Attention, Memory, and Wearable Interfaces

Thad E. Starner, Georgia Institute of Technology

"Do you use that while you drive?" This question is among the most common that people ask when they first see my wearable computer with its head-up display. My standard answer is, "No, driving is a highly visual task, so a visual display causes too much conflict. Instead, the wearable should interact through audio while the user is driving." During extended conversations, I might add, "My colleague Brad Rhodes has used his display while driving, but he mounted it in his cap where it only blocks the car's roof. He finds that it is more effective for providing directions while driving than trying to check a map in his lap."

This dialog hints at the deeper issue of the limits of human attention and how wearable interfaces should be developed to complement, not interfere, with normal human capabilities. Most interfaces on desktop computers do not have this problem; desktop interface designers can assume that the user is solely concentrating on the digital task. However, a major advantage of a wearable is that users can take it anywhere and use it anytime. Often the wearable is used to support a primary task such as conversing or fixing a car engine. In such situations, the standard desktop Windows, Icons, Menus, and Pointer interface is hardly ever appropriate. Most current WIMP interfaces require fine motor control and hand and eye coordination—the user must select an item from a series of cascading menus. Even after 15 years of practice with a mouse, I can barely manage the complex menus associated with Windows 2000's "Start" menu. Imagine trying to use such a system while driving or, in the case of wearables, even walking down a street!

Wearable interfaces must be adapted to a wearer's environment and task instead of using default desktop interfaces. A heuristic is that a wearable interface should be designed to maximize performance and minimize investment in attention. However, what does "attention" mean and how do we quantify it?

ATTENTION

What we commonly know as attention consists of several different but interrelated abilities.[1] Focused attention is the ability to concentrate on particular information selected from a mass of data—such as holding a conversation in a noisy restaurant. Here, I am concerned with divided attention, which is the ability to allocate attention to simultaneous tasks. For divided-attention tasks such as driving and navigating, Lund reported that our competence is affected by task similarity, task difficulty, and practice.[1,2]

In some sense, these factors are intuitive. For example, it seems logical that a person can simultaneously perform two easy tasks with less difficulty than two hard tasks. Similarly, the automobile example provides an intuition of task similarity. For the presentation of navigation information while driving, a highly visual task, audio is more dissimilar than video and is thus a better choice. Everyday experience confirms this—automobile drivers routinely converse with passengers but rarely attempt to watch television.

DUAL TASKS

Much cognitive psychology literature explores the human capacity for attention while performing multiple tasks. D.A. Allport and his colleagues performed key dual task experiments in 1972.[3] The researchers divided their subjects into three groups and asked each group to "shadow" prose played into one ear of their headphones. Shadowing simply means orally repeating, word for word, the prose as it is played. Humans are naturally good at this task—think of telephoning an information service (such as a recording of movies and show times) and repeating the message to others in the room. In addition to the shadowing task, Allport's first group also tried to learn a list of words that were played through the other ear of the headphones. The second group attempted to learn the list from a presentation of the same words as presented on a screen. The third group learned the list as pictures representing the words on a screen. A control study showed that the percentage

of participants' word recall without the shadowing task was approximately the same for all presentation methods. However, when performing both tasks, the group that learned the list through pictures outperformed the other groups. In fact, the picture-based recall rate was very close to the respective control group without shadowing.

This experiment suggests that listening to the word list or reading it from a screen uses some of the same attentional resources as shadowing prose. Yet, identifying and memorizing pictures of those same words does not seem to conflict with shadowing; thus, the subjects could perform both tasks simultaneously. In another experiment, the researchers discovered that musicians can shadow prose and simultaneously play the piano while sight reading. In fact, with practice, their performance was not significantly worse on either task as compared to when they performed the tasks separately.[3]

Do these experiments indicate that wearable interfaces should solely concentrate on situations in which humans can easily perform dual tasks? Certainly these experiments suggest that wearable interface designers should carefully consider their user's tasks and environment before deciding on a given modality for interaction. Wearable interfaces might even offer several different user-selectable interaction modes or automatically change interaction modes depending on the user's context.

However, practice also plays a major role in the user's ability to perform dual tasks. For example, the musicians in the study discussed earlier made fewer errors after gaining more experience with the study. Experienced piano players and experienced drivers have performed their respective tasks so often that the tasks have become "automatic." Such automatic processing can be learned in a relatively short time period (at least when compared to the life expectancy of major interface designs).[4] Such learning effects should be considered when designing wearable interfaces.

THE CELL PHONE CONTROVERSY

The controversy over the use of cellular phones in automobiles has increased public awareness of the issues of divided attention. However, few studies have been performed on the effects of mobile phone use in actual driving conditions. One exception is a study by the Japanese National Police Agency for the month of June 1996. Out of the accidents attributable, in part, to a mobile phone, only 16 percent were while the driver was talking on the phone, while 80 percent involved physical manipulation of the phone (for example, looking at the keypad to dial a number or dropping the phone while trying to answer it).[1] Presumably, the amount of time drivers talked on the phone was significantly more than the time spent dialing or answering the phone, reinforcing the idea that the manipulation of the phone is a much more risky behavior than simply talking on it.

Perhaps the major problem with mobile phone use while driving is that the main resources for driving the car—that is, visual and manual attention—are shared with the task of interacting with the physical interface of the phone. Governments are enacting laws to help alleviate the problem. For example, beginning 1 November 2001, it is illegal to use a mobile phone while driving in the State of New York unless the interface is "hands-free." Unfortunately, though, the exact meaning of "hands-free" (such as use of voice dialing or simply that the phone is not supported by the hand while talking) might not be specified in such laws or even in studies that are often quoted on the subject. For example, Donald Redelmeier's and Robert Tibshirani's "Association Between Cellular-Telephone Calls and Motor Vehicle Collisions"[2] is widely reported as finding no difference between the risk associated with hands-free and hand-held cell phones, yet the study groups all cell phones that are not handheld into the same "hands-free" category. With recent advances such as speech control, specific interfaces must be evaluated before classifying a group of products as too risky for their benefits.

REFERENCES

1. United States National Highway Traffic Safety Administration, "An Investigation of the Safety Implications of Wireless Communication in Vehicles," 1997.

2. D. Redelmeier and R. Tibshirani, "Association Between Cellular-Telephone Calls and Motor Vehicle Collisions," *The New England J. Medicine*, vol. 336, no. 7, Feb. 1997, pp. 453–458.

OVERCOMING NOVEL INTERFACES: LEARNING GRAFFITI AND QWERTY

In mobile computing, the Graffiti alphabet (introduced with the Palm PDA or Xerox PARC's Unistroke system) exemplifies how a novel interface can be sufficient for beginners yet benefit significantly from the effects of practice. Beginners often find these stroke-based alphabets difficult and annoying because they differ significantly from their normal writing. Yet, the alphabet is familiar enough so that novice users can quickly access the PDA's basic functionality. With time, users can enter entire words without having to think about individual pen strokes. Reportedly, Unistroke was designed so that users could even write without looking.

Owing to the Palm's success, the pen computer market has blossomed over the past five years. Ironically, in the early 1990s, pen computer manufacturers argued adamantly that users would not accept a handwriting recognition system that did not allow "natural" handwriting because they would not take the time to learn the new interface. Yet, when the Palm appeared with its small form factor and an affordable price, users did change their behavior to gain the desired functionality.

Another example is the QWERTY keyboard. The Internet enticed a new market segment of naive users into computing, yet alphabetically formatted

WEARABLE COMPUTING

next issue

Wearables are often associated with head-up displays. Yet misconceptions abound regarding how such displays are used and the properties that make a good head-up display for a wearable. We'll take a quick look into the psychophysics and practical use of these tools.

keyboards are still difficult to find. Beginners can initially gain enough functionality on a QWERTY keyboard to tolerate the seemingly bizarre layout of characters and with time can convert their "hunt and peck" typing into something that resembles automatic processing. Such users quickly reach the stage where they can converse while entering a URL.

AUGMENTING MEMORY

What wearable-based functionality will users find so compelling that they will learn new interfaces? The same need that pushed the PDA into the market—augmenting memory—will help drive wearables. Often, the original Palm was used as a rolodex, a calendar, and as scrap paper (in the form of electronic ink)—all functions that assist the user's memory over different time scales. Obviously this functionality is compelling: billions of dollars worth of mobile devices have been sold. Yet recently I have seen evidence of users abandoning their devices. Whenever I give an invited talk, I ask how many members of the audience own PDAs. Given the technical leanings of the audience, approximately 50 percent will raise their hands. However, when I ask how many of these PDA owners actually have the device with them in the auditorium, only half of the owners raise their hands. While only an informal experiment, this 50 percent abandonment rate of a relatively accepted product class is surprising. Why is this?

My students and I are working on a survey on the use (and disuse) of mobile scheduling devices (such as paper day planners, PDAs, scrap paper, and so on) to examine this question. Preliminary results indicate that access time to the interface might play a major role.[5] Specifically, it took an average of 20 seconds for our subjects to remove a PDA from a pocket or bag, boot it, and reach the point where they could enter an appointment. Perhaps users value PDA functionality, but the inconvenience of accessing the device inhibits its casual use. Another example of the inconvenience of the PDA interface is that it is very difficult to use a PDA to write down an idea while walking. The interface requires a hand to hold the device, the other hand to use the stylus, both eyes, and careful attention to navigate the interface to where a note can be saved. Careful studies into PDA use are needed, but such discrepancies between desired functionality and ease of use might present an interesting business opportunity for wearable computing companies.

Another possible problem with PDA interfaces is that they could interfere with the user's main task, such as conversation. I'm very interested in this possibility. I developed my wearable's interface, first as a junior member of BBN Technologies' speech group and later as a junior graduate student at MIT, to help remember important parts of conversations with senior personnel and students. For example, before having my wearable, I found it embarrassing to have to email a colleague to ask him to send me a needed computer command that he had just told me in the hallway. With a wearable computer, I found that I could type `tar -xvf /dev/nrst1`—the command to extract a backup tape on a Sun at the time—during a 30-second passing conversation.

While I find wearables' note-taking abilities compelling, there remains a lingering concern in many people's minds on the cognitive load of the keyboard interface during conversations. Could the act of taking notes actually inhibit the ability to comprehend and contribute to a conversation? Students are very familiar with this conundrum; when attending difficult classes they often complain that they can either pay attention in class and understand the material, or take careful notes, but not both. This problem motivated my use of a head-up display.

In 1991, I tried to make a wearable based on an early laptop. The idea was to design the system so that it could be mounted on the arm (although the project never reached that stage). While typing was much faster than writing, I found that looking up and down between the laptop display and the blackboard was still too distracting from the lecture. The Private Eye, a crisp 720×280 head-up display that could be focused to appear at different distances, provided the solution. The head-up display let the user pay attention to the blackboard and the lecturer while maintaining a background awareness that the text was being entered correctly. The same idea, overlaying a virtual image on a real one, is used in fighter jets to avoid the extra load of the pilot moving his head and switching his attention from his instrumentation to his cockpit view.

Yet I still have not made a convincing argument that note taking does not significantly interfere with conversation or even the process of understanding a lecture. Perhaps typing interferes with the process of encoding memories, as some physical tasks have been shown to do[6] as Schacter described in *The Seven Sins of Memory*?[7] Or, perhaps the dual tasks of typing and conversing are more similar to the earlier piano studies?

However, there is a third possibility. In my experience, I find that I can take notes while conversing if the notes are short enough. In the few moments after a conversation (for example, while walking to my next meeting), I will rearrange the notes and complete any that are incomplete. However, if I try to take word-for-word dictation during a con-

versation, I can only type one sentence before I cannot concentrate on the conversation. I hypothesize that I can take short notes successfully (as opposed to dictation) because I quickly shift my attention between two related tasks. Conversation is slow and full of redundancies when compared to reading. If a colleague says something of interest, I can devote my attention for a few seconds to typing a few words into my wearable to remind me of the concept. My attention switches back to the conversation in time to catch the next concept. This method fails, however, during involved research conversations (see the Herbert Simon sidebar). When the conversation is full of specific details I wish to remember and it continues to introduce new concepts, I can lose track.

Fortunately, the rules of conversational speech allow for recovery in socially graceful ways. Phrases such as "Hold on, are you saying that ..." and "What was that number again?" are already common in everyday speech. Only with the most complex concepts and difficult conversations do I need to say "Hold on for a second, let me get that down." Interestingly, though, I cannot switch rapidly between reading email and an unrelated conversation. Perhaps the disjoint topics conflict for my attention. Or perhaps the situation is analogous to the shadowing task in which the subjects were simultaneously learning a list of words from text on a screen.

I'm afraid this essay raises many more questions than it answers. Yet, I feel that the issue of attention and memory in wearable interfaces will be key to developing compelling wearable products in the future. Every year the cognitive psychology literature explores and refines key questions in attention and memory. How do we apply the lessons from such laboratory-based studies to the highly mobile world of wearables? Fortunately, the field provides the answer: use wearable hardware to adapt such laboratory procedures to our tasks. I hope we will see more such

studies in International Symposium on Wearable Computers (ISWC), Ubiquitous Computing International Conference (Ubicomp), and Conference on Human Factors in Computing Systems (CHI) in coming years. Perhaps with the information we gain, wearables can avoid some of the pitfalls of current mobile technology and appear on the consumer market with interfaces that truly augment, not interfere, with our everyday lives. ⚏

HERBERT SIMON ON ATTENTION

The late Herbert Simon, Noble Prize and Turing Award winner, once wrote

What information consumes is rather obvious: it consumes the attention of its recipients. Hence a wealth of information creates a poverty of attention and a need to allocate that attention efficiently among the overabundance of information sources that might consume it.[1]

While the quote is from 1971, Simon had been talking about such ideas since the 1950s. Simon introduced the concept of *satisficing*, which states that people make decisions that are "good enough" instead of "optimal" in the face of a glut of choices or information. Otherwise, we would spend all of our time simply processing the possibilities. Today, we are faced with more information than ever. Intelligent agents on wearable computers can help us remember and attend to the information that is most important. However, agent designers must take care that their designs help the problem instead of exacerbating it.

A sampling of Simon's books

Administrative Behavior (1947), *Models of Man* (1957), *Organizations* (coauthored with James March, 1958), and *The Sciences of the Artificial* (1969).

REFERENCE

1. H. Simon, "Designing Organizations for an Information-Rich World," *Computers, Communications and the Public Interest*, Martin Greenberg, ed., Johns Hopkins Press, Baltimore, 1971, pp 40–41.

REFERENCES

1. N. Lund, *Attention and Pattern Recognition*, Routledge, East Sussex, UK, 2001.

2. M. Eysenck and M. Keane, *Cognitive Psychology: A Student's Handbook*, 3rd ed., Lawrence Erlbaum Associates, Hove, UK, 1995.

3. D.A. Allport, B. Antonis, and P. Reynolds, "On the Division of Attention: A Disproof of the Single Channel Hypothesis, *Quarterly J. Experimental Psychology*, vol. 24, 1972, pp. 225–235.

4. W. Schneider and R. Shiffrin, "Controlled and Automatic Human Information Processing: Detection, Search, and Attention," *Psychological Review*, vol. 84, 1977, pp. 1–66.

5. B. Wong, T. Starner, and R. McGuire, "Towards Conversational Speech Recognition for a Wearable Computer-Based Appointment Scheduling Agent," tech report, GVU Center, Georgia Institute of Technology, 2002, no. 02–17.

6. T. Shallice et al., "Brain Regions Associated with Acquisition and Retrieval of Verbal Episodic Memory, *Nature*, vol. 368, 1994, pp. 633–635.

7. D. Schacter, *The Seven Sins of Memory*, Houghton Mifflin, Boston, 2001.

Thad E. Starner is an assistant professor of computing at the Georgia Institute of Technology, where he directs the Contextual Computing Group in the Institute's College of Computing. Contact him at thad@cc.gatech.edu.