Commentary on Section 4. Eye tracking in human-computer interaction and usability research: Ready to deliver the promises.

Robert J.K. Jacob, Ph.D.
jacob@cs.tufts.edu
Department of Computer Science
Tufts University, 161 College Avenue, Medford, Mass. 02155, USA

Keith S. Karn, Ph.D.
keith.karn@usa.xerox.com
Xerox Corporation, Industrial Design / Human Interface Department
1350 Jefferson Road, Mail Stop 0801-10C, Rochester, NY 14623, USA
    and
Center for Visual Science / Department of Brain and Cognitive Sciences
University of Rochester, Rochester, NY, USA

Introduction

This section considers the application of eye movements to user interfaces—both for analyzing interfaces, measuring usability, and gaining insight into human performance—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 & Steel (1905) applied motion picture photography to record the temporal aspects of eye movements in two dimensions. Their technique 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 & 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" (p. 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 U.S. 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 and 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 (U.S. Airforce / Honeywell Corporation and U.S. 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 //2x// 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 70s 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, 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, 1989, Levine, 1984, Levine, 1981). 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, 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, 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; and 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 & 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 can be used as 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:

---

[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.

- "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).

- "…[T]he eyetracking system has a promising future in usability engineering" (Benel, Ottens & Horst, 1991, p.465).

- "…[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).

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

- "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).

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 70s 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 & 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 (in press).

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 probably 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 & Juliano (1999, 2000) and Goldberg & 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

---

[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 & Cozijn and by Hyönä, Lorch & 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.

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 & 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, Jones & Milton, 1950; Card, 1984; and 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 20 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."

- *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.

---

[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.

Table 1. Summary of 20 usability studies incorporating eye tracking.

Authors / Date	Users and Tasks	Eye Tracking Related Metrics
Fitts, Jones & Milton, 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 & 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 & 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 & 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 & Horst, 1991	7 PC users. Viewing web pages.	• Gaze % (proportion of time) on each area of interest • Scan path
Backs & 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 & Kuto, 1992	7 young adults. Confirm sales receipts (unit price, quantity, etc.) on various screen layouts.	• Scan path direction • Number of instances of backtracking
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 & 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 & 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 & 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 & Holmes, 2002	8 university students with web experience. Passively view web pages.	• Scan path
Goldberg, Stimson, Lewenstein, Scott & 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 & 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

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, Chapter TBD, 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 & Kotval (1998); Kotval & Goldberg, 1998; and Zülch & Stowasser (Chapter TBD in 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, Jones & Milton (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, Jones & Milton, 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, Jones & Milton, 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 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, Mennie & Rusted (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 & 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 & 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 & 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 scientists 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, Jones & Milton, 1950; Crosby & Peterson, 1991; Card, 1984; Altonen, et al., 1998) and individual differences (Yarbus, 1965/1967; Card, 1984; Andrews & Coppola, 1999). Eye tracking seems like an especially useful tool to study repetitive or well-practiced tasks and "power usability" (Karn, Perry & Krolczyk, 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 as discussed above. 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.