the experimental mobile phone (taken from a similar experiment).

(three children in each of the experimental conditions), eight said they had already used the mobile phone of their parents and 13 children reported to possess a mobile phone of their own. Those children, who reported to have or use mobile phones, used the phones mostly for making and receiving calls and sending text messages.

To ensure that differences in performance can be traced back to the differential effects of instructions and age groups, respectively, instead of other prevailing distinctions, the children's previous history of technical usage was surveyed. First, the three experimental groups were analyzed with respect to the experience with technical devices (Table 1), using non-parametric tests (Kruskal–Wallis).

As found, the three experimental groups did not differ significantly with respect to the experience with technical devices (mobile phone: $z = 2.3$; $p > 0.05$; PC: $z = 0.3$; $p > 0.05$; VCR/DVD: $z = 0.2$; $p > 0.05$). Furthermore, the reported ease of using these devices (mobile phone: $z = 1.2$; $p > 0.05$; PC: $z = 1.5$; $p > 0.05$; VCR/DVD: $z = 0.08$; $p > 0.05$) and the rated interest in technology in general ($z = 1.9$; $p > 0.05$) was also not found to be different across the groups. Furthermore, the three experimental groups were also comparable with respect to children's spatial ability ($F(2,27 < 1$; n.s.)) and short-term memory capacity ($F(2,27 = 1$; n.s.)). Thus it can be assumed that user characteristics are equally distributed over the three experimental groups.

For the experience with technology, age differences will also be reported in this section. From Table 2 it becomes obvious that overall no age differences were present. All age groups reported to use the PC most often (once per day), whereas mobile phones and VCR devices were used more seldom (once or twice per month). With respect to the ease of using the devices, the mobile phone was judged to be most easy, and, according to children's statements, the PC and the VCR were "rather easy" to use. Finally, in all age groups, the interest in technology was rated as "rather high". Statistical testing revealed only for the ease of using mobile phones marginally significant age group differences ($z = 5.5$; $p < 0.1$), showing that the ease of using mobile phones was slightly judged to

Table 1
User characteristics (gender, age, experience with technical devices, reported ease of using it as well as children's interest in technology in the three experimental conditions)

	Step-by-step	Menu structure	Exploration
Gender	4 Boys, 6 girls	6 Boys, 4 girls	5 Boys, 5 girls
Age	11.5 (1.7)	11.5 (1.8)	11.4 (1.7)
Frequency of using is 1 = several times a day; 2 = 1 × a day; 3 = 1–2 × a week; 4 = 1–2 × a month; 5 = less than once a month			
Mobile phone	3.4 (1.8)	4.7 (1.5)	4.2 (1.4)
PC	2.4 (1.4)	2.5 (1.1)	2.3 (1.2)
VCR	3.9 (0.9)	4.1 (1)	4 (1.5)
Ease of using is 1 = easy; 2 = quite easy; 3 = quite difficult; 4 = difficult			
Mobile phone	1.2 (0.8)	0.9 (1)	0.9 (0.7)
PC	1.7 (0.7)	1.5 (0.5)	1.4 (0.7)
VCR	1.4 (1)	1.6 (1)	1.5 (1)
Interest in technology (1 = low; 2 = quite low; 3 = quite high; 4 = high)	3.2 (0.9)	3 (0.7)	3.4 (0.7)

Given are means (and standard deviations).

Table 2
Age differences in the frequency of using technical devices, the reported ease of use as well as children's interest in technology

	9–10 year olds	11–12 year olds	13–14 year olds
Frequency of using is 1 = several times a day; 2 = 1 × a day; 3 = 1–2 × a week; 4 = 1–2 × a month; 5 = less than once a month			
Mobile phone	3.8 (1.5)	4.4 (1.3)	4.1 (2.1)
PC	2.4 (1)	2.5 (1.1)	2.3 (2.3)
VCR	3.9 (1.2)	3.9 (0.9)	4.2 (1)
Ease of using is 1 = easy; 2 = quite easy; 3 = quite difficult; 4 = difficult			
Mobile phone	1.3 (0.7)	1.2 (1)	0.6 (0.5)
PC	1.2 (0.4)	1.8 (0.8)	1.6 (0.5)
VCR	1.1 (0.9)	1.7 (1.2)	1.6 (0.9)
Interest in technology (1 = low; 2 = quite low; 3 = quite high; 4 = high)	3.4 (0.7)	3.1 (0.6)	3.1 (0.9)

Given are means (and standard deviations).

be easier by the older children compared to the other two age groups. All other statistic test did not reveal significant differences between age groups.

2.7. Procedure

At the beginning of the experiment, all children were told that the study aimed at evaluating how well current mobile phones are designed in order to be easily usable, especially by children. It was emphasized that it was not them, the children, who were being tested, but to find out what can be done to facilitate the usage of technical devices. In addition, the children were instructed to work as thoroughly and as fast as they could. In order to reinforce the children and to keep up their motivation, a "Congratulations" – message appeared on the screen when they had successfully solved a task.

Each child was randomly allocated to one of the three experimental conditions. In the step-by-step and the menu structure group, the specific type of instruction the child would thereafter use was presented and explained referring to a sample instruction. It was carefully ensured that the children knew exactly what the respective instruction aimed at and that they understood what they had to do in each task. The children in the step-by-step and the menu structure condition were shown the respective instruction for a period of 20 s after being explained what they had to do in the tasks and before starting to process each single task. The children belonging to the exploration group were encouraged to explore the functionalities of the simulated mobile phone for a time period of 5 min before being presented the tasks, but did not receive further guidance. As was observed by the experimenter, the children in fact explored the phone vividly and highly motivated.

Afterwards, children were asked to solve the same four tasks in a slightly modified manner for a second time, this time receiving neither the menu structure nor the step-by-step instruction. A time limit of 10 min was set for the processing of each task. This was done to get a detailed insight into users' navigation behavior and to prevent feelings of frustration, which could arise when children did not achieve to solve the majority of tasks. After processing the tasks on the mobile phone, the children were given a pause, where they spontaneously commented on the phone tasks, what they felt, what had attracted their attention in the menu or difficulties they had experienced. After the little break, children completed the two cognitive tests determining their spatial ability as well as their verbal short-term memory capacity. Overall, it can be stated that the children had much fun participating in the study, even though some of them stated after the experiment that the phone sometimes was "pretty hard to handle". The children were gratified for their participation with a little toy.

3. Results

In the following, the results are given. To assess the usefulness of the different instructions, learnability effects were of central interest, i.e. the users' increase or decrease in performance when solving tasks for the second time after having used different instructions. Thus, first, the effects of the different instructions on learnability are reported, and then effects of age as well as their interaction with the instructions are addressed. Furthermore, the relationship between the children's spatial ability and memory, gender and children's experience with technical devices and their efficiency using the mobile phones is looked at to determine possible reasons for the differences in users' performance.

3.1. Preliminary remarks with respect to statistical testing

The key findings, which were aimed at, were the interactions between children's age and the type of instruction used. For calculating interaction effects, analyses of variances are usually applied (as this procedure allows determining interacting effects). Due to the small sample size ($N = 30$), though, some cautionary notes have to be considered. Not only that averages might be rather unstable but also the statistical power of finding effects might be low. In order to meet this problem, the main effects were confirmed by non-parametric procedures (Wilcoxon tests for dependent and Kruskall–Wallis tests for independent measurements). Consistently, the results by non-parametric as well as parametric testing revealed the same effects to be statistically significant as well as the same effects to be not significant, respectively. Moreover, the homogeneity of variances was tested by Levene-tests. As found, the number of tasks solved, the time on task and the number of detour steps showed in-significant values (with F-values ranging from 1.9 to 1.2, and p-levels ranging from 0.14 to 0.31), thus variance homogeneity can be assumed and outcomes using analyses of variance statistics should not lead to misinterpretation. Solely, for the number of hierarchical returns variance homogeneity of variances was not present throughout (with Levene-values ranging from F-values between 0.02 and 8 and p-levels from n.s. to $p = 0.01$). Even though this measure, another fruitful parameter of children's menu navigation, was not dropped: As the other dependent variables showed the very same pattern of results, it was assumed that the outcomes could be interpreted reliably. Taken these preliminary tests as a basis, ANOVAs for repeated measurements were run for assessing learnability effects, with instruction type, age and level of spatial ability as independent factors and time on task, detour steps and number of hierarchical steps back as dependent factors. Revealing interrelations between user characteristics and performance, correlation analyses were executed. The level of significance was set at $p < 5\%$ for all tests carried out. Outcomes within the less restrictive significance level ($p < 10\%$) were referred to as marginally significant.

3.2. Effects of the instruction types

First, the main effect of instruction type was found not to be significant. However, this is nothing to be surprised of as we expected that the instruction types would differentially affect performance, depending on the children' age and their abilities to mentally incorporate the knowledge by the instruction types.

The participants' effectiveness solving the eight tasks with, on average, 7.1 (step-by-step; SE = 1.3), 7.5 (menu structure; SE = 0.8) and 7.4 (exploration; SE = 0.9) tasks solved out of eight tasks was comparably high in all groups. Therefore, in the following only the efficiency in terms of time on task, the detour steps as well as the hierarchical returns are considered as dependent variables. It was of interest to see which of the three navigational aids showed the most profound increase in performance from the first to the second trial.

Regarding the *time* needed to process the tasks, a significant learning effect was shown ($F(1,21) = 11.8$; $p < 0.05$) with an interaction between instruction type and number of trial ($F(2,21) = 5.4$; $p < 0.05$) as well as between age and trial ($F(2,21) = 12.5$; $p < 0.05$), but no interaction between all three factors. In Fig. 7 (left side), the outcomes are illustrated. Though not being the most correct representation, line diagrams depict the interaction, simply because the interactive nature of the three instruction types in the two trials can be demonstrated more clearly. There were two conditions showing an increase in performance from the first to the second trial: Using the menu structure, children were by 36% faster in the second compared to the first trial (first trial: $M = 579.4$ s; SE = 176.3 s; second trial: $M = 371$ s; SE = 127.4 s). The same was found in the exploration condition, yielding a still higher benefit: the amelioration from the first to the second trial was 44% (first trial: $M = 638$ s; SE = 176.3 s; second trial: $M = 356$ s; SE = 127.4 s). In contrast, in the step-by-step condition, the children showed a decrease in performance, with a 13% longer time on task, increasing from, on average 508 s (SE = 176.3 s) in the first to 574 s (SE = 127.4 s) in the second trial.

For the number of detour steps ($F(2,21) = 4.4$; $p < 0.05$) and hierarchical returns ($F(2,21) = 11.05$; $p < 0.05$) significant interactions between instruction type and number of trial (learnability) were found, too. As visualized in Fig. 7 (center: number of detour

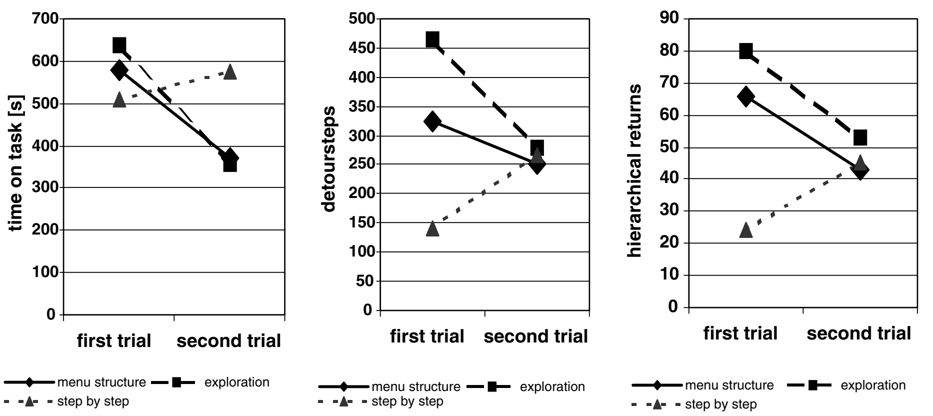

Fig. 7. Learnability effects of the different instructions regarding time on task (left), detour steps (center) and hierarchical returns (right).

steps; right side: number of hierarchical returns), the step-by-step condition was the only one showing no improvement in the second trial, but rather a worsening of performance: children made 267 detour steps (SE = 76.9) and 45 steps back (SE = 13) in the second trial compared to 141 detour steps (SE = 108.6) and 24 returns (SE = 19.8) in the menu hierarchy in the first trial (increase of 90%). In contrast, the menu structure led to 250 detour steps (SE = 76.8) and 43 steps back (SE = 13.1) in the second trial compared to 324 detour steps (SE = 108.6) and 66.3 steps back (SE = 19) in the first trial (decrease of 23% for detour steps, 35% for steps back). Also in the exploratory condition less steps were needed in the second trial, 278.7 detours steps (SE = 76.9) and 53 steps back (SE = 13.1), which corresponds to a decrease of 40% (34% for steps back) from the 463.7 detour steps (SE = 108.6) and 79.8 steps back (SE = 19.8) that were needed in the first trial. Interestingly though, the step-by-step condition yielded, with respect to the number of detour steps and hierarchical returns, the best performance in the first trial.

3.3. Effects of age

From a developmental point of view, it is insightful to see whether children of different age groups showed a different performance. This is illustrated in Fig. 8 (again, the illustration by line diagrams was chosen, even though no evident relation is given between measures in the first and second trial). As was expected by developmental status, the 9–10 year olds were in fact the group showing the lowest performance, and the children aged 13–14 years were found to be the best. When taking the time the children needed in the first and second run together (Fig. 8, left side), the youngest children were in total by 320% slower ($M = 1856.2$ s; SE = 122.8 s) and the intermediate age group about 100% slower ($M = 874$ s; SE = 116.2 s) than the oldest group ($M = 437$ s; SE = 111 s). This difference showed to be significant as determined by univariate analyses of variance ($F(2,27) = 11.2$; $p < 0.05$). For the detour steps, another significant age effect was found ($F(2,27) = 10.7$; $p < 0.05$). The youngest children carried out, on average, 1091 detour steps (SE = 80.5), the 11–12 year olds made 456 (SE = 76) and the oldest group 260 detour steps (SE = 72). It should be noted that the youngest group made 30 times more detour steps

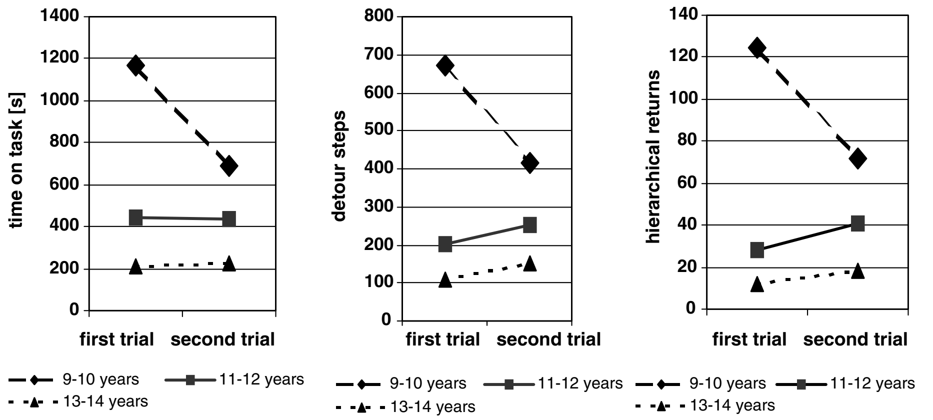

Fig. 8. Learnability effects in the different age groups regarding time on task (left), detour steps (center) and hierarchical returns (right).

than the 36 steps that were necessary to solve the tasks in the most direct way. Even though distinctly better, the middle-aged children still carried out 13 times as many and the oldest children still 7 times as many detour steps than the amount of steps actually necessary. Moreover, there was a significant effect of age on the number of hierarchical returns carried out ($F(2,27) = 12.3$; $p < 0.05$). The youngest group made 202 steps back (SE = 13.6), the 11–12 year olds 81 (SE = 12.9) and the children aged 13–14 years executed only 44 hierarchical steps back (SE = 11.9). Age effects are shown in Fig. 8.

Focusing now on learnability effects as mentioned before, they showed to interact with the children's age. The biggest increase in performance from the first to the second trial was observed in the youngest group: processing the tasks a first time, they made on average 673.7 detour steps (SE = 89.7) and 127.4 hierarchical steps (SE = 15.3) back needing 1169 s (SE = 122.7 s). In the second trial, only 417 detour steps (SE = 71.7) and 74.2 steps back (SE = 11.9) were carried out in about half the time ($M = 686$ s; SE = 122.8). In contrast, the older age groups (the 11–12 and the 13–14 year olds) did not improve in the second trial compared to the first.

3.4. Effects of spatial ability

In this section, the effects of children's spatial abilities on their performance using the phone are focused at, referring to the strong benefits of spatial abilities on navigation performance reported in the literature for adults. By median split two groups were formed: children having scores >41.5 points were allocated to a "high" spatial ability group, and those having scores <41.5 points were allocated to a "low" ability group. The outcomes for efficiency measure are illustrated in Fig. 9 (left side: time on task, center: detour steps; right side: number of hierarchical returns).

As to be seen in Fig. 9, spatial abilities have a major impact on children's performance (time on task): ($F(1,28) = 5.2$; $p < 0.05$); detour steps: ($F(1,28) = 8.3$; $p < 0.05$); hierarchical returns: ($F(1,28) = 10.2$; $p < 0.05$). Children belonging to the low-spatial ability group needed 796.5 s to process the tasks in the first trial, whereas children with high spatial abilities processed tasks in less than half the time ($M = 354$ s; SE = 129.3). However, both groups showed a significant learning effect ($F(1,28) = 4.9$; $p < 0.05$), which was not differ-

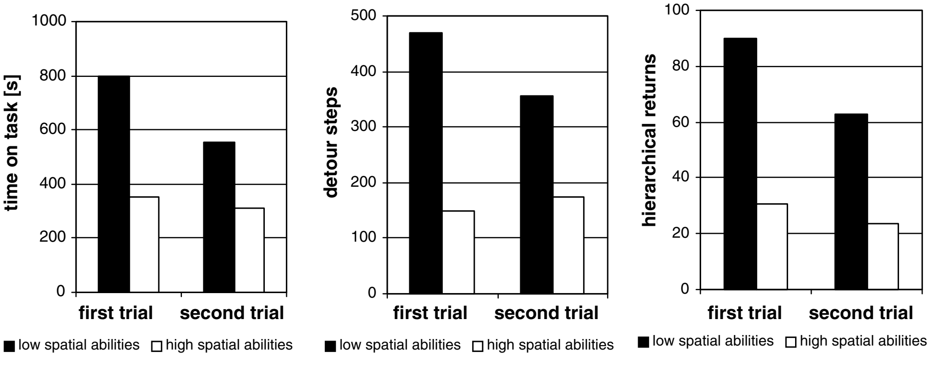

Fig. 9. Effects of spatial abilities on performance regarding time on task (left), detour steps (center) and hierarchical returns (right). Black bars indicate the group with lower spatial abilities (scores <41.5), white bars the group with high spatial abilities (scores >41.5).

ent for the two groups (no interaction effect). In the second trial, the children with high spatial abilities processed the four tasks in 311.3 s (SE = 100.4 s), while the low-spatial abilities children needed still 555.8 s (Fig. 9, left side). Looking now at detour steps and hierarchical returns (Fig. 9, center and right side), the very same pattern was revealed. Children belonging to the low-spatial ability group carried out 468.9 detour steps (SE = 83.7) and 90 hierarchical returns (SE = 14.6) in the first run, and still 354.9 detour steps (SE = 56.9) and 63 hierarchical returns (SE = 9.6) in the second run. Facing the 36 steps that were necessary to solve the four tasks on the most direct way, the detouring of these children is considerable. In contrast, children of the high spatial ability group performed distinctly better. In the first run, they "only" made 150.1 (SE = 83.8) detour steps and 30.9 hierarchical returns (SE = 9.6), and in the second run 175.5 detour steps (SE = 56.9) and 23.6 returns (SE = 14.6) in the menu. There was no significant effect of learnability for both measures, though, the interaction of learnability and spatial ability turned out to be statistically significant for the number of hierarchical returns ($F(1,28) = 4.4$; $p < 0.05$). Apparently, the benefit by learning was not equal for both groups, but was higher for the children having lower values of spatial ability.

3.5. Effects of short-term memory

Finally, the effects of short-term memory are reported. Taking the attained values in the memory test, a group with high (score >7 points) and a group with low verbal memory (score <7 points) abilities were formed through median split. In Fig. 10, the findings are illustrated (left side: time on task, center: detour steps; right side: hierarchical returns). For none of the measures, verbal memory affected efficiency interacting with the mobile phone in a significant way. Not even learnability measures were found throughout (only for the time, the children needed to process the tasks ($F(1,28) = 4.8$; $p < 0.05$)). Concluding, it must be assumed that verbal memory is not a major factor affecting children's menu navigation performance.

Summarizing the outcomes, three key findings can be outlined. The step-by step instruction, providing only landmark knowledge without information of the menu structure,

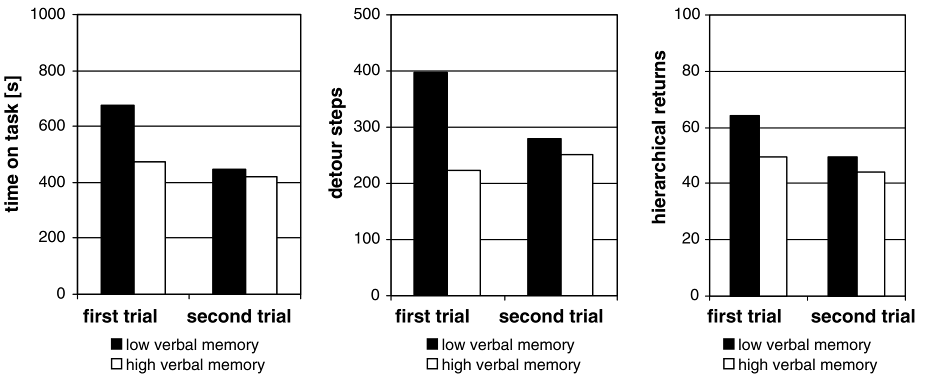

Fig. 10. Effects of short-term memory on performance regarding time on task (left), detour steps (center) and hierarchical returns (right). Black bars indicate the group with lower memory ability (scores <7), white bars the group with high memory ability (scores >7).

reveals to have only a limited sustainability. Furthermore, having high spatial abilities benefits menu navigation performance distinctly. Third, younger children show a distinctly lower competence using a mobile phone and profit less from instructions providing survey knowledge compared to older children. Even though the age effects on navigation performance conform to expectations due to age-related cognitive development, it is a basic question what actually is the origin of the developmental effect that is carried by age. Thus, for a deeper insight into the nature of the age effects, the interrelation between the children's cognitive abilities, experience with technical devices and age on the one hand and the navigational performance on the other hand is looked at.

3.6. Age-related differences in cognitive abilities and their interrelations with performance measures

To gain insights into user characteristics that are carried by the three age groups, in the following the outcomes of psychometric tests and their interrelation to the experience with technical devices and the reported ease of using them, are reported. To complete the picture, interrelations with gender are given.

A first analysis was concerned with the age-related effects of spatial abilities and verbal memory. In Fig. 11, the interrelations between spatial abilities (left side) and short-term memory (right side) as well as children's age are depicted. It can be seen from Fig. 11 that the interrelation of spatial ability and age was much more pronounced than the relation between verbal memory abilities and age. None of the younger children reached scores for spatial abilities above the median (>41.5 points), whereas with respect to memory abilities the younger ones were able to reach the "high score group" (above median, >7 points). Apparently, verbal memory abilities seem to be fully developed at an earlier point in time within the developmental process than spatial abilities.

Accordingly, a high correlation was found between age and spatial abilities ($r = .71$; $p < 0.01$): The older the children, the more developed their abilities to mentally represent spatial relations were. In contrast, the correlation between age and short-term memory was lower ($r = .28$; n.s.) and yielded no statistically significant interrelation. From this result it can be derived that the ability to memorize and recall does not vary systematically with children's age. Though, the level of verbal memory was significantly related to the level of spatial ability ($r = 0.39$; $p < 0.05$). Gender effects on spatial ability were not identified, showing boys and girls to be equally able to handle spatially demanding tasks.

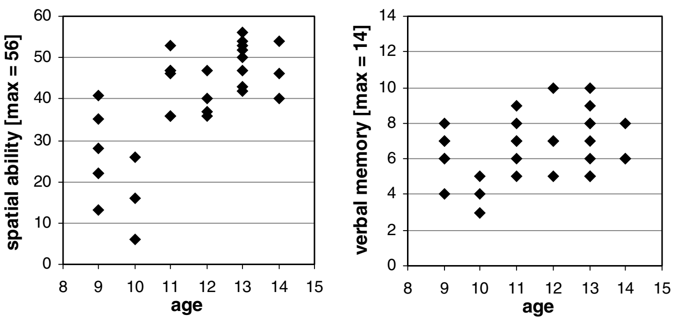

Fig. 11. The interrelation of age and spatial abilities (left side) and verbal memory abilities (right side).

The interest in technology was correlated with the frequency of using a PC ($r = 0.41$; $p < 0.05$). Children, who often use the PC reported to have a higher interest in technology, but neither the PC usage nor the interest in technology showed an interrelation with age. The latter is very insightful, as it suggests that difference in performance between the younger and older children cannot be attributed to their differing exposure to technical devices. The reported ease of using mobile phones was significantly related to the frequency using them ($r = -0.63$; $p < 0.00$), but it also showed a correlation to age ($r = -0.38$; $p < 0.05$) and the level of spatial ability ($r = -0.49$; $p < 0.01$). Thus, using a mobile phone was rated the easier, the more frequently the phone is used, the older the children were and also, the higher their spatial ability level was. Memory abilities did not show significant relations to the level of experience with technical devices and the reported ease of using them. Concluding, age effects were prominently found with respect to the level of spatial abilities and to the reported ease of using mobile phones. Beyond the boys' higher frequency of using a PC, gender effects were not present.

To analyze interrelations between the participants' age, cognitive abilities and performance using the mobile phone, Pearson correlation coefficients were calculated for these variables. Also, interrelations of gender and performance were assessed. Finally, the frequency and the reported ease of using a PC turned out to be an important facet of children's previous experience with technical devices. The correlation outcomes are depicted in Table 3.

Children's spatial abilities were correlated with time on task ($r = -.58$; $p < 0.05$), detour steps ($r = -.59$; $p < 0.05$) and hierarchical steps back ($r = -.56$; $p < 0.05$). Thus, high spatial abilities help the children to easily handle a mobile phone which means they need less time and detour steps as well as steps back to higher menu levels when solving the tasks. Short-term memory, in contrast, seems to be less important for users' efficiency: The correlations of verbal memory with efficiency measures range between $r = -.22$ (n.s., for hierarchical steps back) and $r = -.38$ ($p < 0.05$, for time on task). Even though gender is discussed as another major factor within the usage of technical devices, the boys and girls in our experiment showed a comparable performance with respect to their effectiveness and efficiency using a mobile phone. The frequency and the ease with which the children reported to use a PC were also interrelated with performance (frequency of using a PC: $r = .43$; $p < 0.05$ (number of tasks solved); $r = -.34$; $p < 0.1$ (time on task); Ease of using

Table 3
Correlations between age, gender, cognitive abilities, experience using a PC and the performance when interacting with the mobile phone

	Tasks solved	Time on task	Detour steps	Hierarchical returns
Age	$.50^{***}$	$-.62^{***}$	$-.59^{***}$	$-.64^{***}$
Gender	.29	.04	.04	.03
Spatial ability	$.54^{**}$	$-.58^{***}$	$-.59^{***}$	$-.56^{**}$
Short-term memory	$.31^{**}$	$-.38^{**}$	$-.33^{**}$	$-.22$
Frequency of PC usage	$.43^{**}$	$-.34^{*}$	$-.1$	$-.1$
Ease of using a PC	$-.26$	.25	$.48^{***}$	$.48^{***}$

Pearson values were used for interval-scaled data, Spearman-Rho values for ordinal-scaled data and Cramérs indices for nominal-scaled data.

* $p < 0.1$.

** $p < 0.05$.

*** $p < 0.01$.

a PC: $r = .48$; $p < 0.01$, for both, number of detour steps and hierarchical returns). Thus, while gender and memory abilities showed no significant impacts on performance, higher experience with the PC as well as higher spatial ability contributed to a good navigation performance.

From these findings it may be concluded that, on the one hand, age effects on the efficiency interacting with the mobile phone may be explained by increasing spatial abilities. On the other hand, the reason for the differing usefulness of the three instructions for the different age groups may be due to the instructions' demands on users' spatial abilities. Therefore, multivariate analyses of variances with instruction and spatial ability as independent factors were carried out. Dependent variables were the number of detour steps and the number of hierarchical returns in the first as well as in the second trial. Those variables had shown to be very sensitive for difficulties in navigation and orientation.

3.7. Interaction effects between instruction type and users' spatial ability

According to the omnibus F-tests, the effect of the instruction type alone did not reach statistical significance ($F(8,44) = 1.8$; $p > 0.05$), but spatial ability ($F(4,21) = 4.8$; $p < 0.01$) as well as the interaction of spatial ability and instruction type ($F(8,44) = 2.4$; $p < 0.05$) had significant effects on users' performance. This interaction is illustrated in Fig. 12 for

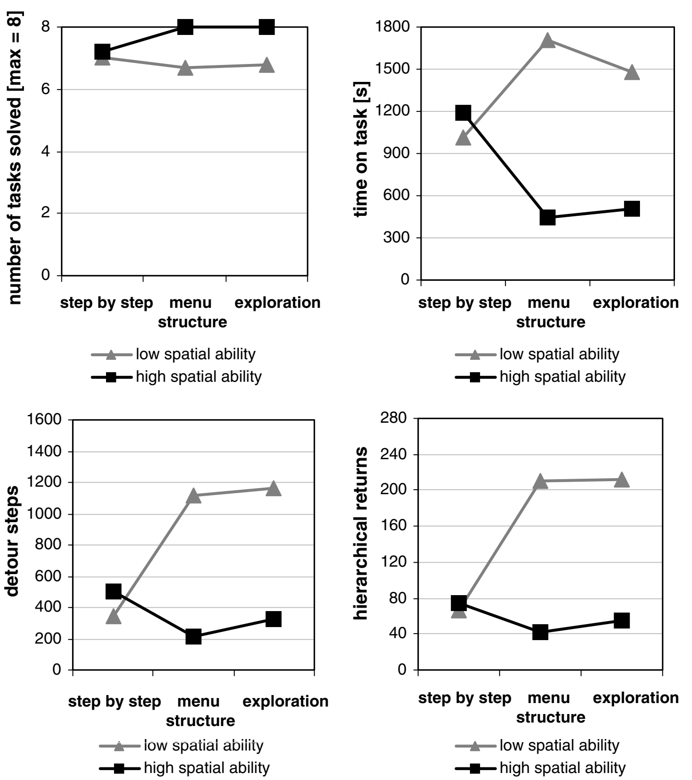

Fig. 12. The interaction between spatial ability and type of instruction on the number of tasks solved (upper left), the time on task (upper right), the number of detour steps (lower left) and hierarchical returns (lower right).

all dependent measures (number of tasks solved, time on task, number of detour steps and returns in menu hierarchy).

As can be seen, independently of the measure, children with high spatial ability (black lines) showed the lowest performance in the step-by-step condition and the best performance in the menu structure condition. Focusing on the number of detour steps and the returns in hierarchy, they made, on average, 498 detour steps (SE = 200.1) and 74 returns in menu hierarchy (SE = 35.6) in the step-by-step condition. In the menu structure, performance was significantly better, with only 213 detour steps (SE = 163.4) and 41.8 returns (SE = 29). Contrary, the children who had attained lower scores in the Mosaic Test (gray lines) showed the inverse pattern. The best performance was found in the step-by-step condition (detour steps: $M = 347.3$; SE = 163.4; hierarchical returns: $M = 65.8$; SE = 29.1). In the menu structure, however, the children with lower spatial abilities were extremely disadvantaged, carrying out about three times as many detour steps ($M = 1115.3$; SE = 200.1) and hierarchical returns ($M = 210.5$; SE = 35.6) compared to the step-by-step condition. Apparently, the understanding of the conceptual knowledge, which is provided by the menu structure, requests a concept of space and spatial knowledge, respectively. The same applies to the exploration condition: Children with lower spatial abilities were not able to benefit from the possibility to explore the menu prior to the experiment and to incorporate the knowledge of the menu into menu navigation. In the exploration condition, the children with lower spatial abilities executed, on average, 1162 detour steps (SE = 179) and made 211.6 returns in menu hierarchy (SE = 31.9), in contrast to the children with high spatial abilities, who took advantage of the exploration of the menu in advance (detour steps: $M = 322.2$; SE = 179.1; returns in menu hierarchy: $M = 322.2$; SE = 179.1).

4. Discussion

Three different types of instructions were under study regarding their effect on children using a mobile phone. On the basis of a supposed analogy between navigation in the mobile phone's hierarchical menu and navigation in the natural environment the instructions were created with reference to theories of spatial orientation, which state that different kinds of knowledge are relevant. The first instruction was a step-by-step instruction of the menu functions that have to be selected to solve a task, which – in terms of theories of spatial orientation – contains mainly landmark information. Secondly, a tree-map of the menu structure containing the names of the functions that need to be selected as well as their spatial relationship in the menu was given to participants. This instruction provides – apart from landmarks – survey knowledge of the menu. Thirdly, children were invited to explore the functionalities of the mobile phone on their own before processing the tasks. Through exploration of the menu survey knowledge is also supposed to develop, as well as the other types of spatial knowledge – route and landmark knowledge.

4.1. Which instruction type was most helpful for children's navigation performance?

In this section the question is addressed which of the instruction types performed best, thus having the most sustainable effect in the longer run. The results showed that the learning effect (the increase of performance from the first to the second trial) was strongest in the exploration and the menu structure condition, the two conditions that provide survey knowledge either directly (menu structure) or indirectly (exploration). While in the menu map the survey knowledge is directly instructed, the benefit of the exploration condition is that through active navigation of the surroundings survey knowledge is built up (Thorndyke & Hayes-Roth, 1982), even though also route and landmark knowledge are mediated by exploration. Hence, it can be concluded that survey knowledge is much more sustainable than landmark knowledge. Children, who have only received step-by-step information on the names of the functions to be selected, representing landmark knowledge, had an inferior performance in the second trial, when solving tasks without help, whereas children who have explored the menu or received information on the menu structure improved their performance from the first to the second trial. In recent studies dealing with the navigation performance in younger and older adults (e.g. Ziefle & Bay, 2006) it was found that the benefit of survey knowledge is fortified when convening with high spatial abilities. The specific benefit is explained by the persons' higher abilities to spatially integrate the structural information. To further substantiate this assumption also for the children examined in this experiment, it was analyzed whether there is an interaction effect between spatial abilities and the effects of the different instructions. This was indeed the case: In the second trial, when the children did not receive any help and had to use the knowledge acquired in the first trial, children with low-spatial abilities performed best when having used the step-by-step instruction before. In contrast, when receiving survey knowledge through the menu structure in the first trial or when having explored the menu, low-spatial ability children made more than twice as many detour steps in the second trial compared to the step-by-step condition. High spatial ability children on the other hand, show the best performance after having used the menu structure or explored the menu and perform worst when depending on the information provided by the step-by-step instruction: Being supported by landmark information only (step-by-step instruction) they made more than three times as many detour steps as those who had the opportunity to build up survey knowledge. The specific advantage of an aid containing survey knowledge for persons having high spatial abilities confirms findings with older adults interacting with mobile phones reported in recent studies (Ziefle & Bay, 2006).

Even though the findings support the importance of spatial orientation for navigation in hierarchical menus and confirm theoretical approaches in this context (Sein et al., 1993; Thorndyke & Goldin, 1983; Thorndyke & Hayes-Roth, 1982), a cautionary note should be considered with respect to the exploration condition. Though effects of the exploration were found to be strong, it is a basic question why the exploration did not show stronger advantage for the children compared to the map of the menu (menu structure). The exploration should – theoretically arguing – have had the strongest benefit, as all three types (landmark, route and survey knowledge) are mediated, while the menu structure aid contained mainly survey knowledge, but also landmark knowledge. This was not the case and both, the exploration and the menu map, performed equally well. Two arguments can be referred to here, one is of theoretical and one of methodological impact. Theoretically, survey knowledge could represent the strongest facet among the three spatial orientation components and, further hypothesizing, could extenuate with the benefits stemming from route or landmark knowledge. If so, it is understandable why exploration and the menu structure aid came to comparable positive effects. Even though this explanation seems to be reasonable, however, a methodological argument has to be ruled out first before this explanation can be adopted. On the basis of the present data, it cannot be finally excluded that the exploration of a broad menu, consisting of many functions, may have overtaxed the children's comprehensibility and their ability to deal with the broad functionality (especially as only 5 min were given to explore the menu). Thus, the exploration could also have led to some form of information overload, which resulted in the inability to extract the relevant information, especially for the younger children. Determining if this supposition is true, it must be analyzed which menu routes were actually visited by the children in the exploration and how far and deep the children struggled through the menu. Presumably, the strength of the advantage of the exploration condition might depend on *how* the children were wandering around while exploring. Hence it seems to be very continuative for further research to log and analyze also the navigation paths the children take while they explore the menu.

4.2. Which were the effects of age on navigation performance?

Not all children showed the same pattern of performance, neither with respect to the overall performance nor with respect to the different instruction types. The younger children being in the stadium of concrete operations indeed showed a distinctly lower performance than the older ones. The performance of the older age groups, both categorized as being in the stadium of formal operations, was found to lie closer to each other when compared to the younger group, but the relative developmental status within the same stadium was shown to be distinguishable, though. This can be taken from the fact that the 13–14 year olds children still outperformed the 11–12 year olds with respect to all efficiency measures. Though, as age just carries the respective cognitive abilities, the performance increment of the older children and the decrements in younger children should be scrutinized, enlightening the nature of the age-related changes. It was shown that the age groups did not differ in their reported frequency using technical devices as the PC; thus, different exposure to technology should be excluded as reason for age effects on performance. Instead, the impact of age-related cognitive abilities come into fore. Verbal memory ability was not significantly interrelated with age and, further, had a comparably lower impact on navigation. In contrast, spatial abilities were found to be crucial, not only as the older children had distinctly higher abilities than the younger ones, the spatial abilities were also revealed to be of crucial impact for a proper navigation with the mobile phones. Thus, the age effects found here cannot be separated from the differences in spatial abilities of the children. The fact that spatial abilities do have a major effect on the navigation performance in technical devices is by far not a new finding, but was found many times for younger and older adults (e.g. Bay & Ziefle, 2003a, 2003b, 2004; Downing et al., 2005; Kim & Hirtle, 1995; Norman, 1994; Ziefle & Bay, 2006; Ziefle et al., 2004). It is a new finding, however, and this can be regarded as a valuable addendum provided by the present study, spatial abilities do not only have impact on navigation performance for the developmentally descending process (older adults), but can be regarded as a major player also for children's navigation performance, thus, in the ascending process of cognitive development.

Conforming age-related differences with respect to spatial abilities, it was shown that the youngest children did not profit from the instructions providing survey knowledge. Instead, they performed best when given a step-by-step instruction, which contains only landmarks but no information about the arrangement of functions within the menu. Both findings may be explained by differences in the ability to mentally manipulate spatial relations, because as stated by Shemyakin (1962), the ability to mentally represent survey knowledge is only fully developed at the age of 12 years. A correlation between age and mental rotation ability undertaken in the present study showed that spatial ability increases with children's age. The children aged 9–14 under study were found to be subject to a developmental change, that is, between 9 and 14 years of age the ability to mentally rotate objects increases. Corresponding with earlier findings of (Bay & Ziefle, 2003a, 2003b), the increasing spatial abilities presumably enable the children to build up an adequate mental representation of the menu structure of the mobile phone. To understand and efficiently use survey knowledge, spatial abilities are necessary, too. This explains why young children could not profit from the menu structure and the exploration where they were supposed to build survey knowledge and form a mental representation of the menu.

4.3. The role of gender and computer experience for navigation performance

According to findings in the literature it was assumed that girls as well as children with less experience interacting with technical devices show inferior performance using the mobile phone. In our data, the gender effect could not be supported. Girls and boys showed to be equally capable of solving the mobile phone tasks in an efficient way. Computer experience, however, was strongly related to navigation performance in the phone. The more often children reported using a PC the better was their performance. Still, differences in computer experience seem not to correlate with age since the reported frequency using a PC did not differ between age groups. Interestingly, an interrelation between computer experience and spatial abilities were found. Highly spatially skilled children reported to use more frequently technical devices such as PCs. The nature of this interrelation – does high spatial ability promote their use of technical devices or is spatial ability developed as consequence of frequent PC usage? – seems worthwhile further exploration. It may provide insightful hints for pedagogic contexts and educating children for the technology prone future: A deeper understanding which abilities underlie the successful interaction with technical devices is of cardinal interest, as the utilization of technological devices is no longer voluntary or restricted to a specific user group, but has already become an essential requirement also for children.

4.4. Conclusions

From the findings it may be concluded that the building up of mental spatial representations in form of survey knowledge of the menu structure is a major issue when using a mobile phone for all users, however, especially for child users, as they are a major future user group, this should be considered in phone design. The process of building up an appropriate representation and gaining survey knowledge has to be supported by instructions and manuals. Theories of spatial orientation in the natural environment can give useful insights and hints towards how such instructions should be designed. For very young children, that is children aged between 9 and 10 years it is of interest whether a simplified form of the diagram of the menu structure could be more helpful. If they have suffered from information overload through the menu structure provided in the present experiment, a diagram containing less detail could prove to be much more efficient. Recapitulating, we could show in the described experiment that for children aged 11 or older a diagrammatic instruction providing survey knowledge by presenting the menu's structure as well as the location of the functions that have to be selected to solve a task could be as efficient as a free exploration of the phone for a period of 5 min. Taking into consideration that children were only allowed to look at the instruction for 20 s – in total for a period of 80 s in the four tasks compared to the 5 min of exploration – this is a very good result.

A final note is concerned with the overall performance of the children observed here. One might have expected that the mobile phone tasks would have been processed without too much friction losses. Even if this might be true for the effectiveness (as most of the children solved about or slightly more than 7 out of 8 tasks), it should be taken into account that children were given 10 min for each task, plenty of time compared to the time periods users take in real life. Looking at efficiency measures however, it was found that even the 13–14 year old kids carried out a lot of detouring routes instead of solving the tasks the shortest way possible. For the four phone tasks to be processed twice, the children carried out 130 detouring steps and returned 17 times to higher levels in menu hierarchy after all. Given the reference of 36 steps that were necessary to solve the four tasks on the most direct path, this seems to be quite substantial. Conceivably, this shows that mobile phones – at least in the way being currently designed – impose a high cognitive load even on 13–14 year old children.

4.5. Limitations of the present study and future research

Finally, some limitations of the present study are considered that should be pursued for future research. One limitation refers to the small sample size. Due to the rather complex experimental design aiming at uncovering interaction effects between the utility of instructions and children's age as well as their cognitive developmental status, future studies should replicate the findings with more children of different ages. This would strengthen the overall statistical power. Even though, the consistency of the findings within this study and also across studies dealing with the effects of instruction types in other user groups (Bay, 2003; Bay & Ziefle, 2004; Ziefle & Bay, 2006), corroborates the stability of the results found here. In future research though, still more experimental work should be conducted to corroborate the results presented here. A second point refers to the type of device examined. The mobile phone is only one example of a whole group of devices with small screens and only few keys to interact with the complex menu. To further substantiate the importance of survey knowledge for interacting with this type of technology, experiments putting light on the effects of different types of instructions on children's performance interacting with other devices should be carried out. Future studies should continue in this line of research and assess for example which level of detail is optimal for children's efficiency interacting with a mobile phone. Maybe it is enough to show them a simplified diagram of a menu tree. That a diagram of the menu tree without the names of the functions to be selected can be still more efficient than a menu tree with the function labels as shown in a former study with adults as participants (Bay, 2003). Which specific information is most helpful for children, however, still remains to be shown. Furthermore, as many users are reluctant to spend time reading manuals, another promising procedure is to include orientation aids into the displays of mobile phones. Due to their small size this presents a considerable challenge. Though, first attempts were already carried out (Ziefle & Bay, 2006) and proved to be very efficient especially for older adults, who are – as children – a user group very sensitive for suboptimal user interface design. If and how these navigation aids can help children interacting with the devices without imposing an information overload on them is a timely topic that needs further research.

Acknowledgements

Special thanks to Ralf Barkmeyer for recruiting participants and carrying out the experiments as well as to the children's teacher, Doerthe Rothfeld, who supported us to conduct our research by motivating and instructing children and their parents. Furthermore, the authors thank Sarah Hatfield for her corrections of the paper and Preethy Pappachan as well as Judith Strenk for preparing the figures. A final thank goes to Paul. A. Kirschner and two anonymous reviewers for their critical and constructive comments on an earlier version of this manuscript.

References

Allen, G. L. (2004). *Human spatial memory. Remembering where*. Mahwah, New Jersey: LEA.

Bay, S. (2003). Cellular phone manuals: Users' benefit from spatial maps. In G. Cockton & P. Korhonen (Eds.), *Extended abstracts of the 2003 Conference on Human Factors in Computing Systems. CHI 2003* (pp. 662–663). ACM.

Bay, S., & Ziefle, M. (2003a). Design for all: User characteristics to be considered for the design of phones with hierarchical menu structures. In H. Luczak & K. J. Zink (Eds.), *Human factors in organizational design and management – VII* (pp. 503–508). Santa Monica: IEA.

Bay, S., & Ziefle, M. (2003b). Performance in mobile phones: Does it depend on proper cognitive mapping? In D. Harris, V. Duffy, M. Smith, & C. Stephanidis (Eds.), *Human centred computing. Cognitive, social and ergonomic aspects* (pp. 170–174). Mahwah, New Jersey: LEA.

Bay, S., & Ziefle, M. (2004). How instructions influence novice users' interaction with mobile phones. In H. M. Khalid, M. G. Helander, & A. W. Yeo (Eds.), *Work with computing systems* (pp. 388–393). Kuala Lumpur: Damai Sciences.

Bay, S., & Ziefle, M. (2005). Children using cellular phones. The effects of shortcomings in user interface design. *Human Factors, 47*(1), 158–168.

Beckwith, L., Kiyssinger, C., Burnett, M., Wiedenbeck, S., Lawrance, J., & Blackwell, A. et al. (2006). Tinkering and gender in end-user programmers debbuging. In *Proceedings of the ACM conference on human factors in computing systems (CHI 2006)*, pp. 231–240.

Billingsley, P. A. (1982). Navigation through hierarchical menu structures: Does it help to have a map? In *Proceedings of the human factors and ergonomics society 26th annual meeting* (pp. 103–107). Santa Monica, CA: Human Factors Society.

Carroll, J. M. (1990). *The Nurnberg funnel: Designing minimalist instruction for practical computer skill*. Boston: MIT.

Carroll, J. M., & Mack, R. L. (1985). Exploring a word processor. *Human Computer Interaction, 1*, 283–307.

Downing, R. E., Moore, J. L., & Brown, S. W. (2005). The effects and interaction of spatial visualization and domain expertise on information seeking. *Computers in Human Behaviour, 21*, 195–209.

Egan, D. (1988). Individual differences in human–computer-interaction. In M. Helander (Ed.), *Handbook of human–computer-interaction* (pp. 543–568). Amsterdam: Elsevier.

EN ISO 9241-11 (1997). *Ergonomic requirements for office work with visual display terminals. Part 11: Guidance on usability*. Berlin, Germany: Beuth.

Frese, M., Albrecht, K., Altmann, A., Lang, J., Van Papstein, P., Peyerl, R., et al. (1988). The effects of an active development of the mental model in the training process: Experimental results in a word processing system. *Behaviour and Information Technology, 7*(3), 295–304.

Frese, M., Brodbeck, E., Heinbokel, T., & Mooser, C. (1991). Errors in training computer skills: On the positive function of errors. *Human Computer Interaction, 6*(1), 77–93.

Informationszentrum Mobilfunk (2004). *Studien und Umfragen zum Thema Mobiles Leben* [studies and surveys dealing with mobile living]. Retrieved June, 2004, http://www.izmf.de/html/de/6318.html.

Kim, H., & Hirtle, S. C. (1995). Spatial metaphors and disorientation in hypertext browsing. *Behaviour and Information Technology, 14*(4), 239–250.

Mayer, R. E. (1986). Models for understanding. *Review of Educational Research, 59*(1), 43–64.

Mobileyouth.org (2005). *5 Million school children own a mobile phone in UK*. Retrieved June, 2005, http://www.mobileyouth.org/my_item.php?mid=1471.

Norman, K. L. (1994). Spatial visualization – A gateway to computer-based technology. *Journal of Special Education Technology, 12*(3), 195–206.

Piaget, J. (1983). Piaget's theory. In P. Mussen (Ed.), *Handbook of child psychology*. New York: Wiley & Sons.

Rodriguez, M. A., & Polson, P. G. (2004). Creating usable wordless instructions for performing complex one-time tasks: Effects of violating the rules. In *Proceedings of the human factors and ergonomics society 48th annual meeting* (pp. 941–945). Human Factors and Ergonomic Society.

Rose, C. (1985). *Accelerated learning*. Aylesbury: Accelerated Learning Systems.

Sein, M. K., Olfman, L., Bostrom, R. P., & Davies, S. A. (1993). Visualisation ability as a prediction of user learning success. *International Journal of Man–Machine-Studies, 39*, 599–620.

Shayer, M., & Wylam, H. (1978). The distribution of Piagetian stages of thinking in British middle- and secondary school children. *British Journal of Educational Psychology, 46*, 164–173.

Shemyakin, F. N. (1962). Orientation in space. In B. G. Anan'yev et al. (Eds.), *Psychological sciences in the USSR, Vol. 1, Part 1 (Rep. No. 11466)* (pp. 379–384). Washington: US Office of Technical Reports.

Siegel, A. W., & White, S. H. (1975). The development of spatial representations of large-scale environments. In H. Reese (Ed.), *Advances in child development and behaviour* (pp. 10–55). New York: Academic.

Taylor, H. A., & Tversky, B. (1992). Spatial mental models derived from survey and route descriptions. *Journal of Memory and Language, 31*, 261–282.

Taylor, H. A., & Tversky, B. (1996). Perspective in spatial descriptions. *Journal of Memory and Language, 35*, 371–391.

Tewes, U. (1983). *Hamburg–Wechsler Intelligenztest für Kinder [Hamburg–Wechsler intelligence test]. Revised. (HAWIK-R)*. Bern, Switzerland: Huber.

Thorndyke, P. W., & Goldin, S. (1983). Spatial learning and reasoning skill. In H. L. Pick & L. P. Acredolo (Eds.), *Spatial orientation. Theory, research, and application* (pp. 195–217). New York: Plenum.

Thorndyke, P. W., & Hayes-Roth, B. H. (1982). Differences in spatial knowledge acquired from maps and navigation. *Cognitive Psychology, 14*, 560–589.

Van den Heuvel-Panheizen, M. (1999). Girls' and boys' problems: Gender differences in solving problems in primary school mathematics in the Netherlands. In T. Nunes & P. Bryant (Eds.), *Learning and teaching mathematics: An international perspective* (pp. 223–253). UK: Psychology Press.

Vicente, K. J., Hayes, B. C., & Williges, R. C. (1987). Assaying and isolating individual differences in searching a hierarchical files system. *Human Factors, 29*(3), 349–359.

Westermann, S. J. (1997). Individual differences in the use of command line and menu computer interfaces. *International Journal of Human Computer Interaction, 9*(2), 183–198.

Ziefle, M., & Bay, S. (2004). Mental models of a cellular phone menu. Comparing older and younger novice users. In S. Brewster & M. Dunlop (Eds.), *Mobile human computer interaction* (pp. 25–37). Berlin: Springer.

Ziefle, M., & Bay, S. (2005). How older adults meet complexity: Aging effects on the usability of mobile phones. *Behaviour and Information Technology, 24*(5), 375–389.

Ziefle, M., & Bay, S. (2006). How to overcome disorientation in mobile phones' menus: A comparison of two different types of navigation aids. *Human Computer Interaction, 21*(4), 393–432.

Ziefle, M., Bay, S., & Schwade, A. (2006). On keys' meaning and modes: The impact of different navigation key solutions on children's efficiency using a mobile phone. *Behaviour and Information Technology, 25*(5), 413–431.

Ziefle, M., Bodendieck, A., & Kuenzer, A. (2004). The impact of user characteristics on the utility of adaptive help systems. In H. M. Khalid, M. G. Helander, & A. W. Yeo (Eds.), *Work with computing systems* (pp. 71–76). Kuala Lumpur: Damai Sciences.

Ziefle, M., Arning, K., & Bay, S. (2006). Cross-platform consistency and cognitive compatibility: The importance of users' mental model for the interaction with mobile devices. In K. Richter, J. Nichols, K. Gajos, & A. Seffah (Eds.) *The Many faces of consistency in cross-platform design, CEUR-WS.org., Vol. 198* (pp. 75–81). http://CEUR-WS.org/Vol-198/.