Information Design Journal 18(1), **22–35**
© 2010 John Benjamins Publishing Company
DOI: 10.1075/idj.18.1.03cho

Charalambos Koutsourelakis & Konstantinos Chorianopoulos

Icons in mobile phones

Comprehensibility differences between older and younger users

Keywords: icon, mobile phone, age, recognition, user study, usability

Mobile phones have become a ubiquitous information and telecommunication terminal used by diverse age groups. Despite research findings that indicate the special needs and skills of senior citizens, contemporary handsets offer limited consideration for them. We investigated whether typical mobile phone icons are suitable for older users, as measured through unaided icon comprehension. For this purpose, we studied 38 users and evaluated 25 icons, which we captured from five different handsets. It was found that there are significant differences in icon comprehension performance between older and younger mobile phone users. We conclude that designers of mobile phone icons have to regard older user groups differently and we describe the qualities of icons suitable to each one of the age groups. Finally, we highlight the need for additional similar research in different ethnic and cultural groups.

Introduction

Early mobile phones functioned just as speech terminals, but newer models have introduced extra functionality, which is depicted through icons. In contrast to desktop computers, mobile phones are a consumer product with a very wide user base. Moreover, contemporary mobile phones are converging with powerful multimedia mobile computers (e.g. Windows Mobile, Apple iPhone, Google Android, Palm Pre), which add a plethora of new features (Keijzers, den Ouden, & Lu, 2008). As a matter of fact, user interface icons have been widely deployed in mobile phone operating systems in order to ease navigation. Given the diversity of mobile phone icon providers (handset manufacturers, platform providers, wireless operators), and the short product life-cycles, are contemporary mobile phone icons suitable for all age groups in terms of subjective comprehensibility?

Although desktop icons have received much attention from researchers and practitioners, mobile phone icons have not been studied enough. As described in the next section, there is a significant body of research on desktop icons, because icons are a fundamental element of graphical user interfaces. Besides desktop icons, Kim and Lee (2005) evaluated mobile phone icons and provided guidelines for the level of abstraction in the visual design of icons. However, their study was not based on available icons but on experimental designs, and they did not consider age differences.

Common sense has linked icons with sophisticated user interfaces and increased perceived usability. Then, the main research question in our work is: Do older users benefit by the design of mobile phone icons?

The objective of this study is to make a comparative

© 2010. John Benjamins Publishing Company
All rights reserved

comprehensibility evaluation for mobile phone icons which are employed in competitive handsets. Moreover, we have included a set of icons by a wireless operator (Vodafone), which has decided to replace the original handset icons with branded ones. We evaluated the icons with young and old users, in terms of subjective comprehensibility.

The rest of this article is organized as follows: (1) Present related work on icons, (2) Explain the methodology of the study, (3) Present the results of the study, (4) Discuss the results in the face of established theories and related findings, and (5) Provide conclusions for practitioners and suggestions for further research.

Related work

The design of graphical user interfaces (GUIs) implies a need for effective icons, which represent actions and objects. Indeed, since the beginning of GUIs, icon design and usability evaluation have been of importance (Rogers, 1989).

Besides the desktop computer, icons have been an important part of other interaction paradigms, such as the Web and mobile phones. In addition, icons have been included in contemporary touch-based interfaces (e.g. Apple iPhone, Google Android). Therefore, it seems that icons will continue to be one of the dominant components in the majority of future user interfaces.

Taxonomies

Researchers have developed several icon taxonomy systems that organize icons based on their graphic elements (Wang, Hung, & Liao, 2007). According to Rogers (1989) there are two types of icons: (1) Data icons represent objects that could be used in actions (e.g. folders, files), and (2) Function icons represent objects that could perform actions (e.g. paintbrush). Moreover, icons can be described according to the representational technique in the following categories: (1) metaphoric (e.g. road sign for falling rocks), (2) paradigm of use (e.g. fork and knife for restaurant), (3) symbolic (e.g. broken glass stands for fragile objects), and (4) abstract representation that should be memorized by the user (e.g. the symbol for electricity or radioactive places). Mobile phone icon designers have been building upon these paradigms (e.g. the "phonebook" icon is metaphoric and the "settings" icon denotes a paradigm of use), but there might be several functions and objects that defy straightforward visualization (e.g. internet, applications).

Besides Rogers (1989), there are alternative approaches to the taxonomy of icons provided by researchers outside of the computing field. According to Peirce (cited in Marcus, 1993), the semiotic dimensions of a sign are: lexical (production), syntactic (arrangement), semantic (references to that for which they stand), and pragmatic (consumption). Signs by definition are icons, indexes, or symbols. Icons are "naturally" meaningful, like a thin pencil line to represent a line. Indexes are signs caused by something and therefore referring to them, like muddy boot prints on the kitchen floor being a trace or index of the children walking through. Symbols are abstract and must be learned. In many cases, icons in graphical user interfaces are not icons at all, but symbols. However, the computer industry has been slow in adopting the terminology of semiotics for computer graphics. At this point it might be late to make appropriate distinctions, but we are highlighting what the sign design community considers technically correct.

Design guidelines

Unfortunately, there is no such thing as a direct mapping between images and words, be that objects or actions (Rogers, 1989). Hopefully, there are several (culture-dependent) assumptions and conventions (e.g. the mean-

© 2010. John Benjamins Publishing Company
All rights reserved

ing of red, or that of an exclamation mark) that help the design process of icons. Icon designs usually depict familiar images, symbols, and simple diagrams, which stand as metaphors of real world objects and activities. Moreover, an icon is usually complemented by a text label, which provides a hint to its meaning. Researchers have established that neither text alone nor image alone is enough (Egido and Patterson, 1988; Haramundanis, 1996; Wiedenbeck, 1999). Instead the combination of text and image makes up a comprehensible icon.

Nevertheless, icons have been criticized for some of their properties. According to Tufte (1990) the information density of an icon is much inferior to that of the respective label alone. In particular, if an icon uses much "ink" and little semantics in relation to a word, then this might be considered a "fat" icon that attempts to portray visual objects. In contrast, one of the main powerful qualities of a well-designed sign is the efficiency in conveying what would take sometimes many words and many more pixels to communicate verbally. Byrne (1993) has demonstrated that very simple icons might be as successful as complex ones. For example, there are actual systems, such as LoCoS and Blisssymbolics, that have been considered elegant and simple. According to Barr, Noble, and Biddle (2003), an icon is successful when the meaning perceived by the users matches the object or action that the designer intended to communicate though the graphical portrayal of that particular image. Further empirical evidence for the design of motor vehicle displays has been provided by Green and Burgess (1980), who describe an elaborate study of pictographic displays and draw conclusions that might be applicable beyond the car industry.

Aging population and mobile phone use

Improving the quality of life of elderly people is an important issue in the vision of an information society for both research and practice. Due to the ubiquity of the mobile phone, it is often assumed that they are inherently easy to use and require no training, regardless of the age group of the user. In particular, a literature search revealed that few public studies have been done on the performance of current mobile icons for older people. Handset manufacturers might be doing usability tests, but they are not publishing results due to the increased competition in this area. Previous research has identified the age differences in mobile phone navigation performance (Arning and Ziefle, 2007), but there is no research on the universal accessibility of mobile phone icons. In addition to the business drivers and legislative requirements for having accessible mobile phones, the mobile phones have also become an internet terminal, as well as platforms for all sorts of computing tasks (Keijzers, den Ouden, & Lu, 2008). Thus, there is a need to ensure that mobile phone icons cater for the skills of both young and old.

Could the ubiquity of the mobile phone be a threat for the universal accessibility of mobile phone icons? In this study, we compare icon performance between age groups for five icon sets.

Methodology

The objective of the study was to evaluate the comprehensibility of mobile phone icons. For this purpose, we selected five types of icons from five mobile phones (25 icons in total). The types of icons were selected to match common functions/objects, such as address book (phonebook), applications, internet, phone settings, calls log. Then, we presented each one of the 25 icons to 38 users (half of them young and the other half mature) of mobile phones and asked them to guess (free-form question) what the meaning of each icon is. Finally, we coded the results and compared the averages between the two age groups. In the following subsections, we present the details of our methodology.

© 2010. John Benjamins Publishing Company
All rights reserved

Approach

Previous approaches to the evaluation of icons have emphasized that it is critical to collect information about the comprehensibility of icons, suitability of icons and user preferences for icons (Eisen, 1990). In this study, we focused on possible differences in comprehension of widely available mobile phone icons between two age groups. For this purpose, we used an open-ended question to assess the perceived meaning of each icon.

In terms of the experimental setting, there are contrasting opinions with regard to the place where the study is performed. For example, Kaikkonen et al. (2005) claim that there is no significant difference in the usability results they found between the lab and the field. On the other hand, Nielsen et al. (2006) claim that there might be different results found for some usability aspects of a mobile phone. This study took place in a lab context using a computer screen, as explained in the following subsections.

Subjects

Nineteen young users were recruited in the study (min. age = 15, max. age = 22, avg. age = 17, std dev. = 1.3). Ten young users were males and nine of them females. Moreover, nineteen mature users agreed to participate in the study (min. age = 30, max. age = 73, avg. age = 53, std dev. = 12.7). Ten mature users were males and nine were females. All young users had used at least one mobile phone in the past. The current handset was the first mobile phone for more than half of the mature users. In particular, mature users are not regarded so only because of their chronological age, but also because they have been late in the adoption of mobile phones. In this study, we have distinguished between older and younger users not in the traditional accessibility sense of age-related disabilities, but in the sense of two different generations of users: one that has grown up with mobile phones and one that did not. In the following subsection, we describe the selection of mobile phones and icons.

Most of the subjects have been born and raised on a major and popular Mediterranean island (Crete, Greece). Thus, they tend to have deep connections to traditions (religion, lifestyle), but at the same time they are open to the hordes of tourists and their diverse cultures that go to the island almost year around. Finally, all subjects were recruited according to convenience sampling from a secondary education school. Thus, young users were students and mature users were teaching staff with various educational backgrounds (science, humanities, arts). Although, in statistical terms, convenience sampling is not considered to be representative of the respective populations, it is very common in usability tests.

Materials

We selected handsets from four different manufacturers. In addition, we selected one more handset that was branded by an international wireless service operator (Vodafone). The selection of the handsets was based on the European market shares of the respective manufacturers (source: IDC 2006). The particular handset models were selected according to convenience.

Table 1. Mobile phone handsets selected according to brand popularity (IDC 2006).

Brand	Handset
Nokia	6230i
Motorola	RAZR v3
Samsung	d500
Sony Ericsson	750i
Vodafone	Motorola V360v

Next, we had to select a consistent set of icons from each mobile phone menu. Mobile phone menus form a hierarchy, which is not consistent among different

© 2010. John Benjamins Publishing Company
All rights reserved

manufacturers. In some cases, UI objects/actions are found at the top level of the menu with an icon (e.g. SMS messaging), while in other cases features are organized in logical groups (e.g. messaging includes MMS, SMS, email) represented by an icon at the top-level menu. Some contemporary handsets also provide animated icons. Moreover, several handsets provide menu customizations, such as: (1) text-only menus, (2) full-screen-icon menus, (3) icon-palette menu, and (4) menu shortcuts. In this study, we only considered top-level menu icons, icon-palette menus and static icons.

We selected five icons from the main menu of each phone (Figure 1). The icons were selected in such a way that: (1) they stand for popular UI objects/actions (e.g. phonebook, applications, organizer, internet, phone settings, calls log), and (2) there are similar icons on other mobile phones. We did not include the short message service (SMS) and the photo camera icon in the study, because a pre-study questionnaire revealed high familiarity with these features. For those phones with animated icons we selected the static version of the icon; that is, the state of the icon when the cursor is not over it. In total, 25 icons were selected. Each icon was coded (A1, A2, etc.) in order to facilitate statistical tests.

Following Rogers (1989), we organized the selected icons in the following categories:

- Metaphoric. For example, icons A1, B1, C1, D1, E1, and D3 employ familiar real-world objects. Also, icons A5, C5, B3, and D5 employ a book to refer to the address book.
- Paradigm of use. Icons A4, C4, D4, and E4 employ a tool to refer to the notion of fixing or changing the settings of the phone.
- Symbolic. Icons A2, B2, C2, and E2 employ an earth to refer to the World Wide Web.
- Abstract. Icons C3, E3, D2, and B5 employ unfamiliar objects or desktop computing symbols.

The icons in Figure 1 have been arranged according to brand (columns) and according to meaning (rows). The attempt to organize the icons according to meaning (e.g. functionality that appears in submenu items) has not been very straightforward. There are some icons that do not have a corresponding one on another mobile phone, so there are some rows which contain icons that do not seem to fit visually, but they do have a relationship with the rest of the row in terms of the meaning. For example, icon E1 (Applications) provides access to a submenu that includes an agenda application, which in the rest of the phones appears as a top-level item-icon. At the same time, there is icon C3 (Applications), which has the same text label as the latter, but provides access to user-installed applications and games. Although two mobile phones have used the same text label for unrelated groups of functionality, we have organized the icons in a matrix according to semantic meaning conveyed through an icon.

There are additional icon text-label inconsistencies between brands: agenda versus organizer, and phonebook versus contacts. In brief, the reason for the lack of consistent appearance of icons in some of the rows is that despite the phones having similar functionality, they have different menu levels to reveal that functionality, and/or different hierarchies, and/or different naming schemes. This inherent difficulty in organizing and providing a nice clean matrix of the study's icons might be one of the explanations for the poor user performance that we found.

Measuring instruments

There are two types of measuring instruments: qualitative and quantitative. Barr, Noble, and Biddle (2003) have proposed an analytical approach which is based on semiotics for the evaluation of icons. Although semiotics might be beneficial for organizing icons into categories, it does

© 2010. John Benjamins Publishing Company
All rights reserved

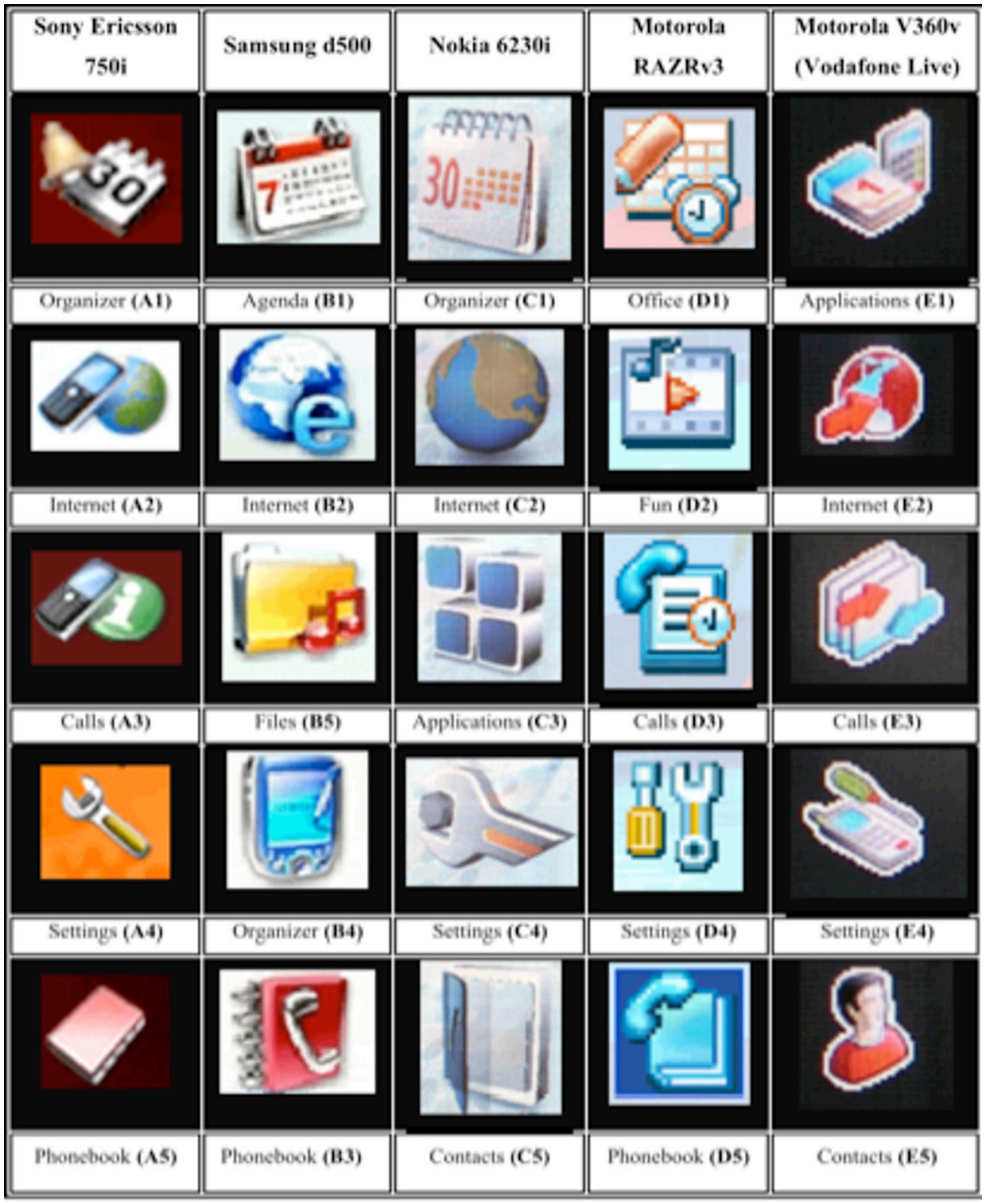

Figure 1. The icons used in the study organized in columns for each of the brands: Sony Ericsson, Samsung, Nokia, Motorola, Vodafone (left to right). Please note that some icons do not have direct association between different brands.

© 2010. John Benjamins Publishing Company
All rights reserved

not afford a quantitative comparison between alternative sets of icons. In order to perform an empirical evaluation of icons among competing mobile phone makers, we employed a quantitative measuring instrument.

Previous icon studies have considered several quantitative measuring instruments. Some researchers have found that icons afford a quicker and easier understandability than the respective text (Collins and Lerner, 1982). In contrast, other researchers claim that there is no difference between the performance of the two, and that the best approach is to combine icons with text labels (Egido and Patterson, 1988). In this study, we evaluated icons without their text labels, in order to measure unaided comprehension rates in different age groups. Although this approach does not provide valid results for the absolute comprehensibility of each icon, it offers a simple and effective instrument, which we used only to compare between user groups.

We used an open-ended question to measure the comprehensibility of each icon (Lohse et al., 1994; Sorenson & Webb, 1991). The answers to open-ended questions were coded in four levels of comprehension: (1) complete comprehension; (2) partial comprehension, when the answer was very close to the real meaning, but did not include the correct text label; (3) wrong comprehension; and (4) no comprehension, which we considered to be representative of the lowest comprehensibility of an icon, the one that does not bring any correlation to the mind of the user. Users could write down up to three guesses about the meaning of each mobile phone icon. Finally, we employed a demographic questionnaire, in order to record basic information about the users (e.g. age, gender).

Procedure

The icons of the study were presented one-by-one on a computer screen. We retained the same resolution between the original icon and the one presented on the computer screen. There were minor differences in the absolute size of the icons, due to the differences in pixels per inch among the mobile phone screens and between mobile screens and computer screens. The icons were displayed in a random order and in full color. Each icon was displayed without any text caption for approximately one minute (total study time was 25 minutes). The subjects were asked to fill in a paper form, which had one question about the meaning of each icon (the respective icon was displayed in grayscale on the paper, in order to assist matching, due the random order of icon display on the computer screen). All questions fitted within one A4 page, in order to avoid page-turning overhead. The comprehension was measured with an open-ended questionnaire, which asked users to guess the meaning of each icon. The answers were coded in four categories as described in the section on measuring instruments.

Results

We tested 25 icons with 38 users. Half of the users were less than 22 years old (with an average age of 17) and the other half of them were more than 30 years old (with an average age of 53).

Firstly, the results of each group indicated a wide variability in the comprehensibility of icons. Figure 2 portrays a comparison of the average comprehension rate for each of the 25 icons between young and mature users. One stands for full comprehension and four stands for no comprehension. Thus, the lowest score denotes the best performance.

Overall, the bar graph for the mature users appears to have ten icons with very low comprehension rates (more than 2.5). In contrast, there are only four icons with very low comprehension rates (more than 2.5) for young users. Moreover, both the minimum (icons B4,

© 2010. John Benjamins Publishing Company
All rights reserved

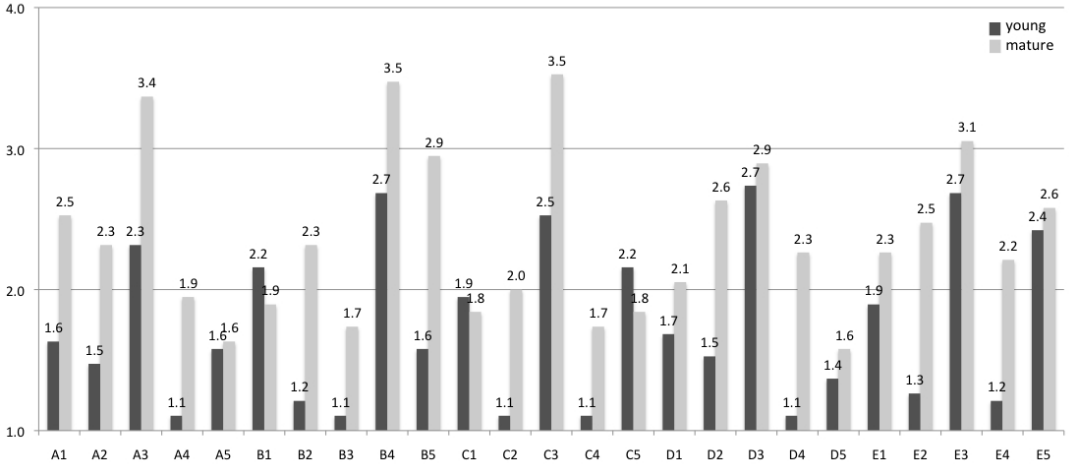

Figure 2. Average comprehension score for each icon compared between mature and young users. The lowest score denotes the best performance (one stands for "full comprehension," two stands for "partial comprehension," three stands for "wrong comprehension" and four stands for "no comprehension").

C3, scored 3.5) and maximum (icons A5, D5, scored 1.6) comprehension scores achieved by the older user group were worse than those for the younger user group (icons B4, D3, E3, scored 2.7, and icons A4, B3, C2, C4, D4, scored 1.1, respectively). In addition, we realized that if we had excluded from the study the youngest of the mature users (those aged between 30 and 45) then the average scores would be much worse for the older user group.

We also employed a t-test (two-tailed, unpaired) to measure whether there is any significant difference in the comprehension of each icon between the two age groups (between-groups design). According to the t-test tables (see Appendix, Tables A1, A2) there is a statistically significant difference for the highlighted pairs of icons, which account for 44% (11 out of 25) of the icons. Therefore, there are considerable differences in the comprehension of icons between young and mature users.

In particular (Figures 3 and 4), we found that there are three "Settings" icons that performed very well only for the young group (D4, E4, A4). Moreover, there is one "Organizer" icon (C1) that was significantly familiar and one "Files" icon (B5) that was significantly unfamiliar, only for the mature group. Finally, we examined the characteristics of icons that performed very well or very badly, regardless of the user group. The "Settings" and "Phonebook" icons were the most comprehensible for both user groups. The "Calls log" icon was the most problematic in many handsets, followed by icons that depict "Applications," "Internet," and "Files." In brief, it seems that mature users are only familiar with the icons that have immediate metaphors in the real world (e.g. phonebook, agenda). At the same time, young users are more familiar with icons that stand for the customization of mobile phones (e.g. "Settings"). In the next section, we provide further discussion of the results of this study.

© 2010. John Benjamins Publishing Company
All rights reserved

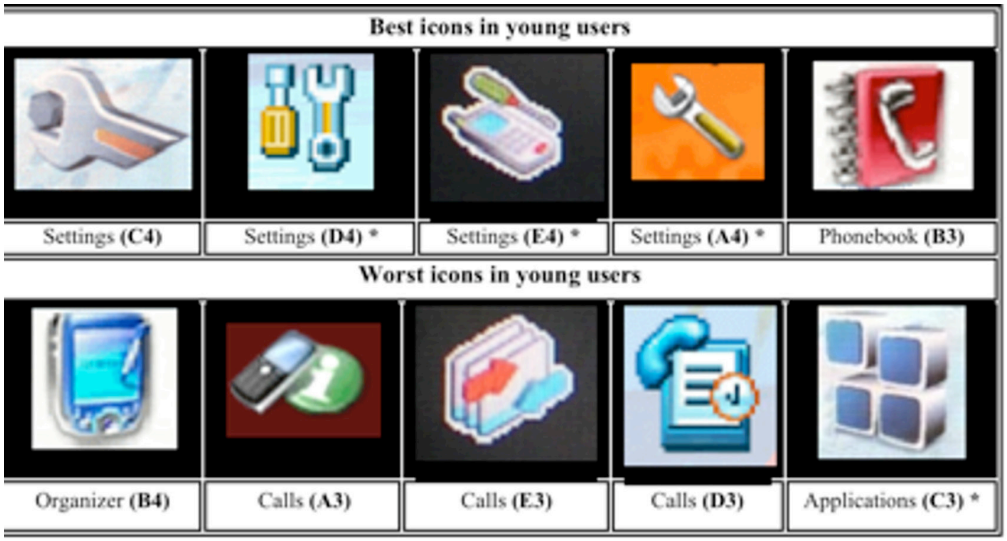

Figure 3. Best and worst five icons in young users' results (asterisk * denotes significant difference between the two user groups).

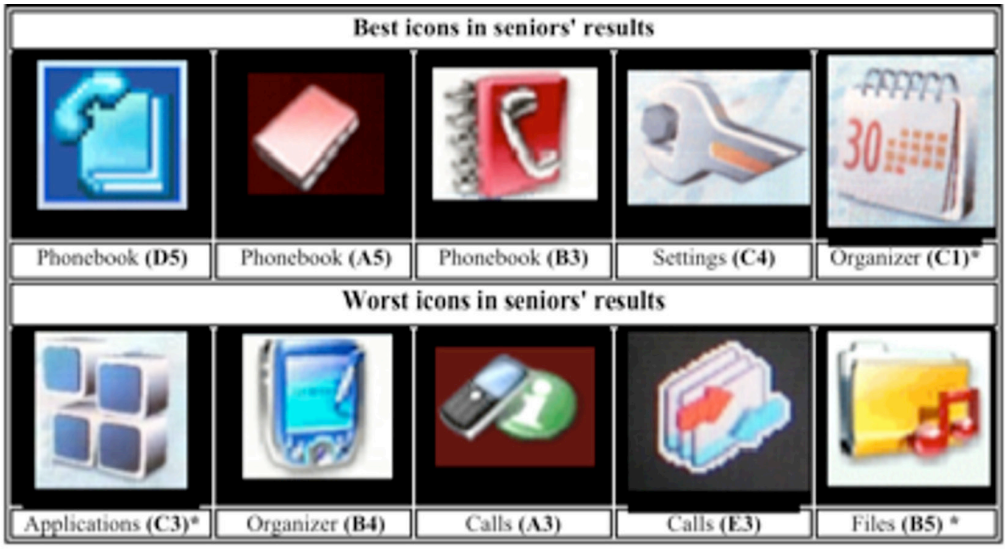

Figure 4. Best and worst five icons in seniors' results (asterisk * denotes significant difference between the two user groups).

© 2010. John Benjamins Publishing Company
All rights reserved

Discussion

We found that, in absolute terms, the performance of mobile phone icons is very much problematic for such a popular interaction device. For example, icons A3, C3, and E3 had very low comprehension rates, while B4 is obviously irrelevant for anyone who does not have experience with electronic organizers, such as "Palm." These icons represent high-level abstract concepts (e.g. "Applications") that do not have an obvious real-world metaphor. In addition, the meaning of "calls log" seems hard to depict visually. Firstly, the notion of "calls log" involves a direct reference to fax and ship logs (e.g. a notebook to record dates and events). Then, it seems that both fax and ship logs might not be as familiar as a metaphor should be for wide consumer understanding, because both are profession-specific. In addition, the use of notebook to depict a log might overlap too much with that of the phonebook or the agenda icon. Therefore, the "calls log" icon seems like a good candidate for redesigning and further evaluation of alternatives. Notably, this issue holds true regardless of age group, which was the main scope of this study.

Although the lower than expected absolute performance (i.e. very low average comprehension rates) could be attributed to the particular experimental set-up (icons were shown on a computer screen, one-by-one, and without any text labels), we also found significant differences between different age groups for the same icon. Therefore, there are icons that are much better than other icons for representing the same functionality for a particular age group. The absolute usability performance of each icon might not be representative of real world use (Ferreira, Noble, & Biddle, 2006), but we consider that the comparative results are trustworthy (within the sampling error and confidence interval reported), as long as the test samples (users and icons) match the real-world use. Since we expect that users perform much better in the real context of use, we provide some justification for the low comprehension scores of icons in this study.

The present study had several limitations that might affect the actual performance of each icon. First, in realistic scenarios of use, icons appear in groups on small mobile phone displays and not individually on a computer screen. Second, the majority of mobile phone menus provide a text label, in addition to the icons, which significantly affects the comprehension performance (Haramundanis, 1996). Finally, besides testing for comprehension of the meaning of each icon, it would have been interesting to also test for distinguishability (e.g. comparison between alternative versions of the same icon) and memorability (e.g. rerun the same study after a few days), which are good subjects for further research.

Nevertheless, the focus of this usability study was to provide evidence on the comparative performance of icons between different age groups. Indeed, this study provides empirical data for the long-established ideal of culturally adapted graphical user interfaces, which was originally proposed by Marcus and Van Dam (1991). Thus the results are not affected by the absolute performance and might be considered trustworthy in terms of relative performance between different age groups. In addition, Kim and Lee (2005) found that there are also differences in icon comprehension between different cultures. Therefore, it is expected that the results of this study might differ both in absolute terms for each icon, and also in comparative tests, if the study was repeated in the context of a different culture.

Conclusion and further research

We found that, despite widespread use and acceptance in desktop computing environments, there are still some issues in the design of icons that concern new application domains, such as mobile phones. Although there are few popular operating systems on the desktop, there are many more in the case of mobile phones and they

© 2010. John Benjamins Publishing Company
All rights reserved

might have to co-exist with simpler second-generation handsets for those users who prefer a no-frills handset (e.g. voice and text). With regard to the whole population, the performance of icons might remain an issue in the design of mobile phone UIs. Nevertheless, the issues revealed in this study highlight an opportunity to customize handsets and UI icons for particular user groups. Besides age differences, the mobile phone is a consumer product that might also be suitable for further customizations, such as style.

In conclusion, it is suggested that mobile phone icons are either standardized or customized for different age groups, at least with regard to those functions/objects that have become established (e.g. messaging, address book, calls log, mobile internet). Nevertheless, there might not be as much motivation to standardize mobile phone icons as there is in safety applications (Hancock et al., 2004). In the case that mobile phone icons are not standardized, then users will experience increased variation in performance across different handsets. In that case, previous studies have indicated that if standardization is not possible, then the designers should at least make the icons as learnable as possible (Moyes & Jordan, 1993).

We plan to continue this line of work with studies on different types of icons and user interface modalities, such as multi-touch displays and car dashboards, which present unique interaction situations. In particular, we plan to repeat the study for the icons found in mobile phone operating systems (e.g. Apple iPhone, Windows Mobile, Google Android, Palm Pre). Moreover, we would like to evaluate the differences among two-dimensional, three-dimensional, and animated mobile phone icons (Baecker, Small, & Mander, 1991). In addition, we would like to compare user preferences between different mobile phone user interface modalities, such as icons, voice, and gestures. Finally, it is suggested that a similar study takes place across time. For example, are handset brands (device manufacturers and wireless operators) changing icons just for the sake of change (style trends), which is very evident in most consumer products?

Acknowledgments

This study was supported by the European Commission Marie Curie Fellowship program (MC-ERG-2008-230894). We are also grateful to the participants of the study and to many constructive comments by the anonymous reviewers.

References

Arning, K., & Ziefle, M. (2007). Understanding age differences in PDA acceptance and performance. *Computers in Human Behavior*, 23(6), 2904–2927.

Baecker, R., Small, I., & Mander, R. (1991). Bringing icons to life. In *Proceedings of CHI '91*(pp. 1–6). New York, NY: ACM.

Barr, P., Noble, J., & Biddle, R. (2003). Icons R icons. In R. Biddle & B. Thomas (Eds.), *Proceedings of the Fourth Australasian User Interface Conference on User Interfaces 2003*, Volume 18 (Adelaide, Australia) (pp. 25–32). ACM International Conference Proceeding Series, vol. 36. Darlinghurst, Australia: Australian Computer Society.

Byrne, M. D. (1993). Using icons to find documents: simplicity is critical. In *Proceedings of CHI '93* (pp. 446–453). New York, NY: ACM.

Collins, B. L., & Lerner, N. (1982). Assessment of fire safety symbols. *Human Factors*, 24(1), 75–84.

Egido, C., & Patterson, J. (1988). Pictures and category labels as navigational aids for catalog browsing. In *Proceedings of CHI '88* (pp. 127–132). New York, NY: ACM.

Eisen, H. A. (1990). ICONER: A tool for evaluating icons. *SIGCHI Bulletin*, 21(3) (Jan. 1990), 23–26.

Ferreira, J., Noble, J., & Biddle, R. (2006). A case of iconic icons. *AUIC2006, Conferences in Research and Practice in Information Technology*, vol. 50, Hobart, Australia.

Green, Paul, & W. T. Burgess (1980). "Debugging a symbol set for identifying displays: Production and screening studies." The University of Michigan, Michigan, UM-HSRI-80-64.

Hancock, H. E., Rogers, W. A., Schroeder, D., & Fisk, A. D. (2004). Safety symbol comprehension: Effects of symbol type, familiarity, and age. *Human Factors*, 46, 183–195.

© 2010. John Benjamins Publishing Company
All rights reserved

Haramundanis, K. (1996). Why icons cannot stand alone. *SIGDOC Asterisk Journal of Computer Documentation*, 20(2) (May 1996), 1–8.

Kaikkonen, A., Kallio, T., Kekäläinen, A., Kankainen, A., & Cankar, M. (2005). Usability testing of mobile applications: A comparison between laboratory and field testing. *Journal of Usability Studies*, 1(1), 4–16.

Keijzers, J., den Ouden, E., & Lu, Y. (2008). Usability benchmark study of commercially available smart phones: Cell phone type platform, PDA type platform and PC type platform. In *Proceedings of the MobileHCI '08* (pp. 265–272). New York, NY: ACM.

Kim, J. H., & Lee, K. P. (2005). Cultural difference and mobile phone interface design: Icon recognition according to level of abstraction. In *Proceedings of MobileHCI '05*, vol. 111 (pp. 307–310). New York, NY: ACM.

Lohse, G. L., Biolsi, K., Walker, N., & Rueter, H. H. (1994). A classification of visual representations. *Communications of the ACM*, 37(12), 36–49.

Marcus, A. 1993. Human communications issues in advanced UIs. *Communications of the ACM*, 36(4), 100–109.

Marcus, Aaron, & Van Dam, Andries (1991). User-interface developments for the nineties. *ACM Communications,* 24(9) (September 1991), 49–57.

Moyes, J., & Jordan, P. W. (1993). Icon design and its effect on guessability, learnability, and experienced user performance. *Proceedings of HCI 93*, 49–60.

Nielsen, C. M., Overgaard, M., Pedersen, M. B., Stage, J., & Stenild, S. (2006). It's worth the hassle!: the added value of evaluating the usability of mobile systems in the field. In *Proceedings of the 4th NordiCHI '06*, vol. 189 (pp. 272–280). New York, NY: ACM.

Rogers, Y. (1989). Icons at the interface: their usefulness. *Interacting with Computers*, 1(1) (Apr. 1989), 105–117.

Sorenson, P. F., & Webb, J. M. (1991). Experimental evaluation of icon quality. *SIGCHI Bulletin*, 23(4), 78.

Tufte, E. R. (1990). *Envisioning information*. Cheshire, CT: Graphics.

Wang, H. F., Hung, S. H., & Liao, C. C. (2007). A survey of icon taxonomy used in the interface design. In *Proceedings of ECCE '07*, vol. 250 (pp. 203–206). New York, NY: ACM.

Wiedenbeck, S. (1999). The use of icons and labels in an end user application program: An empirical study of learning and retention. *Behaviour and Information Technology*, 18(2), 68–82.

About the authors

Charalambos Koutsourelakis is a postgraduate student in the Department of Graphic Design and Multimedia at the Hellenic Open University, Greece. He holds a Diploma in Computer Science and an MA in Multimedia.

Email: koutsourelakis@gmail.com

Konstantinos Chorianopoulos is Marie Curie Fellow and Lecturer in the Department of Informatics at the Ionian University, Corfu, Greece. He holds an MEng (Electronics and Computer Engineering), an MSc (Marketing and Communication), and a PhD (Human–Computer Interaction).

Email: choko@ionio.gr

Contact

Ionian University, Department of Informatics
7 Platia Tsirigoti
49100 Corfu
GREECE

© 2010. John Benjamins Publishing Company
All rights reserved

Appendix A

In a between-groups experimental design (age group, comprehension rate), we measured statistical significance for the average value differences in the icon comprehension data between young and mature users (19 users in each of the groups). In the following tables we have highlighted those cells that have scored less than 0.09 in the t-tests. In this study, the results should be representative of the respective population within a 91% confidence interval.

Table A1. T-test results for Sony Ericsson, Samsung, Nokia

Brand	Sony Ericsson					Samsung					Nokia				
Icon	a1	a2	a3	a4	a5	b1	b2	b3	b4	b5	c1	c2	c3	c4	c5
Two-tailed unpaired t-test (p<)	0.016412	0.114728	1.000000	0.080198	0.244800	1.000000	0.039001	0.425305	0.524127	0.000251	0.060688	0.802663	0.008079	0.114728	0.035195

Table A2. T-test results for Motorola, Vodafone Live

Brand	Motorola					Vodafone Live				
Icon	d1	d2	d3	d4	d5	e1	e2	e3	e4	e5
Two-tailed unpaired t-test (p<)	0.770314	0.007911	0.249996	0.006739	0.853448	0.694714	0.088107	0.764182	0.023867	0.261702

© 2010. John Benjamins Publishing Company
All rights reserved

Appendix B

The table below portrays the number of answers for each of the four levels of coding (no, wrong, partial, or full comprehension) for the two user groups.

mature

a1	a2	a3	a4	a5	b1	b2	b3	b4	b5	c1	c2	c3	c4	c5	d1	d2	d3	d4	d5	e1	e2	e3	e4	e5	
3	7	0	9	10	2	7	12	0	1	3	9	0	11	7	3	1	2	8	11	2	6	3	7	5	full
8	3	2	5	7	17	5	1	0	6	16	5	1	4	10	12	10	2	4	5	13	5	1	6	2	partial
3	5	8	2	1	0	1	5	10	5	0	1	7	2	0	4	3	11	1	3	1	1	7	1	8	wrong
5	4	9	3	1	0	6	1	9	7	0	4	11	2	2	0	5	4	6	0	3	7	8	5	4	nothing

young

a1	a2	a3	a4	a5	b1	b2	b3	b4	b5	c1	c2	c3	c4	c5	d1	d2	d3	d4	d5	e1	e2	e3	e4	e5	
8	15	7	17	14	1	17	17	5	10	1	18	8	17	10	7	10	3	17	15	7	17	6	17	9	full
10	0	1	2	0	15	0	2	0	7	18	0	0	2	1	11	8	1	2	1	9	0	0	1	0	partial
1	3	9	0	4	2	2	0	10	2	0	1	4	0	3	1	1	13	0	3	1	1	7	0	3	wrong
0	1	2	0	1	1	0	0	4	0	0	0	7	0	5	0	0	2	0	0	2	1	6	1	7	nothing

© 2010. John Benjamins Publishing Company
All rights reserved