Commentary on Section 4

Eye Tracking in Human–Computer Interaction and Usability Research: Ready to Deliver the Promises

Robert J. K. Jacob and Keith S. Karn

Introduction

This section considers the application of eye movements to user interfaces, both for analyzing interfaces (measuring usability) and as an actual control medium within a human–computer dialogue. The two areas have generally been reported separately; but this book seeks to tie them together. For usability analysis, the user's eye movements while using the system are recorded and later analyzed retrospectively, but the eye movements do not affect the interface in real time. As a direct control medium, the eye movements are obtained and used in real time as an input to the user–computer dialogue. They might be the sole input, typically for disabled users or hands-busy applications, or they might be used as one of several inputs, combining with mouse, keyboard, sensors, or other devices.

Interestingly, the principal challenges for both retrospective and real time eye tracking in human–computer interaction (HCI) turn out to be analogous. For retrospective analysis, the problem is to find appropriate ways to use and interpret the data; it is not nearly as straightforward as it is with more typical task performance, speed, or error data. For real time use, the problem is to find appropriate ways to respond judiciously to eye movement input, and avoid over-responding; it is not nearly as straightforward as responding to well-defined, intentional mouse or keyboard input. We will see in this chapter how these two problems are closely related.

These uses of eye tracking in HCI have been highly promising for many years, but progress in making good use of eye movements in HCI has been slow to date. We see promising research work, but we have not yet seen wide use of these approaches in practice or in the marketplace. We will describe the promises of this technology, its limitations, and the obstacles that must still be overcome. Work presented in this book and elsewhere shows that the field is indeed beginning to flourish.

History of Eye Tracking in HCI

The study of eye movements pre-dates the widespread use of computers by almost 100 years (for example, Javal, 1878/1879). Beyond mere visual observation, initial methods for tracking the location of eye fixations were quite invasive — involving direct mechanical contact with the cornea. Dodge and Cline (1901) developed the first precise, non-invasive eye tracking technique, using light reflected from the cornea. Their system recorded only horizontal eye position onto a falling photographic plate and required the participant's head to be motionless. Shortly after this, Judd, McAllister and Steel (1905) applied motion picture photography in order to record the temporal aspects of eye movements in two dimensions. They recorded the movement of a small white speck of material inserted into the participants' eyes rather than light reflected directly from the cornea. These and other researchers interested in eye movements made additional advances in eye tracking systems during the first half of the twentieth century by combining the corneal reflection and motion picture techniques in various ways (see Mackworth & Mackworth, 1958 for a review).

In the 1930s, Miles Tinker and his colleagues began to apply photographic techniques to study eye movements in reading (see Tinker, 1963 for a thorough review of this work). They varied typeface, print size, page layout, etc. and studied the resulting effects on reading speed and patterns of eye movements. In 1947 Paul Fitts and his colleagues (Fitts, Jones & Milton, 1950) began using motion picture cameras to study the movements of pilots' eyes as they used cockpit controls and instruments to land an airplane. The Fitts *et al.* study represents the earliest application of eye tracking to what is now known as *usability engineering* — the systematic study of users interacting with products to improve product design.

Around that time Hartridge and Thompson (1948) invented the first head-mounted eye tracker. Crude by current standards, this innovation served as a start to freeing eye tracking study participants from tight constraints on head movement. In the 1960s, Shackel (1960) and Mackworth and Thomas (1962) advanced the concept of head-mounted eye tracking systems making them somewhat less obtrusive and further reducing restrictions on participant head movement. In another significant advance relevant to the application of eye tracking to Human–computer interaction, Mackworth and Mackworth (1958) devised a system to record eye movements superimposed on the changing visual scene viewed by the participant.

Eye movement research and eye tracking flourished in the 1970s, with great advances in both eye tracking technology and psychological theory to link eye tracking data to cognitive processes. See for example books resulting from eye movement conferences during this period (i.e., Monty & Senders, 1976; Senders, Fisher & Monty, 1978; Fisher, Monty & Senders, 1981). Much of the work focused on research in psychology and physiology and explored how the human eye operates and what it can reveal about perceptual and cognitive processes. But publication records from the 1970s indicate a lull in activity relating eye tracking to usability engineering. We presume this occurred largely due to the effort involved not only with data collection, but even more so with data analysis. As Monty (1975) puts it: "It is not uncommon to spend days processing data that took only minutes to collect" (pp. 331–332). Work

in several human factors/usability laboratories (particularly those linked to military aviation) focused on solving the shortfalls with eye tracking technology and data analysis during this timeframe. Researchers in these laboratories recorded much of their work in US military technical reports (see Simmons, 1979 for a review).

Much of the relevant work in the 1970s focused on technical improvements to increase accuracy and precision and reduce the impact of the trackers on those whose eyes were tracked. The discovery that multiple reflections from the eye could be used to dissociate eye rotations from head movement (Cornsweet & Crane, 1973) increased tracking precision and also prepared the ground for developments resulting in greater freedom of participant movement. Using this discovery, two joint military/industry teams (US Airforce/Honeywell Corporation and US Army/EG&G Corporation) each developed a remote eye tracking system that dramatically reduced tracker obtrusiveness and its constraints on the participant (see Lambert, Monty & Hall, 1974; Monty, 1975; Merchant *et al.*, 1974 for descriptions). These joint military/industry development teams and others made even more important contributions with the automation of eye tracking data analysis. The advent of the minicomputer in that general timeframe provided the necessary resources for high-speed data processing. This innovation was an essential precursor to the use of eye tracking data in real-time as a means of Human–computer interaction (Anliker, 1976). Nearly all eye tracking work prior to this used the data only retrospectively, rather than in real time (in early work, analysis could only proceed after film was developed). The technological advances in eye tracking during the 1960s and 1970s are still seen reflected in most commercially available eye tracking systems today (see Collewijn, 1999 for a recent review).

Psychologists who studied eye movements and fixations prior to the 1970s generally attempted to avoid cognitive factors such as learning, memory, workload, and deployment of attention. Instead their focus was on relationships between eye movements and simple visual stimulus properties such as target movement, contrast, and location. Their *solution* to the *problem* of higher-level cognitive factors had been "to ignore, minimize or postpone their consideration in an attempt to develop models of the supposedly simpler lower-level processes, namely, sensorimotor relationships and their underlying physiology" (Kowler, 1990, p. 1). But this attitude began to change gradually in the 1970s. While engineers improved eye tracking technology, psychologists began to study the relationships between fixations and cognitive activity. This work resulted in some rudimentary, theoretical models for relating fixations to specific cognitive processes (see for example work by Just & Carpenter, 1976a, 1976b). Of course scientific, educational, and engineering laboratories provided the only home for computers during most of this period. So eye tracking was not yet applied to the study of human–computer interaction at this point. Teletypes for command line entry, punched paper cards and tapes, and printed lines of alphanumeric output served as the primary form of Human–computer interaction.

As Senders (2000) pointed out, the use of eye tracking has persistently come back to solve new problems in each decade since the 1950s. Senders likens eye tracking to a Phoenix raising from the ashes again and again with each new generation of engineers designing new eye tracking systems and each new generation of cognitive psychologists tackling new problems. The 1980s were no exception. As personal

computers proliferated, researchers began to investigate how the field of eye tracking could be applied to issues of human–computer interaction. The technology seemed particularly handy for answering questions about how users search for commands in computer menus (see, for example, Card, 1984; Hendrickson, 1989; Altonen, Hyrskykari & Räihä, 1998; Byrne *et al.,* 1999). The 1980s also ushered in the start of eye tracking in real time as a means of human–computer interaction. Early work in this area initially focused primarily on disabled users (e.g., Hutchinson *et al*., 1989; Levine, 1981, Levine, 1984). In addition, work in flight simulators attempted to simulate a large, ultra-high resolution display by providing high resolution wherever the observer was fixating and lower resolution in the periphery (Tong & Fisher, 1984). The combination of real-time eye movement data with other, more conventional modes of user–computer communication was also pioneered during the 1980s (Bolt, 1981, 1982; Levine, 1984; Glenn *et al*., 1986; Ware & Mikaelian, 1987).

In more recent times, eye tracking in human–computer interaction has shown modest growth both as a means of studying the usability of computer interfaces and as a means of interacting with the computer. As technological advances such as the Internet, e-mail, and videoconferencing evolved into viable means of information sharing during the 1990s and beyond, researchers again turned to eye tracking to answer questions about usability (e.g., Benel, Ottens & Horst, 1991; Ellis *et al.*, 1998; Cowen, 2001) and to serve as a computer input device (e.g., Starker & Bolt, 1990; Vertegaal, 1999; Jacob, 1991; Zhai, Morimoto & Ihde, 1999). We will address these two topics and cover their recent advances in more detail with the separate sections that follow.

Eye Movements in Usability Research

Why "Rising From the Ashes" Rather Than "Taking Off Like Wildfire?"

As mentioned above, the concept of using eye tracking to shed light on usability issues has been around since before computer interfaces, as we know them. The pioneering work of Fitts, Jones and Milton (1950) required heroic effort to capture eye movements (with cockpit-mounted mirrors and movie camera) and to analyze eye movement data with painstaking frame-by-frame analysis of the pilot's face. Despite large individual differences, Fitts and his colleagues made some conclusions that are still useful today. For example, they proposed that fixation frequency is a measure of a display's importance; fixation duration, as a measure of difficulty of information extraction and interpretation; and the pattern of fixation transitions between displays; as a measure of efficiency of the arrangement of individual display elements.

Note that it was also Paul Fitts whose study of the relationships between the duration, amplitude, and precision of human movements published four years later (Fitts, 1954) is still so widely cited as "Fitts' Law." A look at the ISI Citation Index[1] reveals that in the past 29 years Fitts *et al.*'s 1950 cockpit eye movement study was only cited 16 times[2] while Fitts' Law (Fitts, 1954) has been cited 855 times. So we ask, why has

Fitts' work on predicting movement time been applied so extensively while his work in the application of eye tracking been so slow to catch on? Is it simply a useless concept? We think not. The technique has continually been classified as promising over the years since Fitts' work. Consider the following quotes:

> For a long time now there has been a *great need* for a means of recording where people are looking while they work at particular tasks. A whole series of unsolved problems awaits such a technique (Mackworth & Thomas, 1962, p. 713; emphasis added).

> [T]he eyetracking system has a *promising future* in usability engineering (Benel, Ottens & Horst, 1991, p. 465; emphasis added).

> [A]ggregating, analyzing, and visualizing eye tracking data in conjunction with other interaction data holds *considerable promise* as a powerful tool for designers and experimenters in evaluating interfaces (Crowe & Narayanan, 2000, p. 35; emphasis added).

> Eye-movement analysis does appear to be a *promising* new tool for evaluating visually administered questionnaires (Redline & Lankford, 2001; emphasis added).

> Another *promising* area is the use of eye-tracking techniques to support interface and product design. Continual improvements in . . . eye-tracking systems . . . have increased the usefulness of this technique for studying a variety of interface issues (Merwin, 2002, p. 39; emphasis added).

Why has this technique of applying eye tracking to usability engineering been classified as simply "promising" over the past 50 years? For a technology to be labeled "promising" for so long is both good news and bad. The good news is that the technique must really be promising; otherwise it would have been discarded by now. The bad news is that something has held it up in this merely promising stage. There are a number of probable reasons for this slow start, including technical problems with eye tracking in usability studies, labor-intensive data extraction, and difficulties in data interpretation. We will consider each of these three issues in the following sections.

Technical Problems with Eye Tracking in Usability Studies

Technical issues that have plagued eye tracking in the past, making it unreliable and time consuming are resolving slowly (see Collewijn, 1999; Goldberg & Wichansky, this volume). By comparison to techniques used by Fitts and his team, modern eye tracking systems are incredibly easy to operate. Today, commercially available eye tracking systems suitable for usability laboratories are based on video images of the eye. These trackers are mounted either on the participant's head or remotely, in front of the participant (e.g., on a desktop). They capture reflections of infrared light from

both the cornea and the retina and are based on the fundamental principles developed in the pioneering work of the 1960s and 1970s reviewed earlier. Vendors typically provide software to make setup and calibration relatively quick and easy. Together these properties make modern eye tracking systems fairly reliable and easy to use. The ability to track participants' eyes is much better than with systems of the recent past. There are still problems with eye tracking a considerable minority of participants (typically 10 to 20% cannot be tracked reliably). Goldberg and Wichansky (this volume) present some techniques to maximize the percentage of participants whose eyes can be tracked. For additional practical guidance in eye tracking techniques see Duchowski (2003).

The need to constrain the physical relationship between the eye tracking system and the participant remains one of the most significant barriers to incorporation of eye tracking in more usability studies. Developers of eye tracking systems have made great progress in reducing this barrier, but existing solutions are far from optimal. Currently the experimenter has the choice of a remotely mounted eye tracking system that puts some restrictions on the participant's movement or a system that must be firmly (and uncomfortably) mounted to the participant's head. Of course the experimenter has the option of using the remote tracking system and not constraining the user's range and speed of head motion, but must then deal with frequent track losses and manual reacquiring of the eye track. In typical WIMP (i.e., windows, icons, menus, and pointer) human–computer interfaces, constraining the user's head to about a cubic foot of space may seem only mildly annoying. If, however, we consider human–computer interaction in a broader sense and include other instances of "ubiquitous computing" (Weiser, 1993), then constraining a participant in a usability study can be quite a limiting factor. For example, it would be difficult to study the usability of portable systems such as a personal digital assistant or cell phone, or distributed computer peripherals such as a printer or scanner while constraining the user's movement to that typically required by commercially available remote eye trackers. Recent advances in eye tracker portability (Land, 1992; Land, Mennie & Rusted, 1999; Pelz & Canosa, 2001; Babcock, Lipps & Pelz, 2002) may largely eliminate such constraints. These new systems can fit un-tethered into a small backpack and allow the eye tracking participant almost complete freedom of eye, head, and whole body movement while interacting with a product or moving through an environment. Of course such systems still have the discomfort of the head-mounted systems and add the burden of the backpack. Another solution to the problem of eye tracking while allowing free head movement integrates a magnetic head tracking system with a head mounted eye tracking system (e.g., Iida, Tomono & Kobayashi, 1989). These systems work best in an environment free of ferrous metals and they add complexity to the eye tracking procedure. Including head tracking also results in an inevitable decrease in precision due to the integrating of the two signals (eye-in-head and head-in-world).

We see that currently available eye trackers have progressed considerably from systems used in early usability studies; but they are far from optimized for usability research. For a list, and thorough discussion, of desired properties of eye tracking systems see Collewijn (1999). We can probably safely ignore Collewijn's call for a 500 Hz sampling rate, as 250 Hz is sufficient for those interested in fixations rather

than basic research on saccadic eye movements. See the comparison of "saccade pickers" and "fixation pickers" in Karn *et al.* (2000). For wish lists of desired properties of eye tracking systems specifically tailored for usability research see Karn, Ellis and Juliano (1999, 2000) and Goldberg and Wichansky (this volume).

Labor-intensive Data Extraction

Most eye trackers produce signals that represent the orientation of the eye within the head or the position of the point of regard on a display at a specified distance. In either case, the eye tracking system typically provides a horizontal and vertical coordinate for each sample. Depending on the sampling rate (typically 50 to 250 Hz), and the duration of the session, this can quickly add up to a lot of data. One of the first steps in data analysis is usually to distinguish between fixations (times when the eye is essentially stationary) and saccades (rapid re-orienting eye movements). Several eye tracker manufacturers, related commercial companies, and academic research labs now provide analysis software that allows experimenters to extract quickly the fixations and saccades from the data stream (see for example Lankford, 2000; Salvucci, 2000). These software tools typically use either eye position (computing dispersion of a string of eye position data points known as *proximity analysis*), or eye velocity (change in position over time). Using such software tools the experimenter can quickly and easily know when the eyes moved, when they stopped to fixate, and where in the visual field these fixations occurred. Be forewarned, however, that there is no standard technique for identifying fixations (see Salvucci & Goldberg, 2000 for a good overview). Even minor changes in the parameters that define a fixation can result in dramatically different results (Karsh & Breitenbach, 1983). For example, a measure of the number of fixations during a given time period would not be comparable across two studies that used slightly different parameters in an automated fixation detection algorithm. Goldberg and Wichansky (this volume) call for more standardization in this regard. At a minimum, researchers in this field need to be aware of the effects of the choices of these parameters and to report them fully in their publications.[3]

The automated software systems described above might appear to eliminate completely the tedious task of data extraction mentioned earlier. While this may be true if the visual stimulus is always known as in the case of a static observer viewing a static scene, even the most conventional human–computer interfaces can hardly be considered static. The dynamic nature of modern computer interfaces (e.g., scrolling windows, pop-up messages, animated graphics, and user-initiated object movement and navigation) provides a technical challenge for studying eye fixations. For example, knowing that a person was fixating 10 degrees above and 5 degrees to the left of the display's center does not allow us to know what object the person was looking at in the computer interface unless we keep track of the changes in the computer display. Note that if Fitts were alive today to repeat the study of eye tracking of military pilots he would run into this problem with dynamic electronic displays in modern cockpits. These displays allow pilots to call up different flight information on the same display depending on the pilot's changing needs throughout a flight. Certainly less

conventional human–computer interaction with ubiquitous computing devices provide similar challenges. Recent advances integrating eye tracking and computer interface navigation logging enable the mapping of fixation points to visual stimuli in some typical dynamic human–computer interfaces (Crowe & Narayanan, 2000; Reeder, Pirolli & Card, 2001). These systems account for user and system initiated display changes such as window scrolling and pop-up messages. Such systems are just beginning to become commercially available, and should soon further reduce the burden of eye tracking data analysis.

A dynamically changing scene caused by head or body movement of the participant in, or through, an environment provides another challenge to automating eye tracking data extraction (Sheena & Flagg, 1978). Head-tracking systems are now often integrated with eye tracking systems and can help resolve this problem (Iida, Tomono & Kobayashi, 1989), but only in well defined visual environments (for a further description of these issues see Sodhi, Reimer, Cohen, Vastenburg, Kaars & Kirschenbaum, 2002). Another approach is image processing of the video signal captured by a head-mounted scene camera to detect known landmarks (Mulligan, 2002).

Despite the advances reviewed above, researchers are often left with no alternative to the labor-intensive manual, frame-by-frame coding of videotape depicting the scene with a cursor representing the fixation point. This daunting task remains a hindrance to more widespread inclusion of eye tracking in usability studies.

Difficulties in Data Interpretation

Assuming a researcher, interested in studying the usability of a human–computer interface, is not scared off by the technical and data extraction problems discussed above, there is still the issue of making sense out of eye tracking data. How does the usability researcher relate fixation patterns to task-related cognitive activity?

Eye tracking data analysis can proceed either *top-down* — based on cognitive theory or design hypotheses, or *bottom-up* — based entirely on observation of the data without predefined theories relating eye movements to cognitive activity (see Goldberg, Stimson, Lewenstein, Scott & Wichansky, 2002). Here are examples of each of these processes driving data interpretation:

Top-down based on a cognitive theory. Longer fixations on a control element in the interface reflect a participant's difficulty interpreting the proper use of that control.
Top-down based on a design hypothesis. People will look at a banner advertisement on a web page more frequently if we place it lower on the page.
Bottom-up. Participants are taking much longer than anticipated making selections on this screen. We wonder where they are looking.

Reviewing the published reports of eye tracking applied to usability evaluation, we see that all three of these techniques are commonly used. While a top-down approach may seem most attractive (perhaps even necessary to infer cognitive processes from eye fixation data), usability researchers do not always have a strong theory or hypothesis to

drive the analysis. In such cases, the researchers must, at least initially, apply a data-driven search for fixation patterns. In an attempt to study stages of consumer choice, for example, Russo and Leclerc (1994) simply looked at video tapes of participants' eye movements, coded the sequence of items fixated, and then looked for and found common patterns in these sequences. Land, Mennie and Rusted (1999) performed a similar type of analysis as participants performed the apparently simple act of making a cup of tea. Even when theory is available to drive the investigation, researchers usually reap rewards from a bottom-up approach when they take the time to replay and carefully examine scan paths superimposed on a representation of the stimulus.

To interpret eye tracking data, the usability researcher must choose some aspects (dependent variables or metrics) to analyze in the data stream. A review of the literature on this topic reveals that usability researchers use a wide variety of eye tracking metrics. In fact the number of different metrics is fewer than it may at first appear due to the lack of standard terminology and definitions for even the most fundamental concepts used in eye tracking data interpretation. Readers may feel bogged down in a swamp of imprecise definitions and conflicting uses of the same terms. If we look closely at this mire we see that differences in eye tracking data collection and analysis techniques often account for these differences in terminology and their underlying concepts. For example, in studies done with simple video or motion picture imaging of the participants' face (e.g., Fitts, *et al.*, 1950; Card, 1984; Svensson *et al.*, 1997) a "fixation" by its typical definition cannot be isolated. Researchers usually realize this, but nevertheless, some misuse the term "fixation" to refer to a series of consecutive fixations within an area of interest. In fact, the definition of the term "fixation" is entirely dependent on the size of the intervening saccades that can be detected and which the researcher wants to recognize. With a high-precision eye tracker, even small micro-saccades might be counted as interruptions to fixation (see Engbert & Kliegl, this volume, for a discussion).

Eye Tracking Metrics Most Commonly Reported in Usability Studies   The usability researcher must choose eye tracking metrics that are relevant to the tasks and their inherent cognitive activities for each usability study individually. To provide some idea of these choices, Table 1 summarizes 21 different usability studies that have incorporated eye tracking.[4] The table includes a brief description of the users, the task and the eye tracking related metrics used by the authors. Note that rather than referring to the same concept by the differing terms used by the original authors, we have attempted to use a common set of definitions as follows:

Fixation: A relatively stable eye-in-head position within some threshold of dispersion (typically ~2°) over some minimum duration (typically 100–200 ms), and with a velocity below some threshold (typically 15–100 degrees per second).
Gaze Duration: cumulative duration and average spatial location of a series of consecutive fixations within an area of interest. Gaze duration typically includes several fixations and may include the relatively small amount of time for the short saccades between these fixations. A fixation occurring outside the area of interest marks the end of the gaze.[5] Authors cited in Table 1 have used "dwell."[6] "glance," or "fixation cycle," in place of "gaze duration."

Table 1: Summary of 21 usability studies incorporating eye tracking.

Authors/Date	Users and Tasks	Eye Tracking Related Metrics
Fitts, *et al.* (1950)	40 military pilots. Aircraft landing approach.	• Gaze rate (# of gazes/minute) on each area of interest • Gaze duration mean, on each area of interest • Gaze % (proportion of time) on each area of interest • Transition probability between areas of interest
Harris and Christhilf (1980)	4 instrument-rated pilots. Flying maneuvers in a simulator	• Gaze % (proportion of time) on each area of interest • Gaze duration mean, on each area of interest
Kolers, Duchnicky and Ferguson (1981)	20 university students. Reading text on a CRT in various formats and with various scroll rates.	• Number of fixations, overall • Number of fixations on each area of interest (line of text) • Number of words per fixation • Fixation rate overall (fixations/s) • Fixation duration mean, overall
Card (1984)	3 PC users. Searching for and selecting specified item from computer pull-down menu.	• Scan path direction (up/down) • Number of fixations, overall
Hendrickson (1989)	36 PC users. Selecting 1 to 3 items in various styles of computer menus.	• Number of fixations, overall • Fixation rate overall (fixations/s) • Fixation duration mean, overall • Number of fixations on each area of interest • Fixation rate on each area of interest • Fixation duration mean, on each area of interest • Gaze duration mean, on each area of interest • Gaze % (proportion of time) on each area of interest • Transition probability between areas of interest
Graf and Kruger (1989)	6 participants. Search for information to answer questions on screens of varying organization.	• Number of *voluntary* (>320 ms) fixations, overall • Number of *involuntary* (<240 ms) fixations, overall • Number of fixations on target
Benel, Ottens and Horst (1991)	7 PC users. Viewing web pages.	• Gaze % (proportion of time) on each area of interest • Scan path
Backs and Walrath (1992)	8 engineers. Symbol search and counting tasks on color or monochrome displays.	• Number of fixations, overall • Fixation duration mean, overall • Fixation rate overall (fixations/s)
Yamamoto and Kuto (1992)	7 young adults. Confirm sales receipts (unit price, quantity, etc.) on various screen layouts.	• Scan path direction • Number of instances of backtracking

Table 1: continued.

Authors/Date	Users and Tasks	Eye Tracking Related Metrics
Svensson *et al.* (1997)	18 military pilots. Fly and monitor threat display containing varying number of symbols.	• Gaze duration mean, on each area of interest • Frequencies of long duration dwells on area of interest
Altonen *et al.* (1998)	20 PC users. Select menu item specified directly or by concept definition.	• Scan path direction • *Sweep* — scan path progressing in the same direction • Number of fixations per sweep
Ellis *et al.* (1998)	16 PC users with web experience. Directed web search and judgment.	• Number of fixations, overall • Fixation duration mean, overall • Number of fixations on each area of interest • Time to 1st fixation on target area of interest • Gaze % (proportion of time) on each area of interest
Kotval and Goldberg (1998)	12 university students. Select command button specified directly from buttons grouped with various strategies.	• Scan path duration • Scan path length • Scan path area (convex hull) • Fixation spatial density • Transition density • Number of fixations, overall • Fixation duration mean, overall • Fixation/saccade time ratio • Saccade length
Byrne *et al.* (1999)	11 university students. Choosing menu items specified directly from computer pull-down menus of varying length.	• Number of fixations, overall • First area of interest fixated • Number of fixations on each area of interest
Flemisch and Onken (2000)	6 military pilots. Low-level flight and navigation in a flight simulator using different display formats.	• Gaze % (proportion of time) on each area of interest
Redline and Lankford (2001)	25 adults. Fill out a 4-page questionnaire (of various forms) about lifestyle.	• Scan path
Cowen (2001)	17 PC users with web experience. Search/ extract information from web pages.	• Fixation duration total • Number of fixations, overall • Fixation duration mean, overall • Fixation spatial density
Josephson and Holmes (2002)	8 university students with web experience. Passively view web pages.	• Scan path

Table 1: continued.

Authors/Date	Users and Tasks	Eye Tracking Related Metrics
Goldberg, Stimson, Lewenstein, Scott and Wichansky (2002)	7 adult PC users with web experience. Search/extract information from web pages.	• Number of fixations on each area of interest • Fixation duration mean, on each area of interest • Saccade length • Fixation duration total, on each area of interest • Number of areas of interest fixated • Scan path length • Scan path direction • Transition probability between areas of interest
Albert (2002)	24 intermediate to advanced web users. Web search for purchase and travel arrangements on sites with varying banner ad placement.	• Number of fixations on area of interest (banner ad) • Gaze % (proportion of time) on each area of interest • Participant % fixating on each area of interest
Albert and Liu (in press)	12 licensed drivers. Simultaneous driving and navigation using electronic map in simulator.	• Number of dwells, overall • Gaze duration mean, on area of interest (map) • Number of dwells on each area of interest

Area of interest: Area of a display or visual environment that is of interest to the research or design team and thus defined by them (not by the participant).
Scan path: Spatial arrangement of a sequence of fixations.

When we count up the number of times each metric is used (both from the 21 studies included in Table 1 and the three studies reported by Zülch & Stowasser, this volume), we find the most commonly used metrics listed below. The number in parentheses after each metric is the number of studies in which it is used out of the total of the 24 studies reviewed.

Number of fixations, overall (11)
Gaze % (proportion of time) on each area of interest (7)
Fixation duration mean, overall (6)
Number of fixations on each area of interest (6)
Gaze duration mean, on each area of interest (5)
Fixation rate overall (fixations/s) (5)

Each of these six most frequently used metrics is discussed briefly below. For more detailed discussion of these and other metrics see Goldberg and Kotval (1998); Kotval and Goldberg (1998) and Zülch and Stowasser (this volume).

Number of fixations, overall. The number of fixations overall is thought to be negatively correlated with search efficiency (Goldberg & Kotval, 1998; Kotval & Goldberg, 1998). A larger number of fixations indicates less efficient search possibly resulting from a poor arrangement of display elements. The experimenter should consider the relationship of the number of fixations to task time (i.e., longer tasks will usually require more fixations).

Gaze % (proportion of time) on each area of interest. The proportion of time looking at a particular display element (of interest to the design team) could reflect the importance of that element. Researchers using this metric should be careful to note that it confounds frequency of gazing on a display element with the duration of those gazes. According to Fitts *et al.* (1950) these should be treated as separate metrics, with duration reflecting difficulty of information extraction and frequency reflecting the importance of that area of the display.

Fixation duration mean, overall. Longer fixations (and perhaps even more so, longer gazes) are generally believed to be an indication of a participant's difficulty extracting information from a display (Fitts *et al*, 1950; Goldberg & Kotval, 1998).

Number of fixations on each area of interest. This metric is closely related to gaze rate, which is used to study the number of fixations across tasks of differing overall duration. The number of fixations on a particular display element (of interest to the design team) should reflect the importance of that element. More important display elements will be fixated more frequently (Fitts *et al*, 1950).

Gaze duration mean, on each area of interest. This is one of the original metrics in Fitts *et al.* (1950). They predicted that gazes on a specific display element would be longer if the participant experiences difficulty extracting or interpreting information from that display element.

Fixation rate overall (fixations/s). This metric is closely related to fixation duration. Since the time between fixations (typically short duration saccadic eye movements) is relatively small compared with the time spent fixating, fixation rate should be approximately the inverse of the mean fixation duration.

Other Promising Eye Tracking Metrics   Although the metrics presented above are the most popular, they are not necessarily always the best metrics to apply. Other important metrics to consider include:

Scan path (sequence of fixations) and derived measures such as the transition probability between areas of interest — can indicate the efficiency of the arrangement of elements in the user interface.

Number of gazes on each area of interest — is a simple, but often forgotten, measure. Gazes (the concatenation of successive fixations within the same area of interest) are often more meaningful than counting the number of individual fixations.

Number of involuntary and number of voluntary fixations — Graf & Kruger (1989) have proposed that short fixations (<240 ms) and long fixations (>320 ms) be classified as involuntary and voluntary fixations respectively. Further research is needed to validate this method of classifying fixations.

Percentage of participants fixating an area of interest — can serve as a simple indicator of the attention-getting properties of an interface element.
Time to 1st fixation on target area of interest — is a useful measure when a specific search target exists.

Other aspects of ocular-motor performance such as blinks (e.g., Stern, Boyer & Schroeder, 1994), pupil changes (e.g., Hoeks & Levelt, 1993; Marshall, 1998; Backs & Walrath, 1992), vergence, and accommodation can be exploited. These have been considered annoying problems by most eye movement researchers in the past, but they may be a rich source of data. For example, Brookings, Wilson and Swain (1996) report that blink rate is more sensitive to workload (related to task difficulty) than many other more conventionally used eye tracking measures including saccade rate and amplitude in a demanding visual task (air traffic control).

Other researchers have come up with innovative techniques for analyzing and presenting existing eye tracking metrics. Wooding (2002), for example, has introduced the "Fixation Map" for conveying the most frequently fixated areas in an image. Land *et al.* (1999) refer to "Object-related actions" as a neat way to combine eye tracking data with other participant behaviors such as reaching and manipulation movements. Harris and Christhilf (1980) used an innovative plot of gaze percentage versus average gaze duration to classify types of displays and the ways they are used by pilots. Josephson and Holmes (2002) apply optimal matching (or string-edit) analysis for comparing fixation sequences. To study more truly a participant's fixation of an object, Pelz, Canosa and Babcock (2000) combine times when the eye is moving with respect to the head, but fixed relative to the fixated visual object. Such situations occur when the participant visually pursues a moving object or compensates for head movements via the vestibuloocular reflex (VOR). Salvucci (2000) used an automated data analysis system to test predictions made by various models of cognitive processes.

Difficulties relating eye tracking data to cognitive activity is probably the single most significant barrier to the greater inclusion of eye tracking in usability studies. The most important question to ask when incorporating eye tracking into a usability study is "what aspects of eye position will help explain usability issues?" As discussed above, the most relevant metrics related to eye position vary from task to task and study to study. Sometimes the experimenter has to risk going on a bit of a "fishing expedition" (i.e., collect some eye tracking records and examine them closely in various ways before deciding on the most relevant analyses).

Current and Future Directions for Applying Eye Tracking in Usability Engineering

From the literature reviewed in the preceding sections, we see that the application of eye tracking in usability engineering is indeed beginning to flourish. In this volume, we have additional contributions to this growing field. Goldberg and Wichansky (this volume) provide a thorough introduction to two groups of readers: eye tracking scien-

tists who wish to apply their work to product usability evaluations; and usability engineers who wish to incorporate eye tracking into their studies. Goldberg and Wichansky also provide some practical tips that will be helpful to anyone who is interested in integrating eye tracking in usability engineering. A great deal of more basic research using eye tracking continues to produce results that are applicable to the design of human–computer interaction. The studies by Crosby and Sophian, reported in the current volume, are examples of such work. Here they show that shape comparison can be an effective way to compare ratio data visually. Surprisingly few fixations are needed to compare two shapes. Zülch and Stowasser (current volume) report a series of usability studies using eye tracking. Their laboratory is one of the few of which we are aware that uses these techniques fairly regularly. In their chapter, Zülch and Stowasser report results from a series of studies of industrial manufacturing software where users must solve scheduling problems, search for object or relationship information in list or graphic representation, or search for/extract information from a visual database. Comparing structured and unstructured approaches in problem solving and search tasks, they find differences in eye tracking data.

These studies and the historical work reviewed in previous paragraphs, point to some springboards for further study. We list what we believe to be the greatest opportunities for future work here:

Many usability studies that have incorporated eye tracking have indicated a difference between novice and more experienced participants (Fitts *et al*, 1950; Crosby & Peterson, 1991; Card, 1984; Aaltonen *et al.*, 1998) and individual differences (Yarbus, 1967 [1965]; Card, 1984; Andrews & Coppola, 1999). Eye tracking seems like an especially useful tool to study repetitive or well-practiced tasks and "power usability" (Karn, Krolczyk & Perry, 1997) and the process by which people evolve from novice users to expert users.

When users search for a tool, menu item, icon, etc. in a typical human–computer interface, they often do not have a good representation of the target. Most of the literature in visual search starts with the participant knowing the specific target. We need more basic research in visual search when the target is not known completely. A more realistic search task is looking for the tool that will help me do a specific task, having not yet seen the tool.

More work is needed to resolve the technical issues with eye trackers and the analysis of the data they produce. These issues include constraints on participant movement; tracker accuracy, precision, ease of setup; dealing with dynamic stimuli; and labor-intensive data extraction.

While there is a wealth of literature dealing with fixation patterns both in reading and in picture perception, little data exists on the viewing of pictures and text in combination as they often occur in instruction materials, news media, advertising, multimedia content, etc. (Stolk *et al.*, 1993; Hegarty, 1992; Hegarty & Just, 1993). This seems like fertile ground for the application of eye tracking in usability evaluation.

While there has been considerable use of eye tracking in usability engineering over the 50+ years since Fitts' pioneering work, the concept has not caught on with

anywhere near the popularity of Fitts' Law for human limb movement. We see however, that just in the past ten years, significant technological advances have made the incorporation of eye tracking in usability research much more feasible. As a result, we are already seeing a rapid increase in the adoption of eye tracking in usability labs. We anticipate that future application of these techniques will allow the human–computer interaction design community to learn more about users' deployment of visual attention and to design product interfaces that more closely fit human needs.

Finally we remind the reader always to explore multiple facets of usability. Various measures of usability are necessary to gather the whole picture (see for example Frokjær, Hertzum & Hornbæk, 2000). Eye tracking alone is not a complete usability engineering approach, but it can make a significant contribution to the assessment of usability.

Input from the Eye

Background

We turn now to eye movements as a real time input medium. First, why would one want to use eye movements interactively in a user interface? We can view the basic task of human–computer interaction as moving information between the brain of the user and the computer. Our goal is to increase the useful bandwidth across that interface with faster, more natural, and more convenient communication mechanisms. Most current user interfaces provide much more bandwidth from the computer to the user than in the opposite direction. Graphics, animations, audio, and other media can output large amounts of information rapidly, but there are hardly any means of inputting comparably large amounts of information from the user. Today's user–computer dialogues are thus typically rather one-sided. New input devices and media that use "lightweight," passive measurements can help redress this imbalance by obtaining data from the user conveniently and rapidly. The movements of a user's eyes can thus provide a convenient high-bandwidth source of additional user input.

Eye trackers have existed for a number of years, but their use has largely been confined to laboratory experiments. The equipment is gradually becoming sufficiently robust and inexpensive to consider use in real user–computer interfaces. What is now needed is research in appropriate interaction techniques that incorporate eye movements into the user–computer dialogue in a convenient and natural way.

The simplest solution would be to substitute an eye tracker directly for a mouse — install an eye tracker and use its *x, y* output stream in place of that of the mouse. Changes in the user's line of gaze would directly cause the mouse cursor to move. But the eye moves very differently from the intentional way the hand moves a mouse; this would work poorly and be quite annoying.

There are significant differences between a manual input source like the mouse and eye position to be considered in designing eye movement-based interaction techniques:

Eye movement input is distinctly faster than other current input media (Ware & Mikaelian, 1987, Sibert & Jacob, 2000). Before the user operates any mechanical pointing device, he or she usually looks at the destination to which he or she wishes to move. Thus the eye movement is available as an indication of the user's goal before he or she could actuate any other input device.

"Operating" the eye requires no training or particular coordination for normal users; they simply look at an object. The control-to-display relationship for this device is already established in the brain.

The eye is, of course, much more than a high-speed cursor positioning tool. Unlike any other input device, an eye tracker also tells where the user's interest is focused. By the very act of pointing with this device, the user changes his or her focus of attention; and every change of focus is available as a pointing command to the computer. A mouse input tells the system simply that the user intentionally picked up the mouse and pointed it at something. An eye tracker input could be interpreted in the same way (the user intentionally pointed the eye at something, because he or she was trained to operate this system that way). But it can also be interpreted as an indication of what the user is currently paying attention to, without any explicit input action on his or her part.

This same quality is also a problem for using the eye as a computer input device. Moving one's eyes is often an almost subconscious act. Unlike a mouse, it is relatively difficult to control eye position consciously and precisely at all times. The eyes continually dart from spot to spot, even when its owner thinks he or she is looking steadily at a single object, and it is not desirable for each such move to initiate a computer command.[7]

Similarly, unlike a mouse, eye movements are always "on." There is no natural way to indicate when to engage the input device, as there is with grasping or releasing the mouse. Closing the eyes is rejected for obvious reasons — even with eye tracking as input, the principal function of the eyes in the user–computer dialogue is for communication *to* the user. Eye movements are an example of a more general problem with many new passive or non-command input media, requiring either careful interface design to avoid this problem or some form of explicit "clutch" to engage and disengage the monitoring.

Also, in comparison to a mouse, eye tracking lacks an analogue of the integral buttons most mice have. Using blinks as a signal is a less than ideal solution because it detracts from the naturalness possible with an eye movement-based dialogue by requiring the user to think about when to blink.

Finally, eye tracking equipment is still far less stable and accurate than most manual input devices.

The problem with a simple implementation of an eye tracker interface is that people are not accustomed to operating devices simply by moving their eyes. They expect to be able to look at an item without having the look *mean* something. At first, it is empowering to be able simply to look at what you want and have it happen, rather than having to look at it (as you would anyway) and then point and click it with the mouse. Before long, though, it becomes like the "Midas Touch." Everywhere you look,

another command is activated; you cannot look anywhere without issuing a command. The challenge in building a useful eye tracker interface is to avoid this Midas Touch problem. Ideally, the interface should act on the user's eye input when he or she wants it to and let the user just look around when that's what he wants, but the two cases are impossible to distinguish in general. Instead, researchers develop interaction techniques that address this problem in specific cases.

The key is to make wise and effective use of eye movements. Like other passive, lightweight, non-command inputs (e.g., gesture, conversational speech), eye movements are often non-intentional or not conscious, so they must be interpreted carefully to avoid annoying the user with unwanted responses to his actions. A user does not say or mean much by a movement of the eyes — far less than by a keyboard command or mouse click. The computer ought to respond with correspondingly small, subtle responses. Rearranging the screen or opening a new window would typically be too strong a response. More appropriate actions might be highlighting an object for future action, showing amplifying information on a second screen, or merely downloading extra information in case it is requested.

We have also found in informal evaluation of eye movement interfaces that, when all is performing well, eye gaze interaction can give a subjective feeling of a highly responsive system, almost as though the system is executing the user's intentions before he or she expresses them (Jacob, 1991). This, more than raw speed, is the real benefit we seek from eye movement-based interaction. Work presented in this book shows some of the ways to achieve this goal. For example, Illi, Isokoski, and Surakka present a sophisticated user interface that subtly incorporates the eye movement data with other lightweight inputs, and fuses them to decide more accurately just when the system should react to the eye.

Survey of Past Work in Eye Tracking for HCI Input

Using eye movements for Human–computer interaction in real time has been studied most often for disabled (typically quadriplegic) users, who can use only their eyes for input (e.g., Hutchinson, 1989; Levine, 1981, 1984), report work for which the primary focus was disabled users). Because all other user–computer communication modes are unavailable, the resulting interfaces are rather slow and tricky to use for non-disabled people, but, of course, a tremendous boon to their intended users.

One other case in which real-time eye movement data has been used in an interface is to create the illusion of a better graphic display. The chapter by O'Sullivan, Dingliana, and Howlett covers a variety of new ways to do this for both graphics and behavior. Earlier work in flight simulators attempted to simulate a large, ultra-high resolution display (Tong & Fisher, 1984). With this approach, the portion of the display that is currently being viewed is depicted with high resolution, while the larger surrounding area (visible only in peripheral vision) is depicted in lower resolution. Here, however, the eye movements are used essentially to simulate a better display device, but do not alter the basic user–computer dialogue.

A relatively small amount of work has focused on the more general use of real-time eye movement data in HCI in more conventional user–computer dialogues, alone or in combination with other input modalities. Richard Bolt did some of the earliest work and demonstrated several innovative uses of eye movements (Bolt, 1981; Bolt, 1982; Starker & Bolt, 1990). Floyd Glenn (Glenn *et al.*, 1986) used eye movements for several tracking tasks involving moving targets. Ware and Mikaelian (1987) reported an experiment in which simple target selection and cursor positioning operations were performed approximately twice as fast with an eye tracker than with any of the more conventional cursor positioning devices. More recently, the area has flourished, with much more research and even its own separate conference series, discussed below in this section.

In surveying research in eye movement-based human–computer interaction we can draw two distinctions, one in the nature of the user's eye movements and the other, in the nature of the responses. Each of these could be viewed as *natural* (that is, based on a corresponding real-world analogy) or *unnatural* (no real world counterpart). In eye movements as with other areas of user interface design, it is helpful to draw analogies that use people's already-existing skills for operating in the natural environment and then apply them to communicating with a computer. One of the reasons for the success of direct manipulation interfaces is that they draw on analogies to existing human skills (pointing, grabbing, moving objects in physical space), rather than trained behaviors; virtual reality interfaces similarly exploit people's existing physical navigation and manipulation abilities. These notions are more difficult to extend to eye movement-based interaction, since few objects in the real world respond to people's eye movements. The principal exception is, of course, other people: they detect and respond to being looked at directly and, to a lesser and much less precise degree, to what else one may be looking at. We can view the user eye movements and the system responses separately as natural or unnatural:

User's eye movements: Within the world created by an eye movement-based interface, users could move their eyes to scan the scene, just as they would a real world scene, unaffected by the presence of eye tracking equipment (i.e., natural eye movement). The alternative is to instruct users of the eye movement-based interface to move their eyes in particular ways, not necessarily those they would have employed if left to their own devices, in order to actuate the system (i.e., unnatural or learned eye movements).

Nature of the response: Objects could respond to a user's eye movements in a natural way, that is, the object responds to the user's looking in the same way real objects do. As noted, there is a limited domain from which to draw such analogies in the real world. The alternative is unnatural response, where objects respond in ways not experienced in the real world.

This suggests a taxonomy of four possible styles of eye movement-based interaction:

Natural eye movement/Natural response: This area is a difficult one, because it draws on a limited and subtle domain, principally how people respond to other people's gaze. Starker and Bolt provide an excellent example of this mode, drawing on the

analogy of a tour guide or host who estimates the visitor's interests by his or her gazes (Starker & Bolt, 1990). Another example related to this category is the use of eye movements in videoconferencing systems (Vertegaal, 1999). Here the goal is to transmit the correct eye position from one user to another (by manipulating camera angle or processing the video image) so that the recipient (rather than a computer) can respond naturally to the gaze. The work described above, in which eye movement input is used to simulate a better display, is also related to this category.

Natural eye movement/Unnatural response: In our work (Jacob, 1991), we have used natural (not trained) eye movements as input, but we provide responses unlike those in the real world. This is a compromise between full analogy to the real world and an entirely artificial interface. We present a display and allow the user to observe it with his or her normal scanning mechanisms, but such scans then induce responses from the computer not normally exhibited by real world objects.

Unnatural eye movement/Unnatural response: Most previous eye movement-based systems have used learned ("unnatural") eye movements for operation and thus, of necessity, unnatural responses. Much of that work has been aimed at disabled or hands-busy applications, where the cost of learning the required eye movements ("stare at this icon to activate the device") is repaid by the acquisition of an otherwise impossible new ability. However, we believe that the real benefits of eye movement interaction for the majority of users will be in its naturalness, fluidity, low cognitive load, and almost unconscious operation. These benefits are attenuated if unnatural, and thus quite conscious, eye movements are required.

Unnatural eye movement/Natural response: The remaining category created by this taxonomy is anomalous and not seen in practice.

Current Research in Eye Tracking for HCI Input

There is a variety of ways that eye movements can be used in user interfaces and the chapter in this section by O'Sullivan, Dingliana, and Howlett describes a range of them, including some using retrospective analysis for generating better interactive displays. The work they describe uses eye movement input in order to simulate a better interactive graphic display — where better might mean higher resolution, faster update rate, or more accurate simulation. We could say that, if techniques like this work perfectly, the user would not realize that eye tracking is in use, but would simply believe he or she were viewing a better graphic display than could otherwise be built; the eye movement input ought to be invisible to the user. This is in contrast to other work in eye movement-based interaction, such as that described in the chapter by Illi, Isokoski and Surakka, where the eye specifically provides input to the dialogue and may actuate commands directly.

Gaze-based rendering can be applied in several ways: One would be to make a higher resolution display at the point where the user's fovea is pointed. We know that the eye can see with much higher resolution near the fovea than in the periphery, so a uniform, high-resolution display "wastes" many pixels that the user can't see. However, this requires a hardware device that can modify its pixel density in real time.

Some work has been done along these lines for flight simulators, to create the illusion of a large, ultra-high resolution display in a projected dome display described above (Tong & Fisher, 1984). Two overlapping projectors are used, one covering the whole dome screen, and one covering only the high resolution inset; the second one is attached to a mirror on a servomechanism, so it can be rapidly moved around on the screen, to follow the user's eye movements.

Another way to use eye movements is to concentrate graphics rendering power in the foveal area. Here, a slower, but higher quality ray tracing algorithm generates the pixels near the fovea, and a faster one generates the rest of the screen, but the hardware pixel density of the screen is not altered (Levoy and Whitaker, 1990). However, as graphics processors improve, this technique may become less valuable, and it continues to be limited by the hardware pixel density of the screen. A similar approach can also be used, but applied to the simulation rather than the rendering. Even with fast graphics processors and fixed pixel density, the calculations required to simulate three-dimensional worlds can be arbitrarily complex and time consuming. This approach focuses higher quality simulation calculations, rather than rendering, at the foveal region. Another approach described in the chapter by O'Sullivan, Dingliana, and Howlett uses known performance characteristics of the eye, rather than real time eye tracking, to inform the design of the interface. By knowing when a saccade is likely to occur and that the visual system nearly shuts down during saccades, the system can use that time period to change the level of detail of a displayed object so that the change will not be noticed. Finally, the chapter presents ways to use retrospective analysis of eye movements to build a better interface. The authors analyze a user's eye movements viewing an object and use that information to determine the key perceptual points of that object. They can then develop a graphical model that captures higher resolution mesh detail at those points and less detail at others, so that it can be rendered rapidly but still preserve the most important details.

One of the most interesting points raised here is the scheduler component, which keeps track of the processing time spent on computing the current simulation frame and interrupts the process when the allocated time has been exceeded. The authors state "The simulation process proceeds by using the highest resolution data that the collision handling system was able to compute in the allocated time." This is a novel aspect of algorithms for interactive graphics systems — the need for producing the best computation possible in a fixed time (typically, in time for the next video refresh), rather than producing a fixed computation as fast as possible. We too have found this particularly important in virtual reality interfaces and developed a constraint-based language for specifying such interaction at a high level, which allows an underlying runtime system to perform optimization, tradeoffs, and conversion into discrete steps as needed. This allows us to tailor the response speeds of different elements of the user interface within the available computing resources for each video frame (Jacob, Deligiannidis & Morrison, 1999; Deligiannidis & Jacob, 2002).

We turn now to eye movement-based interaction in which the eye input is a first-class partner in the dialogue, rather than being used to improve the display quality but not affect the dialogue. The chapter by Illi, Isokoski and Surakka describes a variety of real time, interactive uses of eye movements in user interfaces. They provide a

survey of much work in this area and lay out the problems and issues in using eye movements as an input medium.

As we have seen, the Midas Touch or clutch problem plagues the design of eye movement-based interfaces — how can we tell which eye movements the system should respond to and which it should ignore? Illi, Isokoski, and Surakka describe an approach that uses additional physiological measurements from the user to help solve this problem. By measuring other lightweight inputs from the user and combining the information from the eye with the information from these additional sensors, the system can form a better picture of the user's intentions. Here, facial muscle activity is measured in real time with EMG sensors. Eye movements, like EMG and other such measurements, share the advantage that they are easy for the user to generate — they happen nearly unconsciously. But they share the drawback that they are difficult to measure and to interpret. By combining several such imperfect inputs, we may be able to draw a better conclusion about the user's intentions and make a more appropriate response, just as one combines several imperfect sensor observations in the physical world to obtain better information than any of the individual sensors could have yielded alone.

Other recent researchers have also found ways to advance the use of eye movements in user interfaces. The MAGIC approach of Zhai, Morimoto and Ihde (1999) is carefully tailored to exploit the characteristics of the eye and the hand and combine their use to provide good performance. Salvucci and Anderson show a way to improve performance and ease of use by adding intelligence to the underlying software. Their system favors more likely targets to compensate for some of the inaccuracy of the eye tracker (Salvucci, 2000). The chapter by Illi, Isokoski, and Surakka describes both of these techniques in further detail. Other researchers are beginning to study the eye movements of automobile drivers, ultimately leading to a lightweight style interface that might combine eye movements with other sensors to anticipate a driver's actions and needs. Selker, Lockerd and Martinez (2001) have developed a new device that detects only the *amount* of eye motion. While this gives less information than a conventional eye tracker, the device is extremely compact and inexpensive and could be used in many situations where a conventional eye tracker would not be feasible.

Most eye movement-based interfaces respond to the user's instantaneous eye position. We can extend this to more subtle or lightweight interaction techniques by replacing "interaction by staring at" with a "interaction by looking around." An additional benefit of this approach is that, because the interaction is based on recent history rather than each instantaneous eye movement, it is more tolerant of the very brief failures of the eye tracker that we often observe. Our approach here is for the computer to respond to the user's glances with continuous, gradual changes. Imagine a histogram that represents the accumulation of eye fixations on each possible target object in an environment. As the user keeps looking at an object, histogram value of the object increases steadily, while histogram values of all other objects slowly decrease. At any moment we thus have a profile of the user's "recent interest" in the various displayed objects. We respond to those histogram values by allowing the user to select and examine the objects of interest. When the user shows interest in an object by looking at it, the system responds by enlarging the object, fading its surface color out to expose its internals, and hence selecting it. When the user looks away from the

object, the program gradually zooms the object out, restores its initial color, and hence deselects it. The program uses the histogram values to calculate factors for zooming and fading continuously (Tanriverdi & Jacob, 2000).

We also see a natural marriage between eye tracking and virtual reality (VR), both for practical hardware reasons and because of the larger distances to be traversed in VR compared to a desktop display. The user is already wearing a head-mounted display; adding a tiny head-mounted eye tracking camera and illuminator adds little to the bulk or weight of the device. Moreover, in VR, the head, eye tracker, and display itself all move together, so the head orientation information is not needed to determine line of gaze (it is of course used to drive the VR display). Objects displayed in a virtual world are often beyond the reach of the user's arm or the range of a short walk. We have observed that eye movement interaction is typically faster than interaction with a mouse. More significantly for this purpose, we have also found that the time required to move the eye is hardly related to the distance to be moved, unlike most other input devices (Sibert & Jacob, 2000). This suggests that eye gaze interaction is most beneficial when users need to interact with distant objects, as is often the case in a virtual environment. We have indeed observed this benefit in a comparison between eye and hand interaction in VR (Tanriverdi & Jacob, 2000), and this may be a promising direction for further use of eye movement interaction.

Future Directions of Eye Tracking for HCI Input

We have seen that eye movements in HCI have been studied for many years. They continue to appear to be a promising approach, but we do not yet see widespread use of eye movement interfaces or widespread adoption of eye trackers in the marketplace. We should remember that there has historically been a long time lag between invention and widespread use of new input or output technologies. Consider the mouse, one of the more successful innovations in input devices, first developed around 1968 (Engelbart & English, 1968). It took approximately ten years before it was found even in many other research labs; and perhaps twenty before it was widely used in applications outside the research world. And the mouse was based on simple, mechanical principles that were well understood from the start, rather than experimental computer vision algorithms and exotic video processing hardware.

When Jacob (first author of this chapter) and his colleagues started work on this issue in 1988 at the Naval Research Laboratories, they saw two main problems. One was better and less expensive eye tracking hardware, and the other was new interaction techniques and ways to use eye movements in interface design. They began working on solving the second problem, intending that the eye tracker industry would advance the first. They used existing large, clumsy, and expensive eye trackers as a test bed to study interfaces that might someday run on convenient, inexpensive, and ubiquitous new eye trackers. The goal was to develop the techniques that people would use by the time this new equipment appeared. Unfortunately, we are still waiting for the new equipment. Eye trackers are continuing to improve, but rather slowly.

This is partly due to the nature of the eye tracker industry. It consists of small companies without large capital investment. They might sell only tens or hundreds of eye trackers in a year. This makes it difficult to invest in the large engineering effort that would be required to develop a really good, inexpensive unit. But without such a unit, the market will continue to be limited to tens or hundreds per year — a "chicken and egg" problem. The cycle may at last be breaking. The basic hardware components of an eye tracker are a video camera, frame grabber, and a processor capable of analyzing the video in real time. The first two components are not only becoming quite inexpensive but are often incorporated into ordinary computer workstations, mainly for use in teleconferencing. Small video cameras and moderate quality frame grabbers are beginning to appear as standard components of desktop workstations. While they are intended for teleconferencing, they could also be useful for eye tracking. Further, current CPU chips can perform the needed processing directly, unlike earlier systems that required dedicated analogue electronics for the necessary speed.

There are still a few wrinkles. We need a frame buffer with fairly high resolution and pixel depth (though it need not be color). A more difficult problem is that the camera must be focused tightly on the eye; ideally the eye should fill the video frame, in order to get enough resolution to distinguish just where the pupil is pointing. This is usually solved with a servo mechanism to point the camera or else a chinrest to hold the user steady. A camera with wide field of view and very high resolution over the entire field might also solve this problem someday. The final requirement is a small infrared light aimed at the eye to provide the corneal reflection. This is not difficult to provide, but is clearly not a default component of current camera-equipped desktop workstations.

The necessary accuracy of an eye tracker that is useful in a real-time interface (as opposed to the more stringent requirements for basic eye movement research) is limited, since a user generally need not position his or her eye more accurately than about one degree to see an object sharply. The eye's normal jittering (microsaccades) and slow drift movements further limits the practical accuracy of eye tracking. It is possible to improve accuracy by averaging over a fixation, but not in a real-time inter-face. The accuracy of the best current eye trackers that can be used for these applications approaches one-degree useful limit. However, *stability* and *repeatability* of the measurements leave much to be desired. In a research study it is acceptable if the eye tracker fails very briefly from time to time; it may require that an experimental trial be discarded, but the user need not be aware of the problem. In an interactive interface, though, as soon as it begins to fail, the user can no longer rely on the fact that the computer dialogue is influenced by where his or her eye is pointing and will thus soon be tempted to retreat permanently to whatever backup input modes are available. While eye trackers have dropped somewhat in price, their performance in this regard has not improved significantly. Performance appears to be constrained less by fundamental limits, than simply by lack of effort in this narrow commercial market.

One of the most promising current hardware technologies is the IBM Blue Eyes device that combines bright pupil and dark pupil eye tracking in a single compact unit (Zhai *et al.*, 1999). One method will sometimes work better than the other for certain subjects or at a moment when there is an extra reflection or artifact in the image. By

rapidly toggling between these two modes, the device can use whichever gives better results from moment to moment. The Eye-R device of Selker *et al.* (2001) represents another possible direction in eye tracking hardware, toward an inexpensive, widely deployable device.

We can also step back and observe a progression in user interface devices that begins with experimental devices used to measure some physical attribute of a person in laboratory studies. As such devices become more robust, they may be used as practical medical instruments outside the laboratory. As they become convenient, non-invasive, and inexpensive, they may find use as future computer input devices. The eye tracker is such an example; other physiological monitoring devices may also follow this progression.

Conclusions

We have reviewed the progress of using eye tracking in human–computer interaction both retrospectively, for usability engineering and in real time, as a control medium within a Human–computer dialogue. We primarily discussed these two areas separately, but we also showed that they share the same principal challenges with eye tracking technology and interpretation of the resulting data. These two areas intersect in software applications where fixations are both recorded for analysis and the display is changed contingent on the locus of a user's fixation. This sort of technology was pioneered for the study of reading (for a thorough review and current methodological developments in this area see Rayner, 1998; Kennedy, Radach, Heller & Pynte, 2000) and has more recently been applied to more complex visual displays (e.g., McConkie, 1991; McConkie & Currie, 1996; Karn & Hayhoe, 2000) and to virtual environments (Triesch, Sullivan, Hayhoe & Ballard, 2002). Hyrskykari, Majaranta, Aaltonen and Räihä (2000) report a practical example of such an application of gaze dependent displays in human–computer interaction. While a user reads text in a non-native language, the system detects reader difficulties (by monitoring for long fixations and regressions) and then pops up translation help automatically.

Although progress has been slow, the concept of using eye tracking in human–computer interaction is clearly beginning to blossom. New work, described in this section and elsewhere in this volume, provides examples of this growing field. We can see this growth in the establishment of a new conference series covering this area, the Eye Tracking Research and Applications Symposium (ETRA) sponsored by the Association for Computing Machinery (ACM). The new field of perceptual or perceptive user interfaces (PUI) also brings together a variety of work on lightweight or sensing interfaces that, like eye tracking, observe the user's actions and respond. It, too, has established a conference series (PUI), beginning in 1997. Finally, we can observe clear growth within the ACM Human Factors in Computing Systems Conference (CHI), which is the premier conference in the general field of human–computer interaction. The first paper at CHI on eye tracking in HCI appeared in 1987 (Ware & Mikaelian, 1987), two more in 1990 (Starker & Bolt, 1990; Jacob, 1990); followed by one in 1996, three in 1998, four in 1999, and six or more in each of 2000, 2001 and 2002.

From the perspective of mainstream eye movement research, HCI, together with related work in the broader field of communications and media research, appears as a new and very promising area of applied work. Hence, at the last European Conference on Eye Movements in Turku, Finland in September 2001 a significant group of researchers from these areas participated and more were invited to contribute to the current volume. It is obvious that both basic and applied work can profit from integration within a unified field of eye movement research.

In this commentary chapter we have reviewed a variety of research on eye movements in human–computer interaction. The promise of this field was clear from the start, but progress has been slow to date. Does this mean the field is no longer worthwhile? We think not. Application of eye tracking in human–computer interaction remains a very promising approach. Its technological and market barriers are finally being reduced. Reports in the popular press (e.g., Kerber, 1999; Gramza, 2001) of research in this area indicate that the concepts are truly catching on. Work, in these chapters and elsewhere, demonstrates that this field is now beginning to flourish. Eye tracking technology has been used to make many promises in the past 50 years. However, we do believe the technology is maturing and has already delivered promising results. We are quite confident that the current generation of workers in this field will finally break the circle of boom and decline that we have seen in the past and make applications of eye movements in HCI an integrated part of modern information technology.

Acknowledgements

Notes

1 This ISI Citation search includes three indices (*Science Citation Index Expanded, Social Sciences Citation Index, and the Arts & Humanities Citation Index*) for the years 1973 to the present.
2 There have certainly been more than 16 studies incorporating eye tracking in usability research, but we use this citation index as a means of judging the relative popularity of these two techniques that Paul Fitts left as his legacy.
3 The problem of defining and computing eye movement parameters has recently been the subject of intensive methodological debate in the neigbouring area of reading research (see Inhoff & Radach, 1998 and Inhoff & Weger, this volume, for detailed discussions). The

chapters by Vonk and Cozijn and by Hyönä, Lorch and Rinck in the section on empirical research on reading deal with specific problems of data aggregation that to a certain degree also generalize to the area of usability research.

4 The list provided in Table 1 is not a complete list of all applications of eye tracking in usability studies, but it provides a good sense of how these types of studies have evolved over these past 50 years.

5 Some other authors use "gaze duration" differently, to refer to the *total* time fixating an area of interest during an entire experimental trial (i.e., the sum of all individual gaze durations).

6 "Dwell" is still arguably a more convenient word and time will tell whether "dwell" or "gaze" becomes the more common term.

7 In their comprehensive theoretical discussion on saccadic eye movments, Findlay and Walker (1999) distinguish between three levels of saccade generation: automatic, automated, and voluntary. The least well known of these levels is the automated one, representing the very frequent class of saccades made on the basis of learned oculomotor routines. On all three levels the generation of eye movements is mediated by bottom up and top down processing, but only a rather small a minority of eye movmements appears to be made on on a "voluntary" basis.

References

Albert, W. (2002). Do web users actually look at ads? A case study of banner ads and eye-tracking technology. In: *Proceedings the 11th Annual Conference of the Usability Professionals' Association*, Florida.

Albert, W. S., & Liu, A. (in press). The effects of map orientation and landmarks on visual attention while using an in-vehicle navigation system. To appear in: A. G. Gale (ed.), *Vision in Vehicles 8*. London: Oxford Press.

Aaltonen, A. Hyrskykari, A., & Räihä, K. (1998). 101 Spots, or how do users read menus? In: *Proceedings of CHI 98 Human Factors in Computing Systems* (pp. 132–139). ACM Press.

Andrews, T., & Coppola, D. (1999). Idiosyncratic characteristics of saccadic eye movements when viewing different visual environments. *Vision Research*, *39*, 2947–2953.

Anliker, J. (1976). Eye movements: On-line measurement, analysis, and control. In: R. S. Monty, and J. W. Senders (eds), *Eye Movements and Psychological Processes* (pp. 185–199). Hillsdale, NJ: Lawrence Erlbaum Associates.

Babcock, J., Lipps, M., & Pelz, J. B. (2002). How people look at pictures before, during, and after image capture: Buswell revisited. In: *Proceedings of SPIE, Human Vision and Electronic Imaging*, *4662* (pp. 34–47).

Backs, R. W., & Walrath, L. C. (1992). Eye movement and pupillary response indices of mental workload during visual search of symbolic displays. *Applied Ergonomics*, *23*, 243–254.

Bolt, R. A. (1981). Gaze-orchestrated dynamic windows. *Computer Graphics*, *15*, 109–119.

Bolt, R. A. (1982). Eyes at the interface. In: *Proceedings of the ACM Human Factors in Computer Systems Conference* (pp. 360–362).

Benel, D. C. R., Ottens, D., & Horst, R. (1991). Use of an eye tracking system in the usability laboratory. In: *Proceedings of the Human Factors Society 35th Annual Meeting* (pp. 461–465). Santa Monica: Human Factors and Ergonomics Society.

Brookings, J. B., Wilson, G. F., & Swain, C. R. (1996). Psychophysiological responses to changes in workload during simulated air traffic control. *Biological Psychology*, 42, 361–377.

Byrne, M. D., Anderson, J. R., Douglas, S., & Matessa, M. (1999). Eye tracking the visual search of click-down menus. In: *Proceedings of CHI 99* (pp. 402–409). NY: ACM Press.

Card, S. K. (1984). Visual search of computer command menus. In: H. Bouma and D. G. Bouwhuis (eds), *Attention and Performance X, Control of Language Processes*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Collewijn, H. (1999). Eye movement recording. In: R. H. S. Carpenter and J. G. Robson (eds.), *Vision Research: A Practical Guide to Laboratory Methods* (pp. 245–285). Oxford: Oxford University Press.

Cornsweet & Crane (1973). Accurate two-dimensional eye tracker using first and fourth Purkinje images. *Journal of the Optical Society of America*, *63*, 921–928.

Cowen, L. (2001). An eye movement analysis of web-page usability. Unpublished Masters' thesis, Lancaster University, UK.

Crosby, M. E., & Peterson, W. W. (1991). Using eye movements to classify search strategies. In: *Proceedings of the Human Factors Society 35th Annual Meeting* (pp. 1476–1480). Santa Monica: Human Factors and Ergonomics Society.

Crowe, E. C., & Narayanan, N. H. (2000). Comparing interfaces based on what users watch and do. In: *Proceedings Eye Tracking Research and Applications Symposium 2000* (pp. 29–36). New York: Association for Computing Machinery.

Deligiannidis, L., & Jacob, R. J. K. (2002). DLoVe: Using Constraints to Allow Parallel Processing in Multi-User Virtual Reality. In: *Proceedings of the IEEE Virtual Reality 2002 Conference*, IEEE Computer Society Press. (Available at http://www.cs.tufts.edu/~jacob/papers/vr02.deligiannidis.pdf).

Dodge, R., & Cline, T. S. (1901). The angle velocity of eye movements. *Psychological Review*, *8*, 145–157.

Duchowski, A. T. (2003). *Eye Tracking Methodology: Theory and Practice*. London: Springer-Verlag.

Ellis, S., Candrea, R., Misner, J., Craig, C. S., Lankford, C. P., & Hutshinson, T. E. (1998). Windows to the soul? What eye movements tell us about software usability. In: *Proceedings of the Usability Professionals' Association Conference 1998* (pp. 151–178).

Engelbart, D. C., & English, W. K. (1968). A research center for augmenting human intellect. In: *Proceedings of the 1968 Fall Joint Computer Conference* (pp. 395–410). AFIPS.

Findlay, J. M., & Walker, R. (1999). A model of saccade generation based on parallel processing and competitive inhibition. *Behavioral & Brain Sciences 22*(4), 661–721.

Fisher, D. F., Monty, R. A., & Senders, J. W. (eds) (1981). *Eye Movements: Cognition and Visual Perception.* Hillsdale, NJ: Lawrence Erlbaum.

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, *47*, 381–391.

Fitts, P. M., Jones, R. E., & Milton, J. L. (1950). Eye movements of aircraft pilots during instrument-landing approaches. *Aeronautical Engineering Review 9*(2), 24–29.

Flemisch F. O., & Onken, R. (2000). Detecting usability problems with eye tracking in airborne battle management support. In: *Proceedings of the NATO RTO HFM Symposium on Usability of information in Battle Management Operations* (pp. 1–13). Oslo.

Frøkjær, E., Hertzum, M., & Hornbæk, K. (2000). Measuring usability: Are effectiveness, efficiency and satisfaction really correlated? In: *Proceedings of CHI 2000 Human Factors in Computing Systems* (pp. 345–352). ACM Press.

Glenn, F., Iavecchia, H., Ross, L., Stokes, J., Weiland, W., Weiss, D., & Zaklad, A. (1986). Eye-voice-controlled interface. In: *Proceedings of the 30th Annual Meeting of the Human Factors Society* (pp. 322–326). Santa Monica: Human Factors Society.

Goldberg, J. H., & Kotval, X. P. (1998). Eye movement-based evaluation of the computer interface. In: S. K. Kumar (ed.), *Advances in Occupational Ergonomics and Safety* (pp. 529–532). Amsterdam: ISO Press.

Goldberg, J. H., Stimson, M. J., Lewenstein, M. Scott, N., & Wichansky, A. M. (2002). Eye tracking in web search tasks: Design implications. In: *Proceedings of the Eye Tracking Research & Applications Symposium 2002* (pp. 51–58). New York ACM.

Graf, W., & Krueger, H. (1989). Ergonomic evaluation of user-interfaces by means of eye-movement data. In: M. J. Smith and G. Salvendy (eds), *Work with Computers: Organizational, Management, Stress and Health Aspects* (pp. 659–665). Amsterdam: Elsevier Science.

Gramza, J. (2001). What are you looking at? *Popular Science*, March, 54–56.

Harris, R. L., & Christhilf, D. M. (1980). What do pilots see in displays? In: *Proceedings of the Human Factors Society — 24th Annual Meeting* (pp. 22–26). Los Angeles: Human Factors Society.

Hartridge, H., & Thompson, L. C. (1948). Methods of investigating eye movements, *British Journal of Ophthalmology*, *32*, 581–591.

Hegarty, M. (1992). The mechanics of comprehension and comprehension of mechanics. In: K. Rayner (ed.), *Eye Movements and Visual Cognition: Scene Perception and Reading.* New York: Springer Verlag.

Hegarty, M., & Just, M.A. (1993). Constructing mental models of machines from text and diagrams. *Journal of Memory and Language*, *32*, 717–742.

Hendrickson, J. J. (1989). Performance, preference, and visual scan patterns on a menu-based system: Implications for interface design. In: *Proceedings of the ACM CHI'89 Human Factors in Computing Systems Conference* (pp. 217–222). ACM Press.

Hoeks, B., & Levelt, W. J. M. (1993). Pupillary dilation as a measure of attention: A quantitative system analysis. *Behavior Research Methods, Instruments & Computers*, *25*, 16–26.

Hutchinson, T. E., White, K. P., Martin, W. N., Reichert, K. C., & Frey, L. A. (1989). Human–computer interaction using eye-gaze input. *IEEE Transactions on Systems, Man, and Cybernetics*, *19*, 1527–1534.

Hyrskykari, A., Majaranta, P., Aaltonen, A., & Räihä, K. (2000). Design issues of iDict: A gaze-assisted translation aid. In: *Proceedings of the Eye Tracking Research and Applications Symposium 2000* (pp. 9–14). NY: ACM Press.

Iida, M. Tomono, A., & Kobayashi, Y. (1989). A study of human interface using and eye-movement detection system. In: M. J. Smith and G. Salvendy (eds), *Work with Computers: Organizational, Management, Stress and Health Aspects* (pp. 666–673). Amsterdam: Elsevier Science.

Inhoff, A. W., & Radach, R. (1998). Definition and computation of oculomotor measures in the study of cognitive processes. In: G. Underwood (ed.), *Eye Guidance in Reading and Scene Perception*, (pp. 29–53). Oxford: Elsevier

Jacob, R. J. K. (1990). What you look at is what you get: Eye movement-based interaction techniques. In: *Proceedings of the ACM CHI'90 Human Factors in Computing Systems Conference* (pp. 11–18). Addison-Wesley/ACM Press.

Jacob, R. J. K. (1991). The use of eye movements in human–computer interaction techniques: What you look at is what you get. *ACM Transactions on Information Systems*, *9*, 152–169.

Jacob, R. J. K., Deligiannidis, L., & Morrison, S. (1999). A software model and specification language for non-wimp user interfaces. *ACM Transactions on Computer–Human Interaction*, *6*(1), 1–46. (Available at http://www.cs.tufts.edu/~jacob/papers/tochi.pmiw.html [HTML] or http://www.cs.tufts.edu/~jacob/papers/tochi.pmiw.pdf [PDF]).

Javal, E. (1878). Essai sur la physiologie de la lecture. *Annales d'Oculistique*, *79*, 97–117, 155–167, 240–274; *80* (1879), 61–73, 72–81, 157–162, 159–170, 242–253.

Josephson, S., & Holmes, M. E. (2002). Visual attention to repeated Internet images: Testing the scanpath theory on the world wide web. In: *Proceedings of the Eye Tracking Research & Applications Symposium 2002* (pp. 43–49). New York: ACM.

Judd, C. H., McAllister, C. N., & Steel, W. M. (1905). General introduction to a series of studies of eye movements by means of kinetoscopic photographs. In: J. M. Baldwin, H. C. Warren and C. H. Judd (eds), *Psychological Review, Monograph Supplements*, *7*, 1–16. Baltimore: The Review Publishing Company.

Just, M. A., & Carpenter, P. A. (1976a). Eye Fixations and Cognitive Processes. *Cognitive Psychology*, *8*, 441–480.

Just, M. A., & Carpenter, P. A. (1976b). The role of eye-fixation research in cognitive psychology. *Behavior Research Methods & Instrumentation*, *8*, 139–143.

Karn, K., Ellis, S., & Juliano, C. (1999). The hunt for usability. Workshop conducted at CHI 99 Human Factors in Computing Systems, Conference of the Computer–Human Interaction Special Interest Group of the Association of Computing Machinery. Pittsburgh. In: *CHI 99 Extended Abstracts*. (p. 173). NY: ACM Press.

Karn, K., Ellis, S., & Juliano, C. (2000). The hunt for usability: Tracking eye movements. *SIGCHI Bulletin*, November/December 2000 (p. 11). New York: Association for Computing Machinery. (Available at http://www.acm.org/sigchi/bulletin/2000.5/eye.html).

Karn, K., Goldberg, J., McConkie, G., Rojna, W., Salvucci, D., Senders, J., Vertegaal, R., & Wooding, D. (2000). "Saccade Pickers" vs. "Fixation Pickers": The effect of eye tracking instrumentation on research. (Panel presentation). Abstract in: *Proceedings of the Eye Tracking Research and Applications Symposium 2000* (p. 87). NY: ACM Press.

Karn, K., & Hayhoe, M. (2000). Memory representations guide targeting eye movements in a natural task. *Visual Cognition*, *7*, 673–703.

Karn, K., Krolczyk, M., & Perry, T. (1997). Testing for power usability. Workshop conducted at CHI 97 Human Factors in Computing Systems. Conference of the Computer–Human Interaction Special Interest Group of the Association of Computing Machinery. Atlanta. In: *CHI 97 Extended Abstracts*. (p. 235). NY: ACM Press.

Karsh, R., & Breitenbach, F. W. (1983). Looking at looking: The amorphous fixation measure. In: R. Groner, C. Menz, D. Fisher and R. A. Monty (eds), *Eye Movements and Psychological Functions: International Views* (pp. 53–64). Hillsdale, NJ: Lawrence Erlbaum Associates.

Kennedy, A., Radach, R., Heller, D., & Pynte, J. (eds, 2000). *Reading as a Perceptual Process*. Oxford: Elsevier.

Kerber, R. (1999). Cleanup crew has eye on web sites. *The Boston Globe*, August 22, 1999. A1, A16–17.

Kolers, P. A., Duchnicky, R. L., & Ferguson, D. C. (1981). Eye movement measurement of readability of CRT displays. *Human Factors*, *23*, 517–527.

Kotval, X. P., & Goldberg, J. H. (1998). Eye movements and interface components grouping: An evaluation method. In: *Proceedings of the 42nd Annual Meeting of the Human Factors and Ergonomics Society* (pp. 486–490). Santa Monica: Human Factors and Ergonomics Society.

Kowler, E. (1990). The role of visual and cognitive processes in the control of eye movement. In: E. Kowler (ed.), *Eye Movements and their Role in Visual and Cognitive Processes*. Amsterdam: Elsevier Science.

Lambert, R. H., Monty, R. A., & Hall, R. J. (1974). High-speed data processing and unobtrusive monitoring of eye movements. *Behavioral Research Methods & Instrumentation*, *6*, 525–530.

Lankford, C. (2000) Gazetracker™: Software designed to facilitate eye movement analysis. In: *Proceedings of the Eye Tracking Research and Applications Symposium 2000* (pp. 51–55). NY: ACM Press.

Land, M.F. (1992). Predictable eye-head coordination during driving. *Nature*, *359*, 318–320.

Land, M. F., Mennie, N., & Rusted, J. (1999). The role of vision and eye movements in the control of activities of daily living. *Perception*, *28*, 1311–1328.

Levine, J. L. (1981). *An Eye-Controlled Computer*. Research Report RC-8857. New York: IBM Thomas J. Watson Research Center, Yorktown Heights.

Levine, J. L. (1984). Performance of an eyetracker for office use. *Computational Biology and Medicine*, 14, 77–89.

Levoy, M., & Whitaker, R. (1990). Gaze-directed volume rendering. In: *Proceedings of the 1990 Symposium on Interactive 3D Graphics* (pp. 217–223). Utah: Snowbird.

Mackworth, J. F., & Mackworth, N. H. (1958). Eye fixations recorded on changing visual scenes by the television eye-marker. *Journal of the Optical Society of America*, *48*, 439–445.

Mackworth, N. H., & Thomas, E. L. (1962). Head-mounted eye-marker camera. *Journal of the Optical Society of America*, *52*, 713–716.

Marshall, S. P. (1998). Cognitive workload and point of gaze: A re-analysis of the DSS directed-question data. *Technical Report CERF 98–03*. San Diego, CA: Cognitive Ergonomics Research Facility, San Diego State University.

McConkie, G. W. (1991). Perceiving a stable visual world. In: J. Van Rensbergen, M. Devijver and G. d'Ydewalle (eds), *Proceedings of the Sixth European Conference on Eye Movements*. (pp. 5–7). Leuven, Belgium: Laboratorium voor Experimental Psychologie, Katholieke Universiteit Leuven.

McConkie, G. W., & Currie, C. B. (1996). Visual stability across saccades while viewing complex pictures. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 563–581.

Merchant, J. Morrissette, R., & Porterfield, J. L. (1974). Remote measurement of eye direction allowing subject motion over one cubic foot of space. *IEEE Transactions on Biomedical Engineering*, BME-21, 309–317.

Merwin, D. (2002). Bridging the gap between research and practice. *User Experience,* Winter, 38–40.

Monty, R. A. (1975). An advanced eye-movement measuring and recording system. *American Psychologist*, *30*, 331–335.

Monty, R. A., & Senders, J. W. (eds) (1976). *Eye Movements and Psychological Processes.* Hillsdale, NJ: Lawrence Erlbaum.

Mulligan, J. (2002). A software-based eye tracking system for the study of air-traffic displays. In: *Proceedings of the Eye Tracking Research and Applications Symposium 2002* (pp. 69–76). New York: Association for Computing Machinery.

Pelz, J. B., & Canosa, R., (2001). Oculomotor behavior and perceptual strategies in complex tasks. *Vision Research*, *41*, 3587–3596.

Pelz, J. B., Canosa, R., & Babcock, J. (2000). Extended tasks elicit complex eye movement patterns. In: *Proceedings of the Eye Tracking Research and Applications Symposium 2000*. (pp. 37–43) New York: ACM Press.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*, 372–422.

Redline, C. D., & Lankford, C. P. (2001). Eye-movement analysis: A new tool for evaluating the design of visually administered instruments (paper and web). Paper presented at 2001

AAPOR Annual Conference, Montreal, Quebec, Canada, May 2001. In: *Proceedings of the Section on Survey Research Methods*, American Statistical Association.

Reeder, R. W., Pirolli, P., & Card, S. K. (2001). WebEyeMapper and WebLogger: Tools for analyzing eye tracking data collected in web-use studies. In: *Extended Abstracts of the Conference on Human Factors in Computing Systems, CHI 2001* (pp. 19–20). New York: ACM Press.

Russo, J. E., & Leclerc, F. (1994). An eye-fixation analysis of choice process for consumer nondurables. *Journal of Consumer Research*, *21*, 274–290.

Salvucci, D. D. (2000). An interactive model-based environment for eye-movement protocol analysis and visualization. In: *Proceedings of the Eye Tracking Research and Applications Symposium 2000* (pp. 57–63). NY: ACM Press.

Salvucci, D. D., & Anderson, J. R. (2000). Intelligent gaze-added interfaces. In: *Proceedings of the ACM CHI 2000 Human Factors in Computing Systems Conference* (pp. 273–280). Addison-Wesley/ACM Press.

Salvucci, D. D., & Goldberg, J. H. (2000). Identifying fixations and saccades in eye-tracking protocols. In: *Proceedings of the Eye Tracking Research and Applications Symposium 2000* (pp. 71–78). NY: ACM Press.

Selker, T., Lockerd, A., & Martinez, J. (2001). Eye-R, a glasses-mounted eye motion detection interface. In: *Proceedings of the ACM CHI 2001 Human Factors in Computing Systems Conference Extended Abstracts* (pp. 179–180). NY: ACM Press.

Senders, J. W. (2000). Four theoretical and practical questions. Keynote address presented at the Eye Tracking Research and Applications Symposium 2000. Abstract in: *Proceedings of the Eye Tracking Research and Applications Symposium 2000* (p. 8). New York: Association for Computing Machinery.

Senders, J. W., Fisher, D. F., & Monty, R. A. (eds) (1978). *Eye Movements and the Higher Psychological Functions.* Hillsdale, NJ: Lawrence Erlbaum.

Shackel, B. (1960). Note on mobile eye viewpoint recording. *Journal of the Optical Society of America*, *59*, 763–768.

Sheena, D., & Flagg, B. N. (1978). Semiautomatic eye movement data analysis techniques for experiments with varying scenes. In: J. W. Senders, D. F. Fisher and R. A. Monty (eds), *Eye Movements and the Higher Psychological Functions* (pp. 65–75). Hillsdale, NJ: Lawrence Erlbaum.

Sibert, L. E., & Jacob, R. J. K. (2000). Evaluation of eye gaze interaction. In: *Proceedings of the ACM CHI 2000 Human Factors in Computing Systems Conference* (pp. 281–288). Addison-Wesley/ACM Press. (Available at http://www.cs.tufts.edu/~jacob/papers/chi00.sibert.pdf [PDF]).

Simmons, R. R. (1979). Methodological considerations of visual workloads of helicopter pilots. *Human Factors*, *21*, 353–367.

Sodhi, M., Reimer, B., Cohen, J. L., Vastenburg, E., Kaars, R., & Kirschenbaum, S. (2002). On-road driver eye movement tracking using head-mounted devices. In: *Proceedings of the Eye Tracking Research & Applications Symposium 2002* (pp. 61–68). New York: ACM.

Starker, I., & Bolt, R. A. (1990). A gaze-responsive self-disclosing display. In: *Proceedings of the ACM CHI'90 Human Factors in Computing Systems Conference* (pp. 3–9). Addison-Wesley/ACM Press.

Stern, J. A., Boyer, D., & Schroeder, D. (1994). Blink rate: A possible measure of fatigue. *Human Factors*, *36*, 285–297.

Stolk, H. Boon, K., & Smulders, M. (1993). Visual information processing in a study task using text and pictures. In: G. d'Ydewalle and J. Van Rensbergen (eds), *Perception and Cognition* (pp. 285–296). Amsterdam: Elsevier Science.

Svensson, E., Angelborg-Thanderz, M. Sjöeberg, L., & Olsson, S. (1997). Information complexity: Mental workload and performance in combat aircraft. *Ergonomics*, *40*, 362–380.

Tanriverdi, V., & Jacob, R. J. K. (2000). Interacting with eye movements in virtual environ-ments. In: *Proceedings of the ACM CHI 2000 Human Factors in Computing Systems Conference* (pp. 265–272). Addison-Wesley/ACM Press. (Available at http://www.cs.tufts.edu/~jacob/papers/chi00.tanriverdi.pdf [PDF]).

Tinker, M. A. (1963). *Legibility of Print*. Ames, Iowa: Iowa State University Press.

Tong, H. M., & Fisher, R. A. (1984). *Progress Report on an Eye-Slaved Area-of-interest Visual Display*. Report No. AFHRL-TR-84–36, Air Force Human Resources Laboratory, Brooks Air Force Base, Texas. Proceedings of IMAGE III Conference.

Triesch, J., Sullivan, B. T., Hayhoe, M. M., & Ballard, D. H. (2002). Saccade contingent updating in virtual reality. In: *Proceedings of the Eye Tracking Research & Applications Symposium 2002* (pp. 95–102). New York: ACM.

Vertegaal, R. (1999). The GAZE groupware system: Mediating joint attention in multiparty communication and collaboration. In: *Proceedings of the ACM CHI'99 Human Factors in Computing Systems Conference* (pp. 294–301). Addison-Wesley/ACM Press.

Ware, C., & Mikaelian, H. T. (1987). An evaluation of an eye tracker as a device for computer input. In: *Proceedings of the ACM CHI+GI'87 Human Factors in Computing Systems Con-ference* (pp. 183–188). New York: ACM Press.

Weiser, M. (1993). Some computer science issues in ubiquitous computing. *Communications of the ACM 36*(7), 75–84.

Wooding, D. S. (2002). Fixation Maps: Quantifying eye-movement traces. In: *Proceedings of the Eye Tracking Research & Applications Symposium 2002* (pp. 31–36). New York: ACM.

Yarbus, A. L. (1967) *Eye Movements and Vision* (Trans. B. Haigh). New York: Plenum Press. (Original work published 1965).

Yamamoto, S., & Kuto, Y. (1992). A method of evaluating VDT screen layout by eye move-ment analysis. *Ergonomics*, *35*, 591–606.

Zhai, S., Morimoto, C., & Ihde, S. (1999). Manual and gaze input cascaded (MAGIC) pointing. In: *Proceedings of the ACM CHI'99 Human Factors in Computing Systems Conference* (pp. 246–253). Addison-Wesley/ACM Press.