Human Computer Interaction

Dana Spiegel
10/22/1998
danas@mit.edu

Abstract

This paper discusses the research that has been done in the field of Human Computer Interaction (HCI) relating to human psychology. A brief overview of HCI is presented. Specific examples of research in the areas of icons and menus are then reviewed. The results of these experiments and the predictions they make about general human psychology and specific human interaction with computers is discussed. Mental models of user interface interaction are discussed and compared to mental models of real world object. Finally, future directions for research are proposed.

Introduction

Computers have become, for better or worse, integral parts of our lives in every respect. We use them to communicate with others, to write our papers, to monitor the arrival and departure of airplanes from a control room, and to play games, to select but a few of their many uses. Each of these uses requires us, as humans, to interact with these machines. This interaction requires a mode of communication that was generally unknown prior to the introduction of computers into our society. Indeed, this mode of communication was almost entirely non-existent prior to the introduction of the Graphical User Interface (GUI) popularized by the Macintosh computer (Levy *1994)*.

The GUI works by using a representational form of communication to inform the user about the state of the machine and to allow the user to tell the computer which operations to perform. As Steven Levy (1994) succinctly states: “Metaphor, it turns out, is the key to making computers comprehensible.” The truth of this statement is exhibited by the fact that so many computer users utilize this notion of metaphorical thinking without questioning its validity. One merely needs to listen to people discussing how a computer program works to realize this:

> “How do I print my letter?”
> “Open the letter, the select print, and hit ok.”

Clearly these people are discussing a particular action to be taken at a computer. However, their conversation doesn’t hinge on the actual motions involved:

> “Place your hand on the device next to the keyboard. Move that device, which causes movement of the black arrow on the bright TV-like device on your desk, so that the black arrow is on top of the white square, which has text below with the name of your letter, and click twice, in rapid succession, on the button at the top of the device in your hand.”

This description of the interaction of a user with a computer sounds ridiculous to anyone who has used a GUI before because it does away with the notion of metaphorical thinking. Yet it is as correct as the statements made above.

When we talk about using a computer, we talk metaphorically: “my letter” is not the letter itself, but rather the icon on the screen that *represents* the letter. Opening it encapsulates the action of moving the cursor (in itself an abstract entity on the screen, represented in reality by the mouse) over the icon that says “my letter” and double-clicking the mouse button—the computer equivalent of slitting the top of an envelope, removing the paper inside, and unfolding it.

Each of these metaphors helps us, as computer users, maneuver through interaction with this device sitting on our desk. The task of making this interaction seem so effortless is tremendous, especially when some cases are anything but effortless. It draws from many different areas of study—human physiology, mental representation, motor coordination, computer programming—all of which fall under the umbrella of the term *Human Computer Interaction* (or HCI).

To understand the fundamentals of HCI, and how it relates to human psychology and physiology, we ask: “What can we learn about human perception and cognition from studying the way in which humans interact with computers?” To answer this question, this paper investigates two aspects of the computer interface—*icons* and *menus*—and reviews research that has been done on these interface elements in conjunction with how humans use them. These studies relate back to psychology by explaining the results in terms of human mental abilities.

Icons

Webster's Revised Unabridged Dictionary defines an *icon* as “a sign (as a word or graphic symbol) whose form suggests its meaning.” (WWWebster Dictionary 1998) The notion of an icon is a very old one, dating back to the 8th century, when an icon was a representation of a religious figure. (*Encyclopædia Britannica* 1998) On a computer screen, an icon is any graphical representation which refers to an object or an action that can be performed. (Rossi & Querrioux-Coulobmier 1997).

Why do computers use icons? Clearly there are multiple representations for objects in the real world, and the situation is the same on a computer screen. Computers could simply use words, for example "Print," to represent the action of printing a document, or "october_budget.xls" to represent a file. So-called command-line interfaces, such as Unix, do represent commands and files in this way. However, both research and experience show that icons are more powerful representations that allow immediate recognition, increase the speed at which users find objects on the screen, and in some cases conserve screen space when representing objects or actions. (Rossi & Querrioux-Coulobmier 1997)

Furthermore, there are some objects that cannot be represented in textual form. Tullis (1988) talks about the application "MacPaint", which is used to create and represent graphical images. He suggests trying to represent the picture in *Figure 1* using alphanumeric characters. This is clearly a very difficult, if not impossible, task. After all, a picture is worth a thousand words.

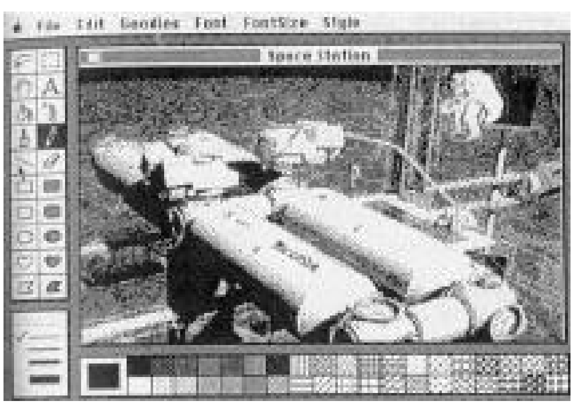

Figure 1 - MacPaint

Icons are not necessarily the best way to represent data, however. Rossi & Querrioux-Coulobmier suggest that "the relationship between an icon and its meaning should be automatic and consequently independent of any learning." (1997) This means that for an icon to work better than another representation (such as a textual description), it needs to evoke implicit understanding of the meaning of the icon.

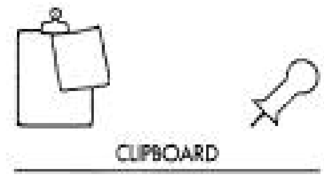

Figure 2 – A clipboard icon

Yamakawa, Miller, & Huchingson (1997) showed that this indeed is the case. In their experiments, they presented subjects with a computer term (either an object, such as “clipboard”, or an action, such as “erase”) and four alternative but similar icons from which subjects had to choose the one that best represented the term. They found that the icons that worked best are those that most concretely represented the object or action and were most visually descriptive. *Figure 2* shows an example in which the computer’s clipboard is best represented by an icon of a physical clipboard. The experiments by Yamakawa, Miller, & Huchingson (1997) showed also that three-dimensional icons were preferred over two-dimensional icons.

These experiments point to specific human cognitive characteristics. Most people live in a three-dimensional visual world. Even those that have use of only one eye or, because of some deficit, have no stereo vision make use of depth cues, such as overlapping figures and object constancy, which are concepts studied by the Gestalt psychology in the early part of the century. (Goldstein 1996) As such, when dealing with the world, three-dimensional representations of objects, so long as they are clear, are the most comfortable. It is therefore not surprising that three-dimensional representations on the computer screen work best for representing objects.

Interacting with the real world requires mental models to predict the outcomes of the actions. (Gray 1999) Extending this concept, interacting with objects and performing actions on a computer require mental models for how that interaction should occur. Icons that are grounded in real life (such as a clock to represent time) require less cognitive work to develop these models, since they are already developed from interaction with the real world.

Menus

A *menu* is a group of visually similar words and/or icons on a computer screen that allows the user to select an action to be performed. Unlike a command-line interface (such as Unix or MS-DOS), “menus have the advantage that users do not have to remember the item they want, they only need to recognize it” (Preece 1994).

To utilize menus, which often appear on the screen in a group with each menu containing a number of menu items (see *Figure 3*), the user must find the menu item that corresponds to the desired action, and then select that menu item. This process involves both memory (the user must remember the mapping of menu item to action performed) and visual search (the user must compare the needed menu item to the menu items available and find the correct one). This process is called entity matching (Preece 1994).

Mills & Prime (1990) studied the effects of rectangular versus circular menus as well as moving versus static menus. Their results showed that for menus with few items, circular menus (see *Figure 4*) are the most efficient form, as all of the menu items are equidistant from the center of the circle, which is where the cursor is located when menu item selection begins. Furthermore, subjects performed much better with static menus than they did with moving menus (where the cursor selecting the menu item stays at a certain location on screen and the menu moves up and down as the subject moves the mouse). This result might be attributable to the increased memory requirements that are needed to traverse moving menus: not only must the subject perform the entity match discussed above, but they also must keep track of which menu items are both above and below the cursor.

Learning effects also play a key role in performance of subjects. Over time, static menus can be learned better than moving menus. (Mills & Prime 1990) A severe example of this was performance on rectangular static menus: Mills & Prime (1990) noted that subjects tended to pay less attention to this type of menu “because of their familiarity with it, as to novel styles.”

Figure 3 – A menu bar

The order of menu items within a menu is also a factor in how subjects perform, and can point to certain memory and visual constraints. If a user already knows the menu item to select (which would mean that the task solely involves searching), alphabetically arranged menus work best. In the case that a user is either unfamiliar with the menu structure, or is not sure of the mapping of menu item to action, then functional organization—which groups menu items that perform similar actions within the same menu, and gives the menu a descriptive title—allows subjects to be most efficient. (Mehlenbacher, Duffy, & Palmer 1989)

Mental Models

Mental models are representations within our minds that allow us to predict the outcome of actions in the world. Any human interaction with the world requires one. The mental model can be as simple as understanding how gravity works when you drop a ball, or as complex as the way you fly a plane. These models allow us to function as thinking beings in the world by providing us a way to understand how the world works. (Gray 1999) (Our exact understanding of how a particular thing works may actually be different from the way a thing actually works; this however is a different topic of discussion).

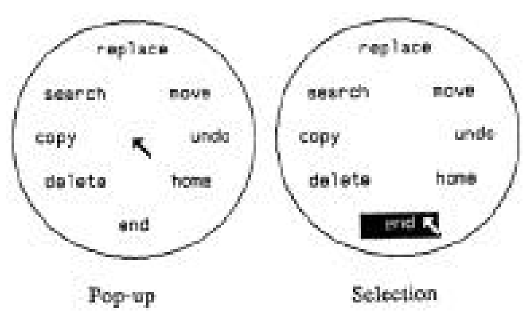

Figure 4 – A circular menu

Humans function similarly when interacting with machines: there is a mental model that predicts how a machine will react given a particular action taken by the user. Though these models describe the behavior of software, which is a non-physical object, they function is the same way. It is important, then, to leverage off of our understanding of mental models in the real world to understand how mental models work in the virtual world of computers.

Mental models of the functioning of an object evolve over time as a person interacts with that object. In the case of computers, people's mental models of a computer's behavior starts to develop upon the person's first use of a computer. Computers have the capacity to present an interface to the user that is unlike anything the user has encountered previously, and so can require the formation of a completely new mental model. Interface designers have taken this fact to heart, and have created interfaces that in many ways behave in a manner similar to real world objects. (Carroll 1991)

By utilizing, at least in part, mental models that have already been developed, interface designers make computers easier to learn. People have a great deal of experience using menus in restaurants. Utilizing a menu to select actions to perform on a computer transfers the mental model people have for menus in the real world into the domain of computer interaction. Indeed, menus are varied, and can be complex interfaces to navigate (as discussed above), yet people use them to interact with computers quite fluidly. Carroll and Rosson (1987) have performed experiments on the ease with which people learn to use menus and have found that much of the overlap between real world and virtual world menus aids in the learning process.

Present and Future Directions

The research done in the field of Human Computer Interaction, in which studies of different icon and menu designs are but a small part, can be used by programmers and designers to create computer interfaces that complement humans' abilities to perceive and understand these visual devices. For example, the absence of moving menus in the Macintosh Operating System and Windows reflect the findings by Mills & Prime on the efficiency of static versus moving menus.

It is curious that most of the research done in the field of HCI starts with the design of some interface, and then progresses to usability testing. There is very little initial thought given to how humans actually process information; instead studies seek to find this out by testing performance on varied interfaces. There is a critical mass of research that has already been done, such that this second method of designing *around* human perception should become more the norm.

A good example of this is color perception: much is understood about the perception of color and the gamut response of the human eye. (Goldstein 1996) A gamut is the compete range or scale of some attribute, in this case color. (WWWebster Dictionary 1998) The gamut response of the human eye is the color range that the eye is sensitive to, which is non-linear. However, computers provide a linear color gamut that does not take into account human perception. The Mac OS provides a color space that contains equal amounts of red, green, and blue. Though a computer monitor can display all of these colors, the human eye has a greater range of color perception in red than in green and blue. The interface for color selection should take this non-linear response into account (Goldstein 1996).

Conclusion

The subject of Human Computer Interaction is very rich both in terms of the disciplines it draws from as well as opportunities for research. Discussed here was just a small subset of the topics contained within HCI. The study of user interface provides a double-sided approach to understanding how humans and machines interact. By studying existing interfaces (such as the graphical user interface or the command line interface), we gain an understanding of how the human mind processes information. We gain insight into how human memory deals with the information presented, as well as its limitations. We also better understand how humans use the visual subsystem to find information.

Alternatively, from studying how human physiology and psychology, we can design better interfaces for people to interact with computers. Work in this domain is only beginning (indeed the number of papers written on this topic has increased in the past few years), and there is much that we don’t yet know about the way the human mind works that would allow more perfect user interfaces to be built.

The study of mental models that allow humans to use these interfaces provides a secondary, higher level approach to understanding Human Computer Interaction. Though mental models are far from concrete objects, we do understand to a how they are used to allow people to interact with the world. By studying them both in the domain of the real world, as well as the domain of the virtual world on a computer screen, we can gain insight into how these models are formed, and how they can be moved from domain to domain.

Bibliography

Carroll, John. M. (1987) *Interfacing Thought*. Cambridge, England. Cambridge University Press.

Carroll, John. M. (1991) *Designing Interaction*. Cambridge, MA. MIT Press.

“icon” Britannica Online. <http://www.eb.com:180/cgi-bin/g?DocF=micro/286/57.html> [Accessed 11 October 1998].

“icon” WWWebster Dictionary. <http://www.m-w.com/cgi-bin/dictionary?book=Dictionary&va=icon> [Accesses 11 October 1998]

“gamut” WWWebster Dictionary. <http://www.m-w.com/cgi-bin/dictionary?book=Dictionary&va=gamut> [Accesses 11 October 1998]

Goldstein, B. (1996) *Sensation and Perception*. Pacific Grove, CA. Brooks/Cole Publishing Company.

Gray, Peter. (1999) *Psychology*. New York, NY. Worth Publishers.

Levy, Steven. (1994). *Insanely Great*. New York, NY: Viking Penguin.

Mehlenbacher, B., Duffy, T., & Palmer, J. (1989) “Finding Information on a Menu: Linking Menu Organization to the User’s Goals”. From *Human-Computer Interaction*. Hillsdale, New Jersey. Lawrence Erlbaum Associates.

Mills, Z. & Prime, M. (1990) “Are All Menus the Same? – An Empirical Study”. From *Human-Computer Interaction — Interact ’90*. North-Holland. Elsevier Science Publishers.

Preece, J. (1994). *Human-Computer Interaction*. New York, NY: Addison-Wesley Publishing Company.

Rossi, J., & Querrioux-Coulombier, G. (1997). “Picture Icon and Word Icon”. From *Human and Machine Perception*. New York, NY: Plenum Press.

Tullis, T. (1988). “Screen Design”. From *Handbook of Human-Computer Interaction*. North-Holland. Elsevier Science Publishers.

Yamakawa, J., Miller, N., & Huchingson, R. (1997). “An Evaluation of Icon Performance Based on User Preferences”. From *Human Factors in Information Systems: The Relationship Between User Interface Design and Human Performance*. Greenwich, CT: Ablex Publishing Corporation.

Richards, W. (Fall 1997). Class notes for 9.34 – Perceptual Information Processing, taken at MIT.