ASSESSMENT OF DIGITAL ICONS

An ergonomic approach.

Marina CARDOSO[1], Sandra RAMALHO[1]
[1] Universidade do Estado de Santa Catarina

ABSTRACT

The key question addressed in this paper is how the designer can and should ensure that the solutions are efficient and designed to meet the goal originally proposed, taking into account users as the centerpiece of the project, since solutions should be designed for them. In order to provide a support for the assessment of digital icons, this article, as part of a larger research, aims at bringing together ergonomics and usability methods and tools and discussing their interactions with the main methods and tests of informational ergonomics, but that have been used in order to measure the comprehension of digital icons by their users, assessing its procedures and how the effects of the usage context have been included in each approach.

KEY WORDS

Icons, comprehension tests, ergonomics.

1. INTRODUCTION

In order to allow the process of transposing the command lines into graphical interfaces in the history of information technology, a short, concise and easily understandable language was necessary. From this perspective, the visual language of pictographs was greatly prominent in this paradigm shift for acting with elements of synthesis, focusing on communication in the case of interfaces with a lot of content, small area and various possible actions to be performed. The use of the so-called "icons" thus became a necessity in this process and provided a very useful tool, not only in the context of large accumulation of information in computer interfaces, but when the support are interfaces with small proportions, such as mobile phones and tablets, which have little room for interaction and must be precise and clear.

Icons are, nowadays, indispensable to the communication process at computerized work environments, always with the intention of facilitating the dialog between user and system, in an iterative process of communication. To investigate this visual communication and its quality in digital media software meets the improvement of the ergonomics and usability of these interfaces, directly impacting on how the user uses them, the ease of such use, and the results and productivity they engender.

2. TESTING DIGITAL ICONS

During the bibliographic research for this paper, we noticed a gap in the design methodology for the design of icons specifically covering steps, procedures and techniques both to design and to assess existing projects, and that require specific details and parameters. The digital nature of the icon, its characteristics, its reduced size, the fact that they communicate actions and visually represent concepts, are displayed on screens and monitors, etc., to generate a series requirements that make the project of icons complex and with specific limitations. There are considerations and recommendations for the design of computer icons and, as well as ergonomic criteria for pictograms in general, established according to Iida and Padovani

apud [Falcão 2006]. But it appears that even contemplating the ergonomic criteria for graphical symbols, some icons do not convey information effectively.

[Horton 1994] notes that following principles can help in this design activity, but does not guarantee a result free from faults. And the way to approach this goal is to test the icons and review them in an iterative process for several cycles until it reaches an efficient and reliable design. However, this is a costly process that takes time and needs to be undertaken from the beginning to the end of a project, and not only as it approaches the deadline for implementing the product / interface, when the cost of correcting problems and reviewing icons increases considerably, as well as teams' and stakeholders' resistance to change.

2.1. Stages and types of tests

Late tests tend to fix only small problems and result in superficial and cosmetic repairs. Therefore, testing constantly and from the beginning is what [Horton 1994] advocates. The author divides the necessary testing into three stages which are overlapped in time and method.

2.1.1 Formative Stage

It is the first stage, in which the goal is to discover what will work best by focusing on rapid prototyping, without refinement, an empirical and iterative development, and involves four main activities: Building, which is the creation of prototypes without refinement; Test, or checking if it works; Analysis, understanding why it failed and what improvements should be made; and Re-design, which is to incorporate what has been learned. These steps should form a cyclic sequence performed several times, not in order to demonstrate the functionality of the icons, but to find their failure and understand how to improve them.

In the field of usability studies, methodological strategies usually adopted in the assessment of user-computer interfaces for their optimization, according to [Queiroz 2006], are also called formative, mainly characterized by continuous processes of research adaptable to both the overall progress and specific aspects of the interface, emerging as a result of research and modification of parts of the interface.

Based on the chronology of phases proposed by [Horton 94], we can relate to each stage most methods that have been used by researchers for digital icons and classify them according to the time of application. In [Formiga 2002], we found a survey of comprehension assessment methods of pictograms from informational ergonomics, but as can be seen in several studies in the area of computer interfaces, they have been applied to evaluate digital icons.

The first one is called Test of Comprehension and shows the degree of correct comprehension of each symbol and, according to [Formiga2002], is the most important test procedure for the development of symbols for public information. The researcher selects a symbol for each concept to be communicated and only the picture is shown, usually printed on individual cards, and the individuals should attribute the meaning that seems most convenient. It is worth noting that only one symbol should be presented per concept in order to avoid comparisons. As an example of using this test to evaluate the comprehension of digital icons, one can point out the study by [Shiraiwa2008], which used this assessment to apply icons to public Internet portals. [Falcão2006] also makes use of this test with some variations to assess the comprehension of digital icons on a DVD already on the market. In these two articles, the icons tested were already implemented on their respective interfaces, Internet and DVD, and thus did not have a formative character.

Interestingly, despite their digital nature, the "pictures" of each icon in both cases were tested with the subjects participating in the research printed on white paper and out of their contexts of use, in accordance with the procedures described in the informational ergonomics methods. Since they relate to signs and signalization, it is common practice to introduce them to subjects in printed form in dimensions similar to those used in signs and warnings.

A second common test is called the Method of Production. According to [Moraes & Formiga2000], in it, participants draw by hand concepts that were expressed verbally or in writing by the researcher. The goal is to analyze the variation between repertoires, according to culture, the social or intellectual level of the participants. It can also be used to assess the difficulty of representing each concept and estimate which graphic elements are used more frequently.

2.1.2 Comparative Stage

According to [Horton 1994], the tests suggested for this phase of the project aim to measure the performance of two or more alternative projects for icons. They require clear and objective criteria to efficiently evaluate the results, which should be well developed and established by consensus among the design team before the test is performed and must help plan how the results will be measured. Comparative tests are most valuable when performed in the intermediate stage of development because they validate ideas generated beforehand, but that are not finished - when prototypes are still subject to refinement and verification prior to the assessment of the complete success of the product. In informational ergonomics, two assessment methods could be classified as suitable to be applied at this stage of icon development.

The first is the Re-identification Method, in which [Moraes & Formiga2000] state that a series of graphic symbols be presented to the test participants so that they speak or write down the meaning of each one. The use of this test is desirable in the middle phase of the project because there are several prototypes to communicate the same concept, and since none of them has yet been refined, their preparation did not demand too much time or effort. Thus, the user can already recognize the icon and assign it with some meaning, and ultimately the researcher or designer has an indication of which icons should be discarded and which should be refined and therefore stand a good chance of being implemented.

The second method, which may be suitable for this stage as well as for the next one, is the Test of Choice. Similar to the test described above, the participants individually elect the symbol that seems to be the best choice for every concept from a range of alternatives presented to them. The assessment results in a percentage by order of preference for the symbols of the same concept. According to the research objective, one, two or three symbols should be elected according to the ranking.

2.1.3 Assessment Stage

It is the phase in which efforts are focused on ensuring the quality of icons developed so far. It should outline the differences between the performance of the designed icons and objectives initially proposed for use, and like all tests, it should establish clearly defined criteria that measure how suitable each icon is. [Horton 94] presents four aspects that must be allocated so the tests are clear and unambiguous: (1) the sample selection, who will be the participants of the test and how to simulate or approach this profile; (2) the scenario, tasks and procedures that the subjects should perform; (3) the method for measurement, or how successes and failures will be measured and (4) the assessment criteria, which is the level or degree of correct results acceptable for the project to be considered sufficient and approved.

HCI assessment methods can provide useful tools, not only in the final stage of a project for icons, such as the so-called Usability Inspections, assessment strategies that are based on analysis and judgment of projects by assessors (ergonomics, software engineers, etc.) that investigate aspects of usability according to a set of criteria, recommendations, rules or heuristics. [Queiroz 2006] points out that some authors insert in these assessments the end-users with knowledge of the context or the task assessed and other professionals with assessing skills. Some of these strategies may be applicable to evaluate projects for icons such as: Systematic Reviews (Walkthroughs), which may occur during the different stages of a system's development, assuming various forms and involving different skills, though its fundamentals remain

unchanged: a group of experts assembled to review and discuss a specific product. Another strategy is the Inspection-Based Design Guidelines, Styleguides and Standards Inspection, which is presented in the form of suggestions and techniques that bring together the knowledge and opinion reflecting current methodological principles applicable to interface projects. And finally, the Heuristic Assessment, which is to study a product in depth, conducted by experts, based on personal experiences and knowledge in order to identify usability problems.

2.2. Ergonomic-cognitive and behavioral assessment methods

The field of ergonomics provides, according to [Stanton2005], a number of methods and techniques that offer a structured approach to analyze and solve design problems, focusing on the emphasis on the analysis of human performance, safety and satisfaction. They may be arranged in five sets according to the type of data they collect: (1) methods that collect data about people represent the collection of physical and physiological data and psychological skills or abilities (cognitive); (2) methods used in developing systems, e.g. data collection using the product with the current design and another one with a new design proposal; (3) methods to assess the performance of a human-machine system represent the collection of data from quantitative and qualitative measures (4) methods to assess demand and effects on people, e.g. the collection of data of the effects of weli-being in a person who performs a particular task, examined at short and long term; (5) and finally, the methods used to develop an ergonomic management program, the supporting strategies and management.

The methods that collect data about individuals can be especially useful for the assessment of human-computer interfaces and the language of their icons. The cognitive and behavioral methods find its theoretical foundation in the disciplines of psychology. These methods provide us with data on perception, cognitive processes and responses to individual or collective stimuli. Information about errors, task performance, goals, decisions, preferences and workload can be gathered in many situations of interaction with objects and their most diverse interfaces.

These techniques provide a wide variety of information about the performance of users. Some of these methods of cognitive analysis, classified as General Analysis Methods, identify general aspects of human factors. According to [Stanton2005], they are: Observation, Individual Interview, Verbal Protocol, Repertory Grids and Focus Group. The methods of ergonomics can be of great help to the team that designs icons, whether using the methods described above in full, isolated or combined, or with informational ergonomics tests. It is possible to note that the comprehension tests indicated somehow already use these techniques from ergonomics, since at applying a questionnaire or assigning a task to the subject, an interview is already being conducted indirectly, taking into account their reactions through observation or discussion. These are ways of gathering data and information that enrich the findings and give clues on how to efficiently achieve our and the user's goals. These are the details about which [Horton 1994] discusses, calling attention to certain precautions that make the tests more realistic.

2.3. Assessment recommendations

Some precautions must be observed when making any type of assessment of digital icons in order to reliably simulate their actual use in operation, when applied to an interface. Making the tests realistic, especially in the final stage, when it is possible to insert them in the interface even if it is not finished, in a context closer to reality, is an important aspect that contributes to increasing the chances of success of the final product and reducing the differences in context, therefore also reducing rework and rough adjustments. Horton points out a few recommendations:

2.3.1 Recruiting reliable test subjects

[Horton 1994] refers to the sample of subjects, which, when well-defined, inclusion criteria must be accurately established, aiming at a homogeneous group that reflects the preferences and behavior of future users. In the case of an existing product, these users should be recruited. Some valid criteria include

selecting people who have physiology (vision, reflexes, etc.), a psychological profile (motivation, skills, etc.) and related professional and educational training, besides having no connections or interests in the success of the product and no knowledge beyond the required for the common user, since it could lead them to overcome or ignore flaws that will hinder or disrupt real users.

2.3.2 Training test subjects

In the case finding subjects similar to real users proves to be impossible, the group should be instructed to simulate and perform the role required. The level of prior knowledge, such as training, work experience or with previous or similar versions of the products, must be balanced as much as possible within the selected sample.

The type of task assigned is also important because it should be clear and realistic, so that miscomprehension does not result in a diagnosis of low-quality interface and icons. Therefore, it is important to clearly define and explain: the purpose of the test; which actions must be performed; the resources available; whether or not online documentation can be used etc.; the deadline for completion etc.; performance objectives, i.e. the criteria for success and information about how they know when the result is achieved; provision of realistic data to be objectively used in the task that does not require skills other than those that will be assessed by the test.

About the observation technique, [Stanton et al. 2005] state that just observing someone can already affect performance. The presence of an observer can get people to demonstrate knowledge about how the product should be used instead of demonstrating how it is actually used. This is especially true when the subjects are taken out of context and of the environment in which they use the product. [Horton94] proposes some measures to minimize these effects. The author suggests testing the subjects in pairs, since it causes less awkwardness; taking the test at the subject's workplace and emphasizing again that the icons are being assessed, and not the people involved in the test.

2.3.3 Subjective Signs

The qualitative results of comprehension tests and assessments state the efficiency and functionality assigned to each icon, but other subjective signs can tell us the reasons for the success of certain icons and the failure of others. In the case of in-person tests, as in the studies cited, where subjects are observed and monitored, subtle clues to their emotional responses and expressions of satisfaction can be detected by the observer. [Stanton et al. 2005] state that observing people as they interact with interfaces or devices to perform tasks is a way of gathering data about errors and the time of performance and provides insight about the difficulty with which the task is carried out.

According to [Stanton et al. 2005], even though our observations are meant to obtain a "feeling" of what is happening, we must be concerned about how to communicate this "feeling". We can either rely on generalizations about a product and say that "most people found the product easy to use", or do it by proving a report. It is necessary to present a record that could be construed as unequivocally as possible, based on a set of observations that can be unambiguously defined. It is possible to choose to record some comments made by the person during the activity, which may consist of a description of what they are doing, the line of thought of their plans and intentions, justifications they provided etc., which could be called a verbal protocol, or it could be done by means of discussions with other users of the product as a Focus Group.

3. FINAL CONSIDERATIONS

In face of the absence of a specific and consistent method for assessing digital icons, the frequent use and application of comprehension tests of symbols from informational ergonomics, signaling, etc, can be found in several studies and research in the field of icon design.

These tests do not foresee applications for digital interfaces for icons that are designed to keep characteristics of legibility, clarity and recognition when they are inserted into the interface, as they will be

seen on screen by means of a light-color system, reduced in size as is typical of icons, often on colored backgrounds etc. These aspects have a significant influence in communication, in their comprehension before the interface context, and place apart the handmade representation, a simple design done by a layman, from the end result that will be processed by the designer and come to the computer screen, integrated to the interface. The necessary adjustments so that "idea" adapts to the language of the digital icon can be so many and so deep that a second assessment, the final icon design, may prove to be insignificant and no longer recognized by the users who initially appointed that picture as the one that best represents a certain concept in their personal opinion.

The problem related to the application of methods not suitable for digital icons is the inconsistency of results, since they do not take into account the effects of context, these influencing aspects mentioned, that directly or indirectly affect the judgment of the studied subjects.

The ergonomic and usability tools available can assist in these investigations, providing more reliable and concrete data about the human factors that affect the results and are subject to measurement. The field of ergonomic assessments of digital interfaces can contribute to the assessment of icons and provide previously established bases, making tests and methods more robust and reliable. According to [Horton 1994], in order to be useful, testing should reliably simulate the actual use of the product. The test subjects should act as real users, and the prototype should evoke the same responses to the current product. This leads us to believe that applying comprehension tests by means of techniques that include showing the subject pictures printed on paper alone, or in different dimensions, different media and out of its context, which is the software interface, can result in responses that from the moment the icon is integrated into the interface are no longer valid.

4. REFERENCES

[Falcão2006] Falcão, Franciane. Avaliação de ícones em interface digital a partir de teste de compreensibilidade e método de produção. Anais ABERGO 2006 (Curitiba/PR: 29 out - 02 out 2006).

[Formiga2002] Formiga, Eliana. Ergonomia informacional: compreensibilidade de símbolos para sinalização de hospitais públicos e unidades de saúde no Rio de Janeiro. Rio de Janeiro, PUC, 2002.

[Horton94] Horton, William. O Livro do Ícone: Símbolos visuais para sistemas de computador e documentação. New York: John Wiley & Sons, 1994.

[Moraes&Formiga2000] Moraes, Anamaria de; FORMIGA, Eliana de L. Avaliação da compreensibilidade de ícones e símbolos através da ergonomia informacional. I Encontro Pan-Americano de Ergonomia/X Congresso Brasileiro de Ergonomia/IEA Endorsed Conference. Salvador. 2000. Anais da ABERGO. ABERGO. Rio de Janeiro. 2000.

[Queiroz2006] Queiroz, J. E. R. Abordagem Híbrida para a Avaliação da Usabilidade de Interfaces com o Usuário. Projeto de Pesquisa da Universidade Federal de Campina Grande, 2006.

[Shiraiwa2008] Shiraiwa, Juliana Couto Silva. UNIVERSIDADE FEDERAL DE SANTA CATARINA Programa de Pós-Graduação em Design e Expressão Gráfica. O reconhecimento de pictogramas em interfaces gráficas digitais pelo usuário idoso: o caso do Portal Pró-Cidadão da Prefeitura Municipal de Florianópolis. Florianopolis, SC, 2008.

[Stanton et al. 2005] Stanton, N. A., BABER, C., YOUNG, M.S. 2005. 'Observation' in N.A. Stanton et al. (eds) Handbook of Human Factors and Ergonomics Methods. Londres : CRC.

[Stanton2005] Stanton, N. A. 2005. 'Behavioral and Cognitive Methods' in Handbook of Human Factors and Ergonomics Methods. Londres : CRC.