Subject name: Interactive Design

Subject code: SEG-14-13

Created By: Ali Akbary

Learning Outcome

Objectives of this chapter are: -

- Explain what cognition is and why it is important for interaction design.
- Discuss what attention is and its effects on our ability to multitask.
- Describe how memory can be enhanced through technology aids.
- Explain what mental models are.
- Show the difference between classic internal cognitive frameworks (e.g., mental models) and more recent external cognitive approaches (e.g., distributed cognition) that have been applied to HCI.
- Enable you to try to elicit a mental model and be able to understand what it means.

Chapter 3 Cognitive Aspects of Interactive Design

Introduction

Imagine it is late in the evening and you are sitting in front of your computer. You have an assignment to complete by tomorrow morning. A 3000-word essay on how natural are natural user interfaces. But you are not getting very far with it. You begin to panic and start biting your nails.

You see two text messages flash up on your smartphone. You instantly abandon your essay and cradle your smartphone to read them. One is from your mother and the other from your friend asking if you want to go out for a drink. You reply straight away to them both. Before you know it, you are back on Facebook to see if any of your friends have posted anything about the party you wanted to go to but had to say no.

FaceTime rings and you see it is your dad calling. You answer it and he asks if you have been watching the football game. You say you are too busy working toward your deadline and he tells you your team has just scored. You chat with him and then say you have to get back to work.

You realize 30 minutes has passed and you return your attention to the essay title. You type 'Natural User Interface' into Google Scholar and click on the top article. You click on the PDF icon for the article and it takes you to another page that requires a login and password. You don't have them for that publisher. You go back to Google Scholar and click on the next link. This time it takes you to the ACM digital library that your university has access to. But before you realize it you have clicked on the BBC Sports site to check the latest score for the football game. Your team has just scored again. Your phone starts buzzing. Two new WhatsApp messages are waiting for you. One is from your dad and another one from your girlfriend. You reply to both and within seconds they text back.

And on it goes. You glance at the time on your computer. It is 3.00 a.m. You really are in a panic now and finally switch everything off except your word processor.

In the past 10 to 15 years, it has become increasingly common for people to be always switching their attention between multiple tasks. At its most extreme form, such behavior has been found to be highly addictive: instead of focusing on our work, we are really waiting for the next hit – be it a new email, text, Facebook posting, news feed, tweet, and so forth. For some, such chronic media multitasking can be debilitating as they are unable to focus their attention on a single task for very long. For others, they have become very adept at using multiple sources of information to perform multiple tasks.

The study of human cognition can help us understand these and other new kinds of computer-augmented behaviors by examining humans' abilities and limitations when interacting with technologies.

In this chapter we examine cognitive aspects of interaction design. Specifically, we consider what humans are good and bad at and show how this knowledge can be used to inform the design of technologies that both extend human capabilities and compensate for their weaknesses. We also look at some of the influential cognitive-based conceptual frameworks that have been developed for explaining the way humans interact with technology.

WHAT IS COGNITION?

There are many different kinds of cognition, such as thinking, remembering, learning, daydreaming, decision making, seeing, reading, writing, and talking.

Norman (1993) distinguishes between two general modes: -

- Experiential cognition
- And reflective cognition.

Kahneman (2011) describes them in terms of fast and slow thinking.

- In experimental cognition, a state of mind in which we perceive, act, and react to events around us intuitively and effortlessly. It requires reaching a certain level of expertise and engagement. Examples include driving a car, reading a book, having a conversation, and playing a video game.
- In contrast, reflective cognition and slow thinking involve mental effort, attention, judgment, and decision making. This kind of cognition is what leads to new ideas and creativity. Examples include designing, learning, and writing a book. Both modes are essential for everyday life. It is useful to think of how the mind works in this way as it provides a basis from which to consider how each can be supported by different kinds of technologies.

Other ways of describing cognition are in terms of the context in which it takes place, the tools that are employed, the artifacts and interfaces that are used, and the people involved. Depending on when, where, and how it happens, cognition can be distributed, situated, extended, and embodied.

Cognition has also been described in terms of specific kinds of processes. These include: -

- Attention
- Perception
- Memory

- Learning
- Reading, speaking, and listening
- problem solving, planning, reasoning, and decision making.

Attention

This is the process of selecting things to concentrate on, at a point in time, from the range of possibilities available.

Attention involves our auditory and/or visual senses: -

- An example of auditory attention is waiting in the dentist's waiting room for our name to be called out to know when it is our time to go in.
- An example of visual attention is scanning the football results as they appear online via a live feed, checking to see whether our team is winning.

Attention allows us to focus on information that is relevant to what we are doing. The extent to which this process is easy or difficult depends on: -

i. Whether we have clear goals
ii. whether the information we need is salient in the environment.

Our Goals

If we know exactly what we want to find out, we try to match this with the information that is available.

- For example, if we have just landed at an airport after a long flight and want to find out who has won the World Cup, we might scan the headlines at the newspaper stand, find the results on our smartphone, call a friend, or ask someone in the street.
- When we are not sure exactly what we are looking for, we may browse through information, allowing it to guide our attention to interesting or salient items. For example, when we go to a restaurant, we may have the general goal of eating a meal but only a vague idea of what we want to eat. We peruse the menu to find things that whet our appetite, letting our attention be drawn to the imaginative descriptions of various dishes. After scanning through the possibilities and imagining what each dish might be like (plus taking into account other factors, such as cost, who we are with, what the specials are, what the waiter recommends, whether we want a two- or three-course meal, and so on), we may then make a decision.

Information Presentation

The way information is displayed can also greatly influence how easy or difficult it is to attend to appropriate pieces of information. Look at Figures below and try the activity (based on Tullis, 1997). Here, the information-searching tasks are very precise, requiring specific answers.


South Carolina

City	Motel/Hotel	Area code	Phone	Rates Single	Rates Double
Charleston	Best Western	803	747-0961	$126	$130
Charleston	Days Inn	803	881-1000	$118	$124
Charleston	Holiday Inn N	803	744-1621	$136	$146
Charleston	Holiday Inn SW	803	556-7100	$133	$147
Charleston	Howard Johnsons	803	524-4148	$131	$136
Charleston	Ramada Inn	803	774-8281	$133	$140
Charleston	Sheraton Inn	803	744-2401	$134	$142
Columbia	Best Western	803	796-9400	$129	$134
Columbia	Carolina Inn	803	799-8200	$142	$148
Columbia	Days Inn	803	736-0000	$123	$127
Columbia	Holiday Inn NW	803	794-9440	$132	$139
Columbia	Howard Johnsons	803	772-7200	$125	$127
Columbia	Quality Inn	803	772-0270	$134	$141
Columbia	Ramada Inn	803	796-2700	$136	$144
Columbia	Vagabond Inn	803	796-6240	$127	$130

Pennsylvania
Bedford Motel/Hotel: Crinaline Courts
(814) 623-9511 S: $118 D: $120
Bedford Motel/Hotel: Holiday Inn
(814) 623-9006 S: $129 D: $136
Bedford Motel/Hotel: Midway
(814) 623-8107 S: $121 D: $126
Bedford Motel/Hotel: Penn Manor
(814) 623-8177 S: $119 D: $125
Bedford Motel/Hotel: Quality Inn
(814) 623-5189 S: $123 D: $128
Bedford Motel/Hotel: Terrace
(814) 623-5111 S: $122 D: $124
Bradley Motel/Hotel: De Soto
(814) 362-3567 S: $120 D: $124
Bradley Motel/Hotel: Holiday House
(814) 362-4511 S: $122 D: $125
Bradley Motel/Hotel: Holiday Inn
(814) 362-4501 S: $132 D: $140
Breezewood Motel/Hotel: Best Western Plaza
(814) 735-4352 S: $120 D: $127
Breezewood Motel/Hotel: Motel 70
(814) 735-4385 S: $116 D: $118

Figure 1 - Two different ways of structuring the same information at the interface: one makes it much easier to find information than the other is not

Activity 3.1

Look at the top screen of Figure 1 and (i) find the price for a double room at the Quality Inn in Columbia, and (ii) find the phone number of the Days Inn in Charleston. Then look at the bottom screen in Figure 1 and (i) find the price of a double room at the Holiday Inn in Bradley, and (ii) find the phone number of the Quality Inn in Bedford. Which took longer to do?

In an early study, Tullis found that the two screens produced quite different results: it took an average of 3.2 seconds to search the top screen and 5.5 seconds to find the same kind of information in the bottom screen. Why is this so, considering that both displays have the same density of information (31%)?

Comment

The primary reason is the way the characters are grouped in the display. In the top screen, they are grouped into vertical categories of information – i.e., place, kind of accommodation, phone number, and rates – that have columns of space between them. In the bottom screen, the information is bunched up together, making it much harder to search through.

Multitasking and Attention

Many of us now spend a large proportion of our time staring at a screen, be it a smartphone, laptop, TV, or tablet. As mentioned in the introduction, while focusing on one task at a screen, we switch constantly between others. For example, every 5 or 10 minutes while writing this chapter, I check my email, breaking off sometimes in mid-sentence to see who has sent me a message and then finding myself diverted to looking at the latest news item or URL recommended to me by a colleague. Like nearly everyone else, I am addicted; I can't stop myself from looking.

But is it possible for us to perform multiple tasks without one or more of them being detrimentally affected? Consider the following. While attending a talk at a conference I watched a student volunteer in front of me deftly switch between four ongoing instant message chats (one at the conference, one at school, one with friends, one at her part-time job), read, answer, delete, and place all new messages in various folders of her two email accounts, check and scan Facebook and her Twitter feeds – while appearing to listen to the talk, take some notes, Google the speaker's background, and open up his publications. When she had a spare moment, she played a game of patience. I must say, I felt quite exhausted just watching her for a few minutes. It was as if she were capable of living in multiple worlds, all at the same time, while not letting a moment go to waste. But how much did she take in of the talk?

There has been much research on the effects of multitasking on memory and attention. A main finding is that it depends on the nature of the tasks and how much attention each demands. For example, listening to gentle music while working can help people tune out background noise, such as traffic or other people talking, and help them concentrate on what they are doing. However, if the music is loud, like Drum and Bass, it can be very distracting. Individual differences have also been found. For example, the results of a series of experiments comparing heavy with light multitaskers showed that heavy media multitaskers (such as the one described above) were more prone to being distracted by the multiple streams of media they are looking at than those who infrequently multitask. The latter were found to be better at allocating their attention when faced with competing distractions (Ophir *et al*, 2009). This suggests that people who are heavy multitaskers are likely to be those who are easily distracted and find it difficult to filter out irrelevant information.

Perception

Perception refers to how information is acquired from the environment via the different sense organs – eyes, ears, fingers – and transformed into experiences of objects, events, sounds, and tastes (Roth, 1986). It is complex, involving other cognitive processes such as memory, attention, and language. Vision is the most dominant sense for sighted individuals, followed by hearing and touch. With respect to interaction design it is important to present information in a way that can be readily perceived in the manner intended.

As was demonstrated in Activity 3.1, grouping items together and leaving spaces between them can aid attention. In addition, many web designers recommend using blank space (more commonly known as white space) when grouping objects together on a screen as it helps users to perceive and locate items more easily and quickly. However, some researchers suggest that too much white space can be detrimental, making it sometimes harder to find information. In a study comparing web pages displaying the same amount of information, but which were structured using different graphical methods, it was found that people took less time to locate items from information that was grouped using a border than when using color contrast. The findings suggest that using contrasting color is not a good way to group information on a screen and that using borders is more effective.

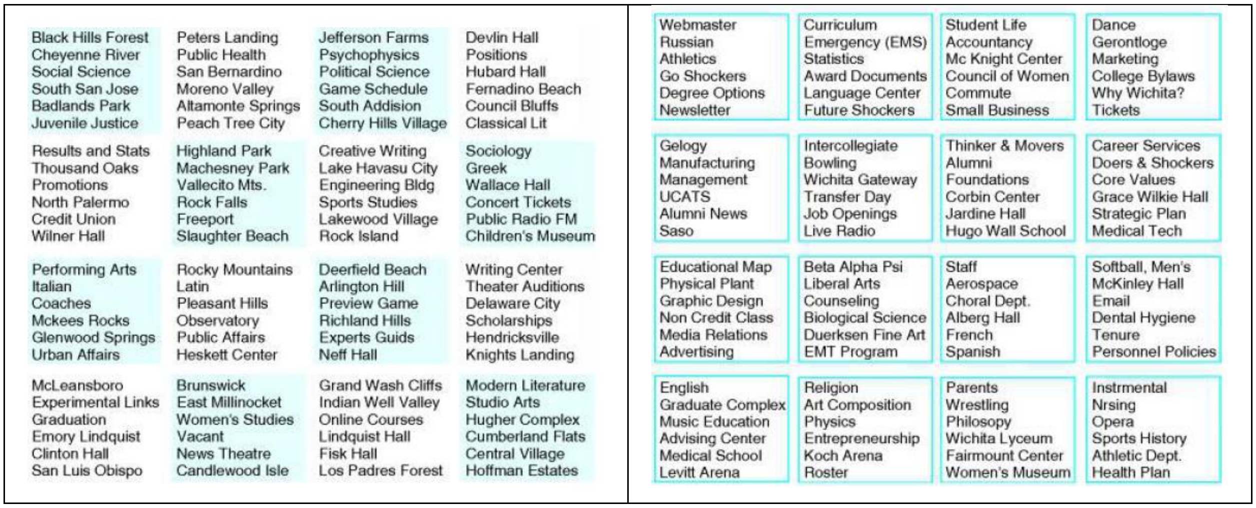

Figure 2 - Two ways of structuring information on a web page. It takes more time for people to find a named item in the top one than in the bottom one, suggesting that using bordering as a grouping method helps searching while using contrasting color hinders it

Design Implications - Perception

Representations of information need to be designed to be perceptible and recognizable across different media: -

- Icons and other graphical representations should enable users to readily distinguish their meaning.
- Bordering and spacing are effective visual ways of grouping information that makes it easier to perceive and locate items.
- Sounds should be audible and distinguishable so users understand what they represent.
- Speech output should enable users to distinguish between the set of spoken words and also be able to understand their meaning.
- Text should be legible and distinguishable from the background (e.g., it is okay to use yellow text on a black or blue background but not on a white or green background).

- Tactile feedback used in virtual environments should allow users to recognize the meaning of the various touch sensations being emulated. The feedback should be distinguishable so that, for example, the sensation of squeezing is represented in a tactile form that is different from the sensation of pushing.

Memory

Memory involves recalling various kinds of knowledge that allow us to act appropriately. It is very versatile, enabling us to do many things. For example, it allows us to recognize someone's face, remember someone's name, recall when we last met them, and know what we said to them last.

It is not possible for us to remember everything that we see, hear, taste, smell, or touch, nor would we want to, as our brains would get completely overloaded. A filtering process is used to decide what information gets further processed and memorized. This filtering process, however, is not without its problems. Often, we forget things we would dearly love to remember and conversely remember things we would love to forget. For example, we may find it difficult to remember everyday things like people's names and phone numbers, or scientific knowledge such as mathematical formulae. On the other hand, we may effortlessly remember trivia or tunes that cycle endlessly through our heads.

How does this filtering process work? Initially, encoding takes place, determining which information is attended to in the environment and how it is interpreted. The extent to which it takes place affects our ability to recall that information later. The more attention that is paid to something and the more it is processed in terms of thinking about it and comparing it with another knowledge, the more likely it is to be remembered. For example, when learning about a topic it is much better to reflect upon it, carry out exercises, have discussions with others about it, and write notes than just passively read a book or watch a video about it. Thus, how information is interpreted when it is encountered greatly affects how it is represented in memory and how easy it is to retrieve subsequently.

Another factor that affects the extent to which information can be subsequently retrieved is the context in which it is encoded. One outcome is that sometimes it can be difficult for people to recall information that was encoded in a different context from the one they are currently in. Consider the following scenario: -

> *"You are on a train and someone comes up to you and says hello. You don't recognize him for a few moments but then realize it is one of your neighbors. You are only used to seeing your neighbor in the hallway of your apartment block and seeing him out of context makes him difficult to recognize initially."*

Another well-known memory phenomenon is that people are much better at recognizing things than recalling things. Furthermore, certain kinds of information are easier to recognize than others. In particular, people are very good at recognizing thousands of pictures even if they have only seen them briefly before. In contrast, we are not as good at remembering details about the things we take photos of when visiting places, such as museums. It seems we remember less about objects when we have photographed them than when we observe them just with the naked eye. The reason for this difference in our ability to remember details about objects is that people don't process as much information about an object when taking photos of it compared with when they are actually looking at it – and hence are not able to remember as much about it later.

Activity 3.2

Try to remember the dates of all the members of your family's and your closest friends' birthdays. How many can you remember? Then try to describe the image/graphic of the latest app you downloaded.

Comment

It is likely that you remembered much better the image, the colors, and the name of the app you downloaded than the birthdays of your family and friends (which most people now rely on Facebook or other online app to remind them of). People are very good at remembering visual cues about things, for example the color of items, the location of objects (e.g., a book being on the top shelf), and marks on an object (e.g., a scratch on a watch, a chip on a cup). In contrast, people find other kinds of information persistently difficult to learn and remember, especially arbitrary material like birthdays and phone numbers.

Memory and Search

The number of documents created, images, music files, and videoclips downloaded, emails and attachments saved, URLs bookmarked, and so on increases every day. Increasingly, people are saving their digital content to the Cloud so that it can be accessed from multiple platforms, but it still needs to be organized in a way that can be easily searched. For example, do they place items in folders or albums or lists? Many people use proprietary storage facilities, such as iCloud, Vimeo, Pinterest, and Flickr, to save their content. A challenge facing these companies is providing interfaces that will enable their users to store their content so they can readily access specific items at a later date, for example a particular image, video, or document. This can be challenging, especially if they have uploaded thousands of them. How do you find that photo you took of your dog spectacularly jumping into the sea to chase a seagull, which you think was taken two or three years ago? It can take ages wading through the hundreds of folders you have, catalogued by date, name, or tag. Do you start by homing in on folders for a given year, look for events, places, or faces, or type in a search term to find it again?

Naming is the most common means of encoding content, but trying to remember a name you created some time back can be very difficult, especially if you have tens of thousands of named files, images, videos, emails, etc. How might such a process be facilitated, taking into account people's memory abilities? Lansdale and Edmonds (1992) suggest that it is profitable to view this kind of remembering as involving two memory processes: -

- **Recall-directed** - This refers to using memorized information about the required content to get as close to it as possible. The more exact this is, the more success the user will have in tracking down the desired content.
- **Recognition-based scanning** - This happens when recall has failed to produce what a user wants and so requires reading through a list.

To illustrate the difference between these two processes, consider the following scenario: -

A user is trying to access a couple of websites she visited the week before that compared the selling price of cars offered by different dealers. The user is able to recall the name of one website, **autobargains.com**. She types this in her web browser and the website appears.

This is an example of successful recall-directed memory. However, the user is unable to remember the name of the second one. She vaguely remembers it was something like **alwaysthecheapest.com**, but typing this in proves unsuccessful. Instead, she switches to scanning her history list and selects the folder labeled more than six days ago. She notices two or three URLs that could be the one desired at the top of the list, and on the second attempt she finds the website she is looking for. In this situation, the user initially tries recall-directed memory and when this fails adopts the second strategy of recognition-based scanning – which takes longer but eventually results in success.

Digital content systems should be designed to optimize both kinds of memory processes. In particular, they should be designed to let people use whatever memory they have to limit the area being searched and then represent the information in this area of the interface so as to maximally assist them in finding what they need. The system should provide the user with a number of ways of encoding documents mnemonically, including time stamping, categorizing, tagging, and attribution (e.g., color, text, icon, sound, or image).

Powerful search engines have gone a long way towards helping people track down the content they want. For example, various search and find tools, such as Android's Bravo SE and Apple's Spotlight, enable the user to type a full or partial name or even the first letter of a file that it then searches for in the entire system, including apps, games, emails, contacts, images, calendars, and applications. Figure 3 shows part of a list of files that Spotlight matched to the phrase 'cognition', prioritized in terms of what I might be looking for, such as documents, web pages, and emails. The categories change depending on the words entered. For example, if someone's name is entered then images, contacts, and websites are prioritized.

Figure 3 - Apple's Spotlight search tool

BOX 3.1 - The Problem with the Magical Number Seven, Plus or Minus Two

Perhaps the best-known finding in psychology (certainly the one that nearly all students remember many years after they have finished their studies) is Miller's (1956) theory that seven, plus or minus two, chunks of information can be held in short-term memory at any one time. By short-term memory he meant a memory store in which information was assumed to be processed when first perceived. By chunks he meant a range of items like numbers, letters, or words. According to Miller's theory, therefore, people's immediate memory capacity is very limited. They are able to remember only a few words or numbers that they have heard or seen. If you are not familiar with this phenomenon, try out the following exercise: read the first set of numbers below (or get someone to read it to you), cover it up, and then try to recall as many of the items as possible. Repeat this for the other sets.

- 3, 12, 6, 20, 9, 4, 0, 1, 19, 8, 97, 13, 84
- cat, house, paper, laugh, people, red, yes, number, shadow, broom, rain, plant, lamp, chocolate, radio, one, coin, jet
- t, k, s, y, r, q, x, p, a, z, l, b, m, e

How many did you correctly remember for each set? Between five and nine, as suggested by Miller's theory?

Chunks can also be combined items that are meaningful. For example, it is possible to remember the same number of two-word phrases like hot chocolate, banana split, cream cracker, rock music, cheddar cheese, leather belt, laser printer, tree fern, fluffy duckling, cold rain. When these are all muddled up (i.e., split belt, fern crackers, banana laser, printer cream, cheddar tree, rain duckling, hot rock), however, it is much harder to remember as many chunks. This is mainly because the first set contains all meaningful two-word phrases that have been heard before and require less time to be processed in short-term memory, whereas the second set are completely novel phrases that don't exist in the real world. You need to spend time linking the two parts of the phrase together while trying to memorize them. This takes more time and effort to achieve. Of course, it is possible to do if you have time to spend rehearsing them, but if you are asked to do it having heard them only once in quick succession, it is most likely you will remember only a few.

By now, you may be thinking ‘Okay, this is interesting, but what has it got to do with interaction design?’ Well, not only does this classic theory have a special place in psychology, it has also made a big impression in Human-computer interaction (HCI). Unfortunately, however, for the wrong reasons. Many designers have heard or read about this phenomenon and think, ‘Ah, here is a bit of psychology I can usefully apply to interface design.’ Would you agree with them? If so, how might people's ability to only remember 7 ± 2 chunks that they have just read or heard be usefully applied to interaction design?

According to a survey by Bailey (2000), several designers have been led to believe the following guidelines and have even created interfaces based on them: -

- Have only seven options on a menu.
- Display only seven icons on a menu bar.

- Never have more than seven bullets in a list.
- Place only seven tabs at the top of a website page.
- Place only seven items on a pull-down menu.

All of these are wrong. Why? The simple reason is that these are all items that can be scanned and rescanned visually and hence do *not* have to be recalled from short-term memory. They don't just flash up on the screen and disappear, requiring the user to remember them before deciding which one to select. If you were asked to find an item of food most people crave in the set of single words listed above, would you have any problem? No, you would just scan the list until you recognized the one (chocolate) that matched the task and then select it – just as people do when interacting with menus, lists, and tabs – regardless of whether they comprise three or 30 items. What the users are required to do here is not remember as many items as possible, having only heard or seen them once in a sequence, but instead scan through a set of items until they recognize the one, they want. This is a quite different task. Furthermore, there is much more useful psychological research that can be profitably applied to interaction design.

Memory Load and Passwords

Phone banking has become increasingly popular in the past few years. It allows customers to carry out financial transactions, such as paying bills and checking the balance of their accounts, at their convenience. One of the problems confronting banks that provide this facility, however, is how to manage security concerns. Anyone can phone up a bank and pretend to be someone else. How do the banks prevent fraudulent transactions?

One solution has been to develop rigorous security measures whereby customers must provide various pieces of information before gaining access to their accounts. Typically, these include providing the answers to a combination of the following: -

- Their zip code or post code
- Their mother's maiden name their birthplace
- The last school they attended the first school they attended
- A password of between five and ten letters a memorable address (not their home)
- A memorable date (not their birthday).

Many of these are relatively easy to remember and recall as they are very familiar. But consider the last two. How easy is it for someone to come up with such memorable information and then be able to recall it readily? Perhaps the customer can give the address and birthday of another member of their family as a memorable address and date. But what about the request for a password? Suppose a customer selects the word ‘interaction’ as a password – fairly easy to remember. The problem is that the bank operators do not ask for the full password, because of the danger that someone in the vicinity might overhear and write it down. Instead, they are instructed to ask the customer to provide specific letters from it, like the seventh followed by the fifth. However, such information does not spring readily to mind. Instead, it requires mentally counting each letter of the password until the desired one is reached. How long does it take you to determine the seventh letter of the password ‘interaction’? How did you do it?

To make things harder, banks also randomize the questions they ask. Again, this is to prevent someone who might be overhearing from memorizing the sequence of information. However, it also means that the customers themselves cannot learn the sequence of information required, meaning they have to generate different information every time they call up the bank.

This requirement to remember and recall such information puts a big memory load on customers. Some people find such a procedure quite nerve-racking and are prone to forget certain pieces of information. As a coping strategy they write down their details on a sheet of paper. Having such an external representation at hand makes it much easier for them to read off the necessary information rather than having to recall it from memory. However, it also makes them vulnerable to the very fraud the banks were trying to prevent, should anyone else get hold of that piece of paper!

Activity 3.3

How else might banks solve the problem of providing a secure system while making the memory load relatively easy for people wanting to use phone banking? How does phone banking compare with online banking?

Comment

An alternative approach is to provide the customers with a PIN and ask them to key this in on their phone keypad, followed by asking one or two questions like their zip or post code, as a backup. Online banking has similar security risks to phone banking, and hence this requires a number of security measures to be enforced. These include that the user sets up a nickname and a password. For example, some banks require answering a question only the user knows the answer to and then typing in three randomly selected letters from a password each time the user logs on. This is harder to do online than when asked over the phone, mainly because it interferes with the normally highly automated process of typing in a password. You really have to think about what letters and numbers are in your password – for example, has it got two letters ‘f's after the number ‘6’ or just one?

Researchers have also investigated whether images could be used instead of alphanumeric for passwords. The idea is based on the principle that recognition is better than recall: users should be able to remember their passwords more accurately if they are required to recognize a set of images from a display that makes up their password than if they have to recall a sequence of alphanumeric. To this end, the graphical authentication approach has been developed, which asks people to select a series of images from different matrices of options. The images can be faces, cartoons, or photos of scenes or objects (e.g., sunset, dog, or even abstract images). To enable the process to be secure, however, requires people selecting a sequence of four to eight images and subsequently being able to recognize each item in the correct sequence. In other words, both recall (of the sequence) and recognition are involved. Studies have shown that while the graphical approach appears an attractive alternative, it has yet to demonstrate convincingly an advantage over the use of alphanumeric. Moreover, it takes much longer to create and subsequently select a sequence of images each time a person logs on than typing in a set of letters and numbers at a keyboard (De Angeli *et al*, 2005).

BOX 3.2 - Digital Forgetting

Much of the research on memory and interaction design has focused on developing cognitive aids that help people to remember; for example, reminders, to-do lists, and digital photo collections. However, there are times when we wish to forget a memory. For example, when someone breaks up with their partner, it can be emotionally painful to be reminded of them through shared digital images, videos, and Facebook friends.

- How can technology be designed to help people forget such memories?
- How could social media, such as Facebook, be designed to support this process?

Sas and Whittaker (2013) suggest designing new ways of harvesting digital materials connected to a broken relationship through using various automatic methods, such as face recognition, that dispose of them without the person needing to personally go through them and be confronted with painful memories. They also suggest that during a separation, people could create a collage of their digital content connected to the ex, so as to transform them into something more abstract, thereby providing a means for closure and helping with the process of moving on.

Computing Aids for Memory Loss

People suffering from memory impairments can find it difficult to complete common household tasks, like cooking and washing up, because they may forget a step or where they were. This can be exacerbated if the person gets interrupted (e.g., the phone rings), and they may end up not including an ingredient or adding the washing-up liquid twice. A prototype system called Cook's Collage was designed to provide surrogate memory support for general cooking tasks (Tran *et al*, 2005). Cameras were mounted underneath cabinets to capture still images of a cooking activity. These were then displayed as a series of images, in the form of a cartoon strip, on a flat-panel display mounted on an eye-level kitchen cabinet (see Figure 4). Preliminary evaluation of the prototype, being used by old people while cooking, showed them using it mainly as an aide-memoire, checking to see whether they had added certain ingredients after being distracted from the cooking task at hand.

Figure 4 - A screenshot of Cook's Collage showing images of a recent cooking activity. The strip is designed to be read backwards, starting with the highlighted image. This shows to the cook that he previously added the 29th scoop (!) of sugar and in the previous image two scoops of soda water

Another computing technology that was used to help people suffering from memory loss (e.g., those with Alzheimer's disease) was the Sense Cam, which was originally developed by Microsoft Research Labs in Cambridge (UK) to enable people to remember everyday events.

This is a wearable camera (the predecessor of Autographer) that intermittently takes photos, without any user intervention, while it is being worn (see Figure 5). The camera can be set to take pictures at particular times; for example, every 30 seconds, or based on what it senses (e.g., acceleration). The camera's lens is fish-eyed, enabling nearly everything in front of the wearer to be captured. The digital images for each day are stored, providing a record of the events that a person experiences. Several studies have been conducted on patients with various forms of memory loss using the device. For example, Hodges *et al* (2006) describe how a patient, Mrs. B, who had amnesia was given a SenseCam to wear. The images that were collected were uploaded to a computer at the end of each day. For the next two weeks, Mrs. B and her husband looked through these and talked about them. During this period, Mrs. B's recall of an event nearly tripled, to a point where she could remember nearly everything about that event. Prior to using the SenseCam, Mrs. B would have typically forgotten the little that she could initially remember about an event within a few days. It is not surprising that she did not want to return the device.

Figure 5 - The SenseCam device and a digital image taken with it

Learning

It is well known that people find it hard to learn by following a set of instructions in a manual. Instead, they much prefer to learn through doing. Graphical User Interface (GUIs) and direct manipulation interfaces are good environments for supporting this kind of active learning by supporting exploratory interaction and, importantly, allowing users to undo their actions, i.e., return to a previous state if they make a mistake by clicking on the wrong option.

There have been numerous attempts to harness the capabilities of different technologies to help learners understand topics. One of the main benefits of interactive technologies, such as web-based learning, eLearning, multimedia, and virtual reality, is that they provide alternative ways of representing and interacting with information that are not possible with traditional technologies, e.g., books. In so doing, they have the potential of offering learners the ability to explore ideas and concepts in different ways. For example, interactive multimedia simulations have been designed to help teach abstract concepts (e.g., mathematical formulae, notations, laws of physics) that students find difficult to grasp. Different representations of the same process (e.g., a graph, a formula, a sound, a simulation) are displayed and interacted with in ways that make their relationship with each other more explicit to the learner.

One form of interactivity that has been found to be highly effective is dynalinking (Rogers and Scaife, 1998). Abstract representations, such as diagrams, are linked together with a more concrete illustration of what they stand for, such as a simulation. Changes in one are matched by changes in the other, enabling a better understanding of what the abstraction means. An early example of its use was software developed for learning about ecological concepts, such as food webs (Rogers *et al*, 2003). A concrete simulation showed various organisms swimming and moving around and occasionally an event where one would eat another (e.g., a snail eating the weed). This was annotated and accompanied by various eating sounds, like chomping, to attract the children's attention. The children could also interact with the simulation. When an organism was clicked on, it would say what it was and what it ate (e.g., 'I am a weed. I make my own food'). The concrete simulation was dynalinked with other abstract representations of the pond ecosystem, including an abstract food web diagram (see Figure 3.6).

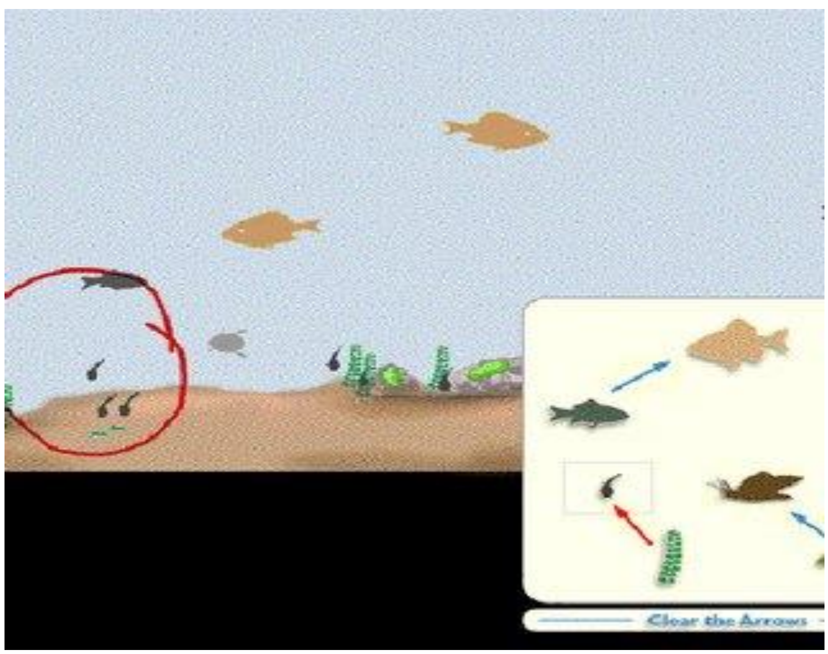

Figure 6 - Dynalinking used in the Pondworld software

Dynalinking has been used in other domains to explicitly show relationships among multiple dimensions where the information to be understood or learned is complex (Sutcliffe, 2002). For example, it can be useful for domains like economic forecasting, molecular modeling, and statistical analyses.

Increasingly, we rely on the Internet and our smartphones to act as cognitive prostheses in the way in which blind people use walking sticks. They have become a cognitive resource that we use in our daily lives as part of the extended mind. Sparrow *et al* (2011) showed how expecting to have Internet access reduces the need and hence the extent to which we attempt to remember the information itself, while enhancing our memory for knowing where to find it online. Many of us will whip out our smartphone to find out who acted in a film, what the name of a book is, what the word in another language is, and so on. Besides search engines, there are a number of other cognitive prosthetic apps that instantly help us find out or remember something, such as Shazam, the popular music recognition app. This has important implications for the design of technologies to support *how* future generations will learn, and *what* they learn.

Design Implications - Learning

- Design interfaces that encourage exploration.
- Design interfaces that constrain and guide users to select appropriate actions when initially learning.
- Dynamically link concrete representations and abstract concepts to facilitate the learning of complex material

Reading, Speaking, and Listening

Reading, speaking, and listening are three forms of language processing that have similar and different properties. One similarity is that the meaning of sentences or phrases is the same regardless of the mode in which it is conveyed. For example, the sentence 'Computers are a wonderful invention' essentially has the same meaning whether one reads it, speaks it, or hears it. However, the ease with which people can read, listen, or speak differs depending on the person, task, and context. For example, many people find listening easier than reading. Specific differences between the three modes include: -

- Written language is permanent while listening is transient. It is possible to re-read information if not understood the first time around. This is not possible with spoken information that is being broadcast.
- Reading can be quicker than speaking or listening, as written text can be rapidly scanned in ways not possible when listening to serially presented spoken words.
- Listening requires less cognitive effort than reading or speaking. Children, especially, often prefer to listen to narratives provided in multimedia or web-based learning material than to read the equivalent text online.
- Written language tends to be grammatical while spoken language is often ungrammatical. For example, people often start talking and stop in mid-sentence, letting someone else start speaking.
- Dyslexics have difficulties understanding and recognizing written words, making it hard for them to write grammatical sentences and spell correctly.

Many applications have been developed either to capitalize on people's reading, writing, and listening skills, or to support or replace them where they lack or have difficulty with them. These include: -

- Interactive books and web-based materials that help people to read or learn foreign languages.
- Speech-recognition systems that allow users to interact with them by using spoken commands (e.g., word-processing dictation, Google Voice Search app, and home control devices that respond to vocalized requests).
- Speech-output systems that use artificially generated speech (e.g., written-text-to-speech systems for the blind).
- Natural-language systems that enable users to type in questions and give text-based responses (e.g., the Ask search engine).
- Cognitive aids that help people who find it difficult to read, write, and speak. Numerous special interfaces have been developed for people who have problems with reading, writing, and speaking (e.g., see Edwards, 1992).
- Customized input and output devices that allow people with various disabilities to have access to the web and use word processors and other software packages.
- Interaction techniques that allow blind people to read graphs and other visuals on the web through the use of auditory navigation and tactile diagrams (Petrie et al, 2002).

Design Implications - Reading, Speaking, and Listening

- Keep the length of speech-based menus and instructions to a minimum. Research has shown that people find it hard to follow spoken menus with more than three or four options. Likewise, they are bad at remembering sets of instructions and directions that have more than a few parts.
- Accentuate the intonation of artificially generated speech voices, as they are harder to understand than human voices.
- Provide opportunities for making text large on a screen, without affecting the formatting, for people who find it hard to read small text.

Problem Solving, Planning, Reasoning, and Decision Making

Problem solving, planning, reasoning, and decision making are processes involving reflective cognition. They include thinking about what to do, what the options are, and what the consequences might be of carrying out a given action. They often involve conscious processes (being aware of what one is thinking about), discussion with others (or oneself), and the use of various kinds of artifacts (e.g., maps, books, pen and paper).

For example, when planning the best route to get somewhere, say a foreign city: -

- We may ask others
- Use a paper map
- Get directions from the web
- Or use a combination of these

Reasoning involves working through different scenarios and deciding which is the best option or solution to a given problem. In the route-planning activity we may be aware of alternative routes and reason through the advantages and disadvantages of each route before deciding on the best one. Many a family argument has come about because one member thinks he knows the best route while another thinks otherwise. Nowadays, many of us offload this kind of decision making (and the stress) onto technology, by simply following the instructions given by a car GPS or a smartphone map app. According to an internal survey carried out by YouGov in March 2014 in the UK, TomTom – which launched the first SatNav in 2004 – has helped 13 million couples avoid navigation arguments in the car!

There has been a growing interest in how people make decisions when confronted with information overload, such as when shopping on the web or at a store. How easy is it to make a decision when confronted with overwhelming choice? Classical rational theories of decision making (e.g., von Neumann and Morgenstern, 1944) posit that making a choice involves weighing up the costs and benefits of different courses of action. This is assumed to involve exhaustively processing the information and making trade-offs between features. Such strategies are very costly in computational and informational terms – not least because they require the decision-maker to find a way of comparing the different options. In contrast, research in cognitive psychology has shown how people tend to use simple heuristics when making decisions (Gigerenzer *et al*, 1999). Atheoretical explanation is that human minds have evolved to act quickly, making just good enough decisions by using fast and frugal heuristics. We typically ignore most of the available information and rely only on a few important cues. For example, in the supermarket, shoppers make snap judgments based on a paucity of information, such as buying brands that they recognize, that are low-priced, or have attractive packaging – seldom reading other package information. This suggests that an effective design strategy is to follow the adage 'less is more' rather than 'more is more,' making key information about a product highly salient.

Thus, instead of providing ever more information to enable people to compare products when making a choice, a better strategy is to design technological interventions that provide just enough information, and in the right form, to facilitate good choices. One solution is to exploit new forms of augmented reality and wearable technology that enable information-frugal decision making and which have glanceable displays that can represent key information in an easy-to-digest form (Rogers, Payne and Todd, 2010).

Dilemma - Can You Make up Your Mind Without an App?

Howard Gardner and Katie Davis (2013) in their book The App Generation note how the app mentality developing in the psyche of the younger generation is making it worse for them to make their own decisions because they are becoming more risk averse. By this they mean that young people are now depending on an increasing number of mobile apps that remove the risks of having to decide for themselves. They will first read what others have said on social media sites, blogs, and recommender apps before choosing where to eat, where to go, what to do, what to listen to, etc. But, relying on a multitude of apps means that young people are becoming increasingly more anxious about making decisions by themselves. For many, their first big decision is choosing which university to go to. This has become an agonizing and prolonged experience where both parents and apps play a central role in helping them out. They will read countless reviews, go on numerous visits to universities with their parents over several months, study the form of a number of league tables, read up on what others say on social networking sites, and so on. But in the end, was all that necessary? They may finally end up choosing where their friends are going or the one, they liked the look of in the first place. Many will have spent hours, weeks, and even months talking about it, reading up on it, listening to lots of advice, and procrastinating right down to the wire. Compared to previous pre-Internet generations, they won't have made the decision by themselves.

Design Implications - Problem Solving, Planning, Reasoning, and Decision Making

- Provide additional hidden information that is easy to access for users who wish to understand more about how to carry out an activity more effectively (e.g., web searching).
- Use simple and memorable functions at the interface for computational aids intended to support rapid decision making and planning that takes place while on the move.

COGNITIVE FRAMEWORKS

A number of conceptual frameworks and theories have been developed to explain and predict user behavior based on theories of cognition. In this section, we outline three early internal frameworks that focus primarily on mental processes together with three more recent external ones that explain how humans interact and use technologies in the context in which they occur. These are: -

Internal

- Mental models
- Gulfs of execution and evaluation
- Information processing

External

- Distributed cognition
- External cognition
- Embodied interaction

Internal

Mental Models

In Chapter 2 we pointed out that a successful system is one based on a conceptual model that enables users to readily learn that system and use it effectively. People primarily develop knowledge of how to interact with a system and, to a lesser extent, how that system works. In the 1980s and 1990s, these two kinds of knowledge were often referred to as a user's mental model.

It is assumed that mental models are used by people to reason about a system and, in particular, to try to fathom out what to do when something unexpected happens with the system or when encountering unfamiliar systems. The more someone learns about a system and how it functions, the more their mental model develops. For example, TV engineers have a deep mental model of how TVs work that allows them to work out how to set them up and fix them. In contrast, an average citizen is likely to have a reasonably good mental model of how to operate a TV but a shallow mental model of how it works.

Within cognitive psychology, mental models have been postulated as internal constructions of some aspect of the external world that are manipulated, enabling predictions and inferences to be made (Craik, 1943). This process is thought to involve the fleshing out and the running of a mental model (Johnson-Laird, 1983). This can involve both unconscious and conscious mental processes, where images and analogies are activated.

Activity 3.4

To illustrate how we use mental models in our everyday reasoning, imagine the following two scenarios: -

1. You arrive home from a holiday on a cold winter's night to a cold house. You have a small baby and you need to get the house warm as quickly as possible.

Your house is centrally heated. Do you set the thermostat as high as possible or turn it to the desired temperature (e.g., 70°F)?

2. You arrive home after being out all night and you're starving hungry. You look in the freezer and find all that is left is a frozen pizza. The instructions on the packet say heat the oven to 375°F and then place the pizza in the oven for 20 minutes. Your oven is electric. How do you heat it up? Do you turn it to the specified temperature or higher?

Comment

- Most people when asked the first question imagine the scenario in terms of what they would do in their own house and choose the first option. A typical explanation is that setting the temperature to be as high as possible increases the rate at which the room warms up. While many people may believe this, it is incorrect. Thermostats work by switching on the heat and keeping it going at a constant speed until the desired set temperature is reached, at which point it cuts out. They cannot control the rate at which heat is given out from a heating system. Left at a given setting, thermostats will turn the heat on and off as necessary to maintain the desired temperature.
- When asked the second question, most people say they would turn the oven to the specified temperature and put the pizza in when they think it is at the right temperature. Some people answer that they would turn the oven to a higher temperature in order to warm it up more quickly. Electric ovens work on the same principle as central heating, and so turning the heat up higher will not warm it up any quicker. There is also the problem of the pizza burning if the oven is too hot!

Why do people use erroneous mental models? It seems that in the above scenarios, they are running a mental model based on a general valve theory of the way something works (Kempton, 1986). This assumes the underlying principle of more is more: the more you turn or push something, the more it causes the desired effect. This principle holds for a range of physical devices, such as faucets and radio controls, where the more you turn them, the more water or volume comes out. However, it does not hold for thermostats, which instead function based on the principle of an on–off switch. What seems to happen is that in everyday life, people develop a core set of abstractions about how things work, and apply these to a range of devices, irrespective of whether they are appropriate.

Using incorrect mental models to guide behavior is surprisingly common. Just watch people at a pedestrian crossing or waiting for an elevator. How many times do they press the button? A lot of people will press it at least twice. When asked why, a common reason given is that they think it will make the lights change faster or ensure the elevator arrives. This seems to be another example of following the ‘more is more’ philosophy: it is believed that the more times you press the button, the more likely it is to result in the desired effect.

Many people's understandings of how technologies and services (e.g., the Internet, wireless networking, broadband, search engines, and computer viruses) work is poor. Their mental models are often incomplete, easily confusable, and based on inappropriate analogies and superstition (Norman, 1983). As a consequence, they find it difficult to identify, describe, or solve a problem, and lack the words or concepts to explain what is happening.

If people could develop better mental models of interactive systems, they would be in a better position to know how to carry out their tasks efficiently, and know what to do if a system started malfunctioning. Ideally, they should be able to develop a mental model that matches the conceptual model. But to what extent is this realistic, given that most people are resistant to spending much time learning about how things work, especially if it involves reading manuals or other documentation? Alternatively, if interactive technologies could be designed to be more transparent, then it might be easier to understand them in terms of how they work and what to do when they don't. Transparency involves including: -

- Useful feedback in response to user input
- Easy-to-understand and intuitive ways of interacting with the system
- In addition, it requires providing the right kind and level of information, in the form of: clear and easy-to-follow instructions
- Appropriate online help and tutorials
- context-sensitive guidance for users, set at their level of experience, explaining how to proceed when they are not sure what to do at a given stage of a task

Dilemma - How Much Transparency?

How much and what kind of transparency do you think a designer should provide in an interactive product? This is not a straightforward question to answer and depends a lot on the requirements of the targeted user groups. Some users simply want to get on with their tasks and don't want to have to learn about how the thing they are using works. In this situation, the interface should be designed to make it obvious what to do and how to use it. Functions that are difficult to learn can be off-putting. Users simply won't bother to make the extra effort, meaning that many of the functions provided are never used. Other users like to understand how the device they are using works in order to make informed decisions about how to carry out their tasks, especially if there are numerous ways of doing something. Some search engines have been designed with this in mind: *"they provide background information on how they work and how to improve one's searching techniques"*

Gulfs of Execution and Evaluation

The gulf of execution and the gulf of evaluation describe the gaps that exist between the user and the interface (Norman, 1986; Hutchins *et al*, 1986). They are intended to show how to design the latter to enable the user to cope with them. The first one – the gulf of execution – describes the distance from the user to the physical system while the second one – the gulf of evaluation – is the distance from the physical system to the user (see Figure 7). Norman and his colleagues suggest that designers and users need to concern themselves with how to bridge the gulfs in order to reduce the cognitive effort required to perform a task. This can be achieved, on the one hand, by designing usable interfaces that match the psychological characteristics of the user (e.g., taking into account their memory limitations) and, on the other hand, by the user learning to create goals, plans, and action sequences that fit with how the interface works.

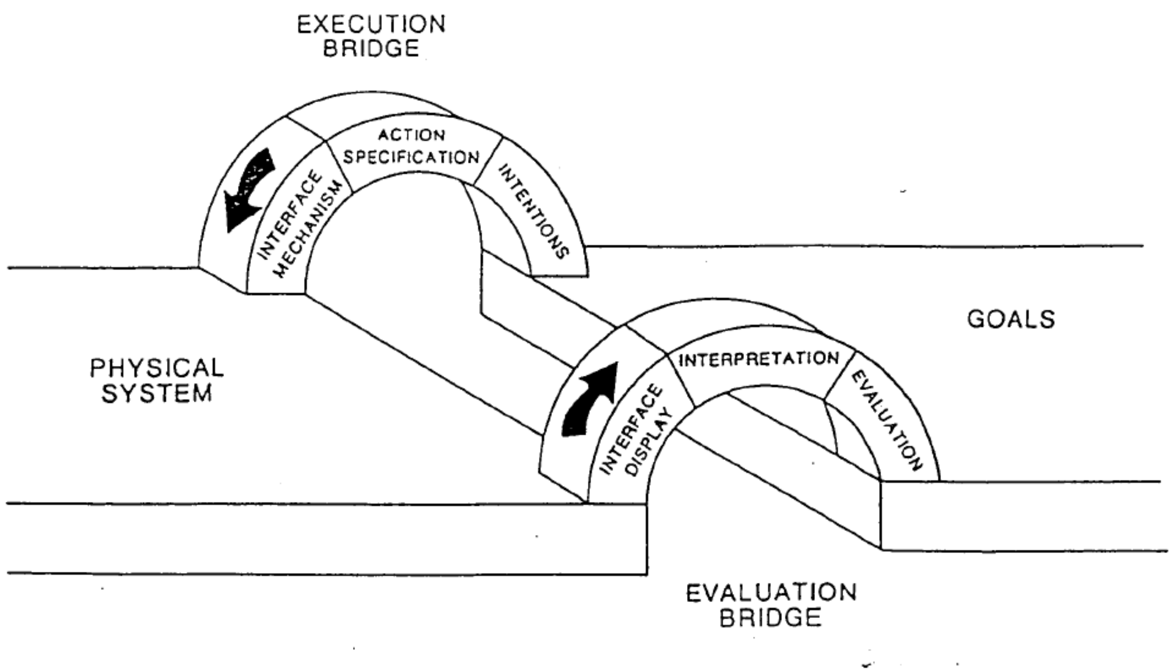

Figure 7 - Bridging the gulfs of execution and evaluation

Source: User centered system design: new perspectives on human-computer interaction by D Norman. Copyright 1986 by Taylor & Francis Group LLC - Books. Reproduced with permission of Taylor & Francis Group LLC.

Information Processing

Another classic approach to conceptualizing how the mind works has been to use metaphors and analogies. Numerous comparisons have been made, including conceptualizing the mind as a reservoir, a telephone network, and a digital computer. One prevalent metaphor from cognitive psychology is the idea that the mind is an information processor. Information is thought to enter and exit the mind through a series of ordered processing stages (see Figure 8). Within these stages, various processes are assumed to act upon mental representations. Processes include comparing and matching. Mental representations are assumed to comprise images, mental models, rules, and other forms of knowledge.

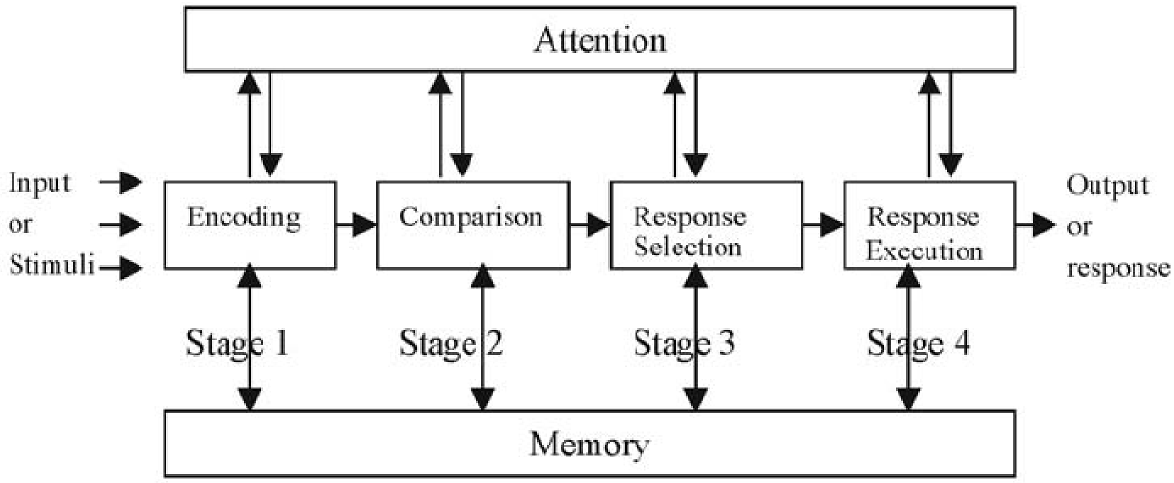

Figure 8 - Human information processing model

The information processing model provides a basis from which to make predictions about human performance. Hypotheses can be made about how long someone will take to perceive and respond to a stimulus (also known as reaction time) and what bottlenecks occur if a person is overloaded with too much information. One of the first HCI models to be derived from the information processing theory was the human processor model, which modeled the cognitive processes of a user interacting with a computer (Card *et al*, 1983). Cognition was conceptualized as a series of processing stages, where perceptual, cognitive, and motor processors are organized in relation to one another (see Figure 9). The model predicts which cognitive processes are involved when a user interacts with a computer, enabling calculations to be made of how long a user will take to carry out various tasks. In the 1980s, it was found to be a useful tool for comparing different word processors for a range of editing tasks.

The information processing approach was based on modeling mental activities that happen exclusively inside the head. Many have argued, however, that they do not adequately account for how people interact with computers and other devices, for example: -

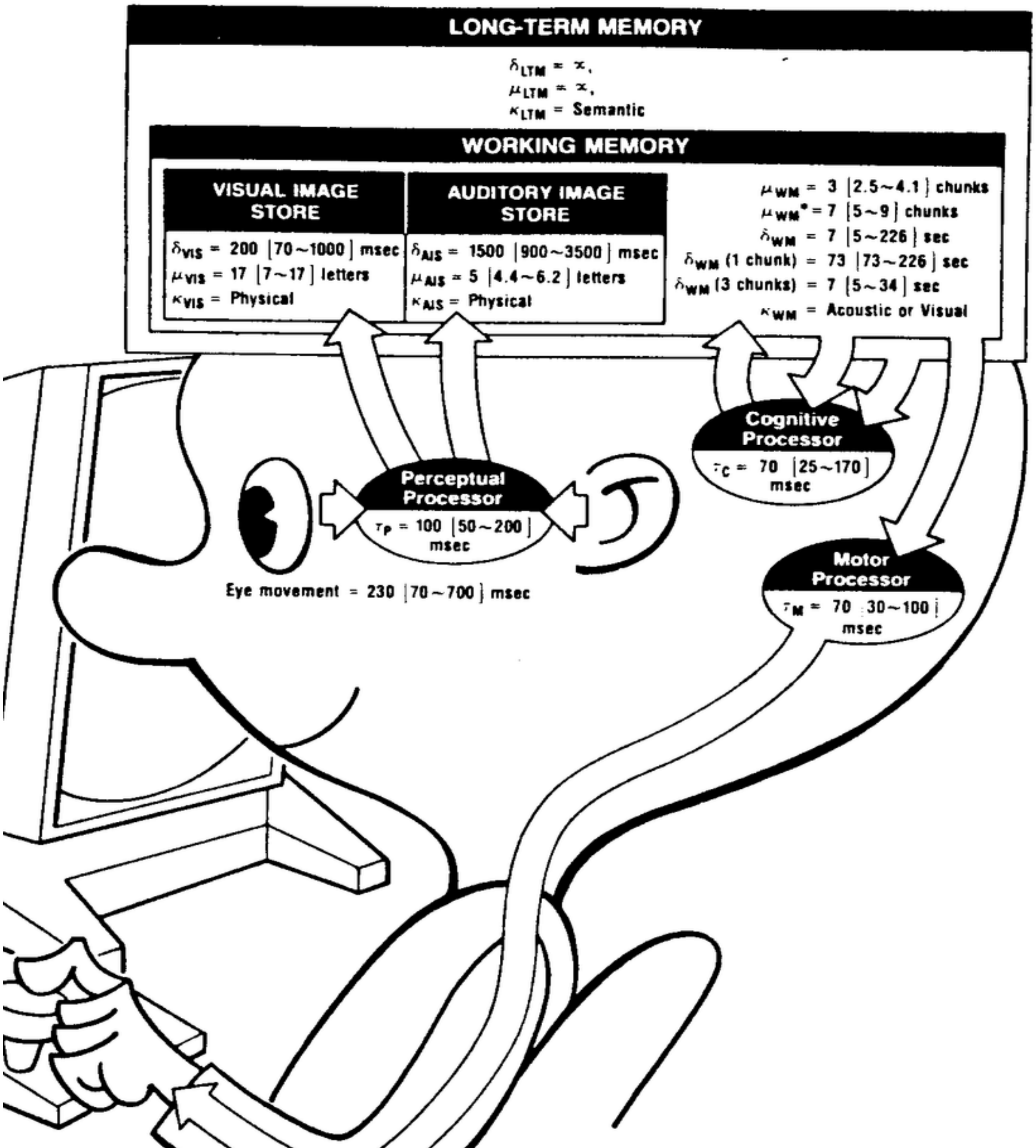

Figure 9 - The human processor model

The traditional approach to the study of cognition is to look at the pure intellect, isolated from distractions and from artificial aids. Experiments are performed in closed, isolated rooms, with a minimum of distracting lights or sounds, no other people to assist with the task, and no aids to memory or thought. The tasks are arbitrary ones, invented by the researcher. Model builders build simulations and descriptions of these isolated situations. The theoretical analyses are self-contained little structures, isolated from the world, isolated from any other knowledge or abilities of the person. (Norman, 1990, p. 5)

Instead, there has been an increasing trend to study cognitive activities in the context in which they occur, analyzing cognition as it happens in the wild (Hutchins, 1995). A central goal has been to look at how structures in the environment can both aid human cognition and reduce cognitive load. The three external approaches we consider next are distributed cognition, external cognition, and embodied cognition.

External

Distributed Cognition

Most cognitive activities involve people interacting with external kinds of representations, like books, documents, and computers – not to mention one another.

For example, when we go home from wherever we have been, we do not need to remember the details of the route because we rely on cues in the environment (e.g., we know to turn left at the red house, right when the road comes to a T-junction, and so on). Similarly, when we are at home we do not have to remember where everything is because information is out there. We decide what to eat and drink by scanning the items in the fridge, find out whether any messages have been left by glancing at the answering machine to see if there is a flashing light, and so on. Likewise, we are always creating external representations for a number of reasons, not only to help reduce memory load and the cognitive cost of computational tasks, but also, importantly, to extend what we can do and allow us to think more powerfully (Kirsh, 2010).

The distributed cognition approach studies the nature of cognitive phenomena across individuals, artifacts, and internal and external representations (Hutchins, 1995). Typically, it involves describing a cognitive system, which entails interactions among people, the artifacts they use, and the environment they are working in (see Figure 10).

An example of a cognitive system is an airline cockpit, where a top-level goal is to fly the plane. This involves: -

- The pilot, captain, and air traffic controller interacting with one another
- The pilot and captain interacting with the instruments in the cockpit
- The pilot and captain interacting with the environment in which the plane is flying (i.e., sky, runway).

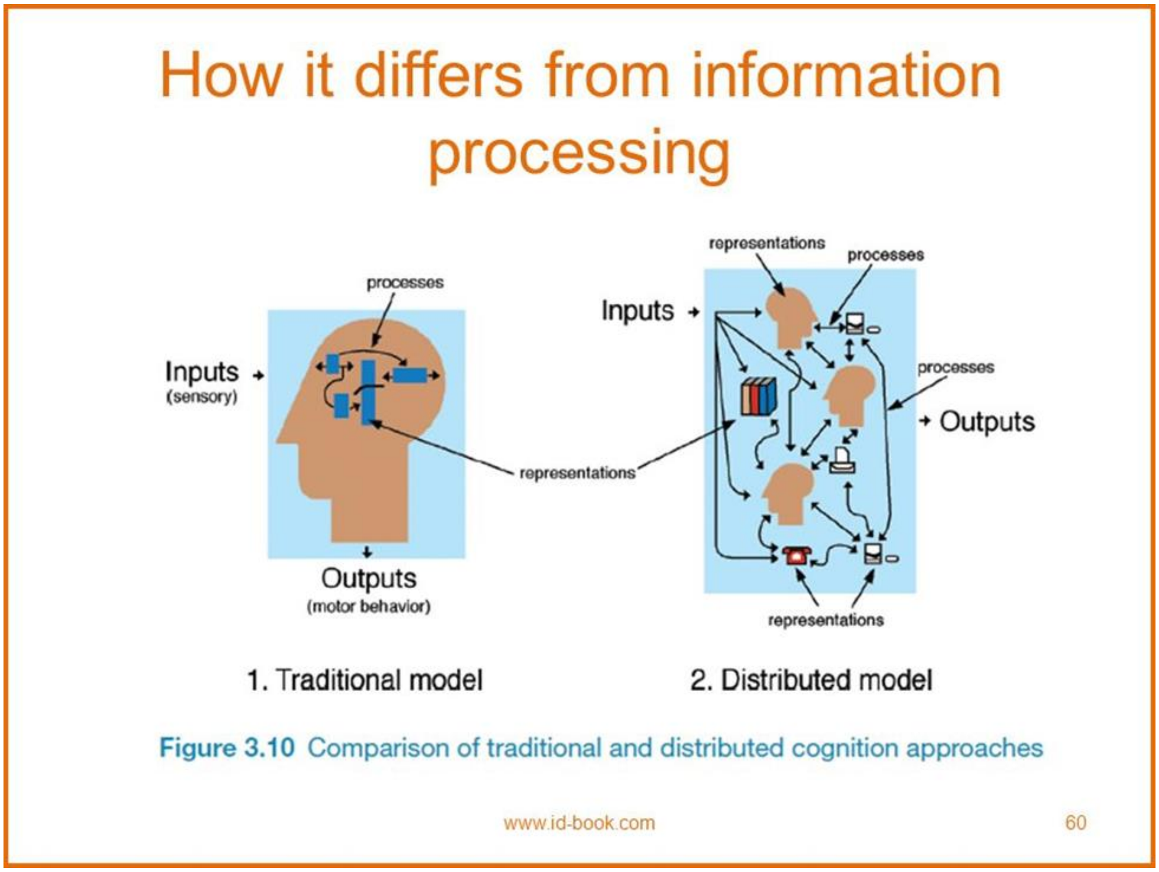

Figure 10 - Comparison of traditional and distributed cognition approaches

A primary objective of the distributed cognition approach is to describe these interactions in terms of how information is propagated through different media. By this is meant how information is represented and re-represented as it moves across individuals and through the array of artifacts that are used (e.g., maps, instrument readings, scribbles, spoken word) during activities. These transformations of information are referred to as changes in representational state.

This way of describing and analyzing a cognitive activity contrast with other cognitive approaches, such as the information processing model, in that it focuses not on what is happening inside the head of an individual, but on what is happening across a system of individuals and artifacts.

For example, in the cognitive system of the cockpit, a number of people and artifacts are involved in the activity of flying to a higher altitude. The air traffic controller initially tells the pilot when it is safe to fly to a higher altitude. The pilot then alerts the captain, who is flying the plane, by moving a knob on the instrument panel in front of them, indicating that it is now safe to fly (see Figure 11). Hence, the information concerning this activity is transformed through different media (over the radio, through the pilot, and via a change in the position of an instrument).

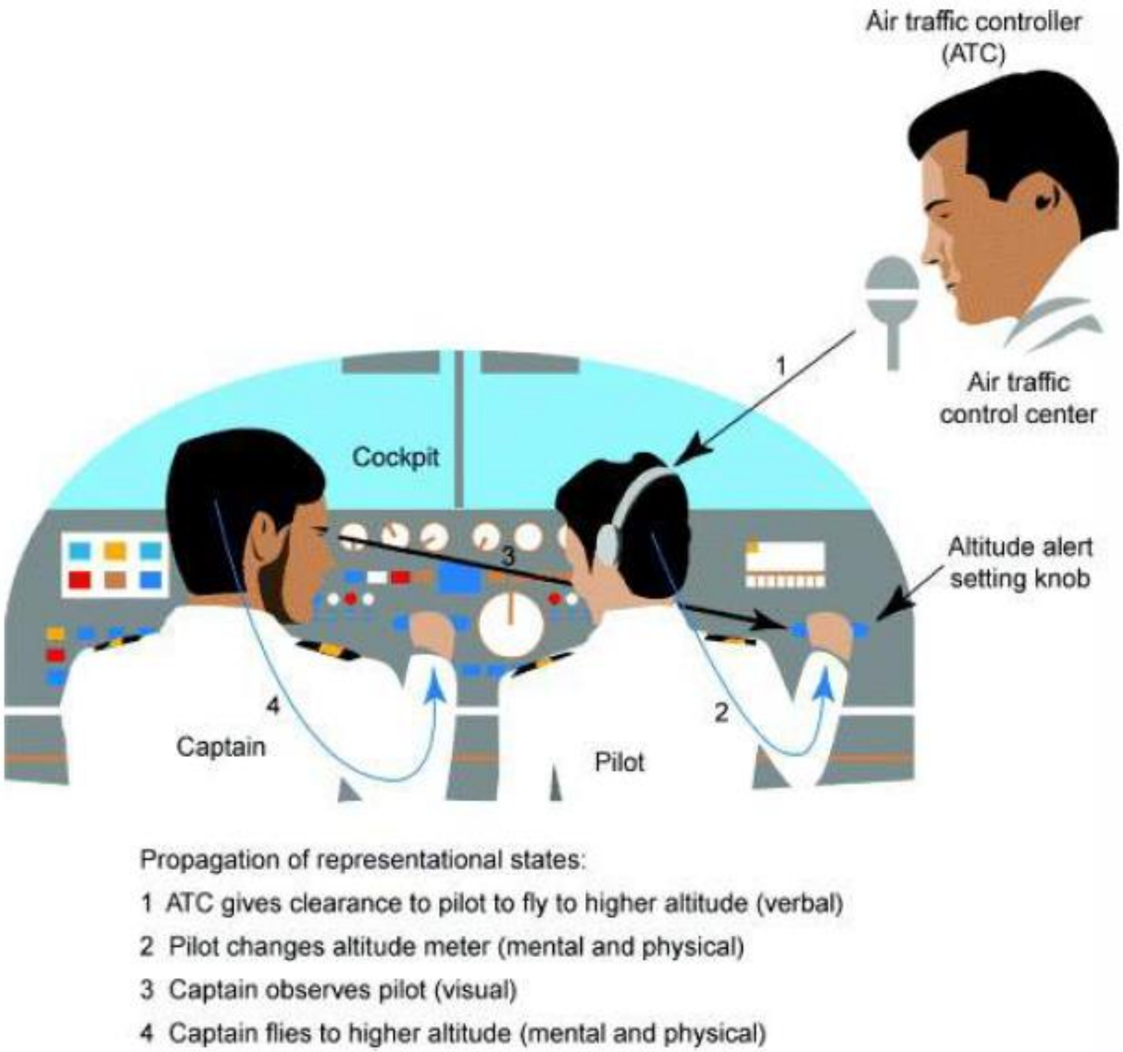

Figure 11 – A cognitive system in which information is propagated through different media

A distributed cognition analysis typically involves examining: -

- The distributed problem solving that takes place (including the way people work together to solve a problem).
- The role of verbal and non-verbal behavior (including what is said, what is implied by glances, winks, and the like, and what is not said).
- The various coordinating mechanisms that are used (e.g., rules, procedures).
- The various ways communication takes place as the collaborative activity progresses.
- How knowledge is shared and accessed.

External Cognition

People interact with or create information through using a variety of external representations, including books, multimedia, newspapers, web pages, maps, diagrams, notes, drawings, and so on. Furthermore, an impressive range of tools has been developed throughout history to aid cognition, including pens, calculators, and computer-based technologies. The combination of external representations and physical tools has greatly extended and supported people's ability to carry out cognitive activities (Norman, 2013). Indeed, they are such an integral part that it is difficult to imagine how we would go about much of our everyday life without them.

External cognition is concerned with explaining the cognitive processes involved when we interact with different external representations (Scaife and Rogers, 1996). A main goal is to explicate the cognitive benefits of using different representations for different cognitive activities and the processes involved.

The main ones include: -

1. Externalizing to reduce memory load
2. Computational offloading
3. Annotating and cognitive tracing.

Externalizing to Reduce Memory Load

Numerous strategies have been developed for transforming knowledge into external representations to reduce memory load. One such strategy is externalizing things we find difficult to remember, such as birthdays, appointments, and addresses. Diaries, personal reminders, and calendars are examples of cognitive artifacts that are commonly used for this purpose, acting as external reminders of what we need to do at a given time, like buy a card for a relative's birthday.

Other kinds of external representations that people frequently employ are notes, like sticky notes, shopping lists, and to-do lists. Where these are placed in the environment can also be crucial. For example, people often place notes in prominent positions, such as on walls, on the side of computer monitors, by the front door, and sometimes even on their hands, in a deliberate attempt to ensure they do remind them of what needs to be done or remembered. People also place things in piles in their offices and by the front door, indicating what needs to be done urgently and what can wait for a while.

Externalizing, therefore, can help reduce people's memory burden by: reminding them to do something (e.g., get something for mother's birthday); reminding them of what to do (e.g., buy a card); and reminding them of when to do something (e.g., send it by a certain date).

A number of smartphone apps have been developed to reduce the burden on people to remember things, including to-do and alarm- based lists. An example is Memory Aid, developed by Jason Blackwood. Figure 3.12 shows a screenshot from it of floating bubbles with keywords that relate to a to-do list.

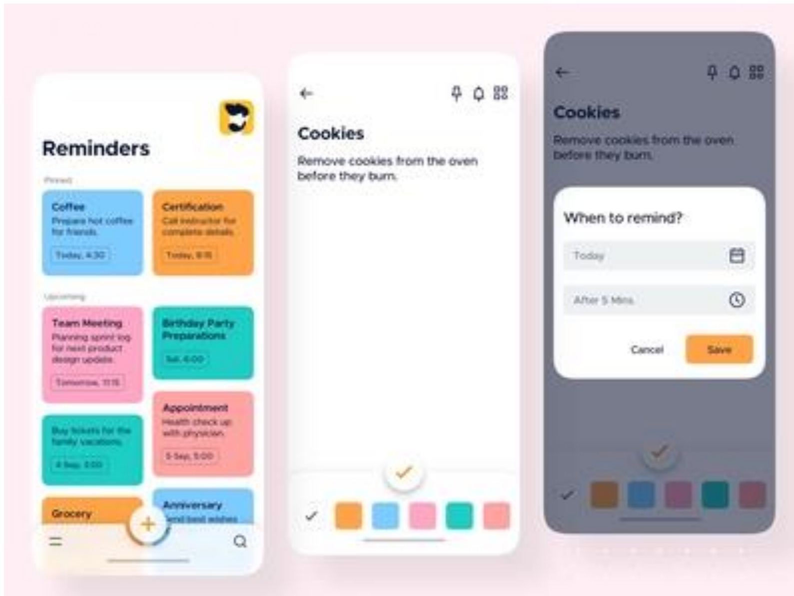

Figure 12 – A screenshot from a smartphone app for reminding users what to do

Computational Offloading

Computational offloading occurs when we use a tool or device in conjunction with an external representation to help us carry out a computation. An example is using pen and paper to solve a math problem.

Activity 3.5

1. Multiply 2 by 3 in your head. Easy. Now try multiplying 234 by 456 in your head. Not as easy. Try doing the sum using a pen and paper. Then try again with a calculator. Why is it easier to do the calculation with pen and paper and even easier with a calculator?
2. Try doing the same two sums using Roman numerals.

Comment

1. Carrying out the sum using pen and paper is easier than doing it in your head because you offload some of the computation by writing down partial results and using them to continue with the calculation. Doing the same sum with a calculator is even easier, because it requires only eight simple key presses. Even more of the computation has been offloaded onto the tool. You need only follow a simple internalized procedure (key in first number, then the multiplier sign, then next number, and finally the equals sign) and then read off the result from the external display.

2. Using Roman numerals to do the same sum is much harder: 2 times 3 becomes II × III, and 234 times 456 becomes CCXXXIV × CDLVI. The first calculation may be possible to do in your head or on a bit of paper, but the second is incredibly difficult to do in your head or even on a piece of paper (unless you are an expert in using Roman numerals or you cheat and transform it into Arabic numerals). Calculators do not have Roman numerals so it would be impossible to do on a calculator.

Hence, it is much harder to perform the calculations using Roman numerals than Arabic numerals – even though the problem is equivalent in both conditions. The reason for this is that the two kinds of representation transform the task into one that is easy and one that is more difficult, respectively. The kind of tool used also can change the nature of the task to being more or less easy.

Annotating and Cognitive Tracing

Another way in which we externalize our cognition is by modifying representations to reflect changes that are taking place that we wish to mark. For example, people often cross things off in a to-do list to show that they have been completed. They may also reorder objects in the environment by creating different piles as the nature of the work to be done changes. These two kinds of modification are called annotating and cognitive tracing:

Annotating involves modifying external representations, such as crossing off or underlining items. Cognitive tracing involves externally manipulating items into different orders or structures.

Annotating is often used when people go shopping. People usually begin their shopping by planning what they are going to buy. This often involves looking in their cupboards and fridge to see what needs stocking up. However, many people are aware that they won't remember all this in their heads and so often externalize it as a written shopping list. The act of writing may also remind them of other items that they need to buy, which they may not have noticed when looking through the cupboards. When they actually go shopping at the store, they may cross off items on the shopping list as they are placed in the shopping basket or cart. This provides them with an annotated externalization, allowing them to see at a glance what items are still left on the list that need to be bought. Some displays (e.g., tablet PCs, large interactive displays, and iPads) enable users to physically annotate documents, such as circling data or writing notes using styluses or their fingertips (see Chapter 6). The annotations can be stored with the document, enabling the users to revisit theirs or others' externalizations at a later date.

Cognitive tracing is useful in situations where the current state of play is in a state of flux and the person is trying to optimize her position. This typically happens when playing games, such as:

In a card game, when the continuous rearrangement of a hand of cards into suits, in ascending order, or collecting same numbers together helps to determine what cards to keep and which to play as the game progresses and tactics change.

In Scrabble, where shuffling around letters in the tray helps a person work out the best word given the set of letters (Maglio *et al*, 1999).

Cognitive tracing has also been used as an interactive function: for example, letting students know what they have studied in an online eLearning package. An interactive diagram can be used to highlight all the nodes visited, exercises completed, and units still to study.

A general cognitive principle for interaction design based on the external cognition approach is to provide external representations at an interface that reduce memory load and facilitate computational offloading. Different kinds of information visualizations can be developed that reduce the amount of effort required to make inferences about a given topic (e.g., financial forecasting, identifying programming bugs). In so doing, they can extend or amplify cognition, allowing people to perceive and do activities that they couldn't do otherwise. For example, information visualizations (see Chapter 6) represent masses of data in a visual form that can make it easier to make cross- comparisons across dimensions. GUIs are also able to reduce memory load significantly through providing external representations, e.g., Wizards and dialog boxes that guide users through their interactions.

Embodied Interaction

The concept of embodied interaction has become popular in interaction design and Human-computer interaction (HCI) since the publication of Dourish's (2001) book "*Where the Action Is*". It is about understanding interaction in terms of practical engagement with the social and physical environment. HCI, which grew out of collaborations between computer scientists and psychologists, initially adopted an information processing perspective. Dourish and others before him, such as Winograd and Flores (1986) and Suchman (1987), criticized this view of cognition as failing to account for the ways that people get things done in real situations. It provides a framing and organizing principle to help researchers uncover issues in the design and use of existing technologies and in the design of new systems.

It has been applied quite broadly to HCI, including work that focuses on the emotional quality of interaction with technology (Höök, 2008), on publicly available actions in physically shared spaces (Robertson, 1997), and on the role of the body in mediating our interaction with technology (Klemmer *et al*, 2006). Others have looked at how to apply a new generation of cognitive theories in interaction design (e.g., Antle *et al*, 2009; Hurtienne, 2009). These theories of embodied cognition are more grounded in the ways that people experience the world through physical interaction, but still emphasize the value of using abstraction from particular contexts.