Interaction Science and the Aging User: Techniques to Assist in Design and Evaluation

Sandra P. Marshall

EyeTracking, Inc., San Diego, CA 92120
smarshall@eyetracking.com

Abstract. The aging user presents unique challenges to designers of new technologies, in part because of physical changes in the user's visual system. This paper reviews a number of these changes and shows how eye tracking can assist interaction science studies, both through traditional eye-movement metrics as well as with pupil-based estimates of cognitive workload. Three studies of older and younger participants are described.

Keywords: Aging, Cognitive Workload, Eye Tracking, Interface Design.

1 Introduction

Interaction science focuses on identifying ways that users interact with and respond to electronic interfaces. Not all users react in the same way to a particular interface, and designing an interface for a broad range of users is often difficult. In particular, building interfaces for older users may be especially challenging. The combination of technological unfamiliarity and physical impairment resulting from the normal aging process may create barriers that dissuade many people in older demographics from using new technologies successfully or perhaps at all. This problem is escalating as computer interactions become more prevalent in employment, education, healthcare, travel, and government settings. Many transactions can only be carried out online, and all demographic groups are interacting in greater numbers with social network sites, smartphone apps, and traditional websites.

Thus, it is important to evaluate continually the success of modern technology systems in supporting older users. Eye tracking is a particularly valuable methodology in interaction science. By measuring eye movements and pupil activity in different digital environments, we can objectively measure the extent to which today's websites, software and other computerized applications have bridged the 'digital divide.' This paper summarizes some of the problems of the aging eye and describes differences observed in younger and older users doing a variety of computer-based tasks.

2 Previous Research

Most research can be broken into two general categories: ophthalmologic research about the eye itself and psychopsychological research about visual perception and

C. Stephanidis and M. Antona (Eds.): UAHCI/HCII 2013, Part II, LNCS 8010, pp. 133–141, 2013.

performance. A large body of literature exists on both topics, with general (but not complete) agreement on the relationships involved.

2.1 Ophthalmologic Research

Many trends have been documented about the effects of aging on the eye. For example, Meisami and colleagues report a reduction in pupil size, the loss in size of visual field (~3.5%), and decreased light utilization by the retina [1].

Much is known about how vision changes with age. Acuity changes are perhaps the most important for individuals seeking to access the internet. Acuity loss with aging is well documented, but research is not consistent about the severity and rate of loss. Importantly, visual acuity appears to decline more slowly in high contrast, high luminance conditions than in low contrast, low luminance conditions [2]. The success of e-readers with older populations may be explained in part because of high contrast screens [3].

Other changes in vision that accompany aging are also important to the issues of universal access and interaction science. Haegerstrom-Portnoy et al. conducted a very large study with more than 1500 participants who were 65 and older [2]. The participants were given a broad battery of visual tests to examine how vision function changes with aging. These researchers report a number of significant findings:

- Glare sensitivity increases with aging, as does recovery from glare exposure
- Contrast sensitivity increases with aging
- Poor stereopsis is more prevalent in the elderly
- Color confusion increases in the very old, especially blue-yellow confusions
- The size of the field of attention shrinks with age
- Peripheral fields decrease with aging, especially under conditions of divided attention
- More attention 'holes' in central visual field are observed in the elderly

How letters and images are presented on a computer display will obviously impact the older user's ability to see and comprehend them. Lack of contrast, presence of high glare, and/or simultaneous display of multiple targets of attention may all negatively impact the user's experience.

2.2 Psychophysiological Research

Much of the psychophysiological research has come from studies using eye tracking. Emphasis here has been largely on how aging affects eye variables such as fixations, saccades, and smooth pursuit. In general, research on fixations shows little loss with aging. The aging eye appears able to fixate a target and to maintain a fixation as easily as the younger eye. Most of the differences observed have been on eye *movement*.

Not surprisingly, older people have slower saccadic reaction times than younger people. Cuiffreda & Tannen report that older individuals could have latencies up to 100 ms longer than younger people ([4], p. 52). Older individuals also show some reduced performance in smooth pursuit tasks (i.e., tasks where they are asked to follow a moving target). Typically, individuals who are engaged in smooth pursuit are able to maintain a relatively stable relationship between eye angular velocity and target angular velocity. This relationship is commonly called 'gain'. Research suggests that older individuals have a reduced gain in smooth pursuit tasks, perhaps as large as 25% [4]. Thus, they fall behind as they engage in smooth pursuit, causing them to show increased saccades (as they attempt to reacquire the target). There is also some evidence that older individuals are more distracted during smooth pursuit, which causes them to drift away from the target. Taken all together, these findings suggest that smooth pursuit might be more cognitively effortful for older individuals than for younger ones.

3 Current Approach

The approach taken here is to contrast the performance of older and younger individuals as they perform relatively routine computer-based tasks. The data come from participants who took part in paid research projects whose objectives did not include age comparisons. The projects were a validation comparison of multiple eye tracking systems using three tasks of varying difficulty, a market research study comparing two different package designs, and a website usability study comparing two different layouts for a shopping website. Two subsamples were obtained from each study: younger participants, aged 31 or younger, and older participants, aged 46 or older. All were screened routinely for eye problems that would disrupt eye tracking. Many wore soft contact lenses or glasses.

3.1 Variables of Interest

The variables of interest for the current study include three aspects of eye data: gaze data, raw pupil data, and workload data. Most of the variables could be obtained from all three studies examined here. They are described briefly below.

Gaze data include variables that describe where individuals looked and for how long. Gaze data come from horizontal and vertical pixel locations of the point of gaze obtained for every measurement (i.e., 60 times per second for an eye tracker sampling at 60 Hz). Three variables are reported here:

- Task Duration: time to complete a designated task. (Not applicable for timed tasks).
- Number of Task Fixations: number of fixations made during completion of the task. (Not applicable for untimed tasks.)
- Mean Fixation Duration: average length of a fixation

Raw pupil data refers to the size of the pupil that is measured continuously as an individual carries out the task._Raw pupil diameter is measured in millimeters and ranges from approximately 2-7 mm.

Cognitive workload is measured here by the Index of Cognitive Activity, a metric derived from raw pupil diameter. Full details are given elsewhere [5]. Basically, to compute the Index, the raw pupil signal is first standardized and then subjected to wavelet analysis followed by a statistical thresholding to remove noise. The result is an Index that is scaled to fall between 0-1, with higher values indicating higher cognitive activity. The Index is continuous and may be measured across any time span (milliseconds, seconds, minutes, hours). The Index values reported here span the entire task under analysis.

3.2 Data Collection

All data were collected with a Tobii TX300 eye tracker. This eye tracker can sample at two different rates, 60 and 300 Hz. Two of the studies (Studies 1 and 3) recorded data at 300 Hz, and Study 2 recorded data at 60 Hz. Fixations were defined with minimum duration of 75 ms and a pixel threshold of 25 for all studies.

4 Studies

4.1 Study 1: Smooth Pursuit and Mental Arithmetic

This study employed three tasks and was originally designed to validate the Index of Cognitive Activity across four different hardware platforms and three different sampling rates [6]. More than 100 individuals participated in the study. Significant differences were found across the three levels of task difficulty for all eye trackers. The easier task was always accompanied by significantly lower cognitive workload than the medium task, which in turn was always accompanied by significantly lower cognitive workload than the difficult task. The Index performed consistently across all platforms and sampling rates, yielding similar values on each with no significant differences across platform or sampling rate.

For the current analysis, data were taken from 18 participants who interacted with a single eye tracker (Tobii TX300). Nine of the participants were older than 47 years, and nine were younger than 32. The age boundaries were selected to allow the greatest number of participants in each age sample while maintaining a difference of at least 15 years between the two groups.

The data are all the eye movements and pupil changes that occurred during participants' completion of the three validation tasks. The first task was a relatively simple smooth pursuit. Five yellow dots were shown on the left half of a blue screen, with one initially flashing red to identify it as the target. The target dot returned to yellow after a few seconds, and all dots then moved in random patterns around the screen for about 45 seconds. At the conclusion of the task, all dots drifted to the bottom of the screen and the participant was asked to identify which was the original target.

The second task repeated the first (with dots moving in a pattern that was different from the first task) and added a mental arithmetic component. Numbers and arithmetic operators appeared on the right side of the display while the dots were moving. The background of the right side was black, and the numbers were white. The numbers and operators remained briefly on the display and then disappeared. The participant's task was to carry out all arithmetic operations and report the final tally at the conclusion of the 45 second task while at the same time tracking the target dot and identifying it at the conclusion.

Finally, the third task contained the first two tasks and added a vigilance component. A small wheel was added to the center of the display, and it rotated either clockwise or counterclockwise. The wheel changed direction every few seconds in a random pattern. The participant was asked to monitor the direction of the wheel. When a tone sounded (every 10 seconds), the participant was to say aloud either whether the wheel had been moving 'clockwise' or 'counterclockwise' when the tone sounded. At the same time, the participant was to compute the mental arithmetic sum on the right display and follow the target dot on the left display as in the second task.

Prior to starting the three tasks, participants were given two opportunities to practice the smooth pursuit of dots to become familiar with that activity. They then began the three-task sequence and completed the tasks in order: easy, medium, and difficult. The entire set took approximately 5 minutes to complete.

4.2 Study 2: Visual Search

In this study, participants were asked to search for specific food items on a grocery shelf image. The images were life size and presented by high definition projection. The original study compared two versions of a package design for one product, and participants completed a series of searches for products by different manufacturers across different grocery shelves. In the current analysis, the factor of package version was ignored. The focus here is on the ways in which older and younger individuals differed as they searched for products. Three different search tasks were analyzed, each requiring the participant to find a specified target item on the grocery shelf.

The original study was based on a sample of 150 participants of varying demographics. Two groups were drawn from this sample, with one group being all individuals older than 47 (N=14) and the other group being all individuals younger than 32 (N=62). The age constraints match those used in Study 1.

4.3 Study 3: Online Shopping

The final study was a website usability project whose objective was to compare two implementations of a shopping website. Participants were asked to shop for several items, using any features on the site as they normally would. Each shopping experience was untimed; a task was considered to be completed when the participant found a satisfactory item for purchase.

A sample of 50 participants was recruited for the original study. Two subsamples were created for the present analysis: participants older than 47 (N=5) and participants younger than 32 (N=24). For comparability with the other studies reported here, three shopping tasks were selected.

5 Results

5.1 Smooth Pursuit

One of the tasks in the three studies allows a glimpse into the aging effect on smooth pursuit. Figures 1 and 2 provide a representative sample of six subjects' gaze as they performed the smooth pursuit required of the first task of Study 1. Each plot spans roughly 45 seconds. The same pattern of pursuit is readily apparent in all plots. However, the plots of the older participants indicate substanially more jitter than those of the younger subjects, as can be seen from the size and density of the various clusters of observations in the plots. There was little variability among the younger participants; their gaze patterns looked much the same. More variability was observed in the older participants, with several older participants having very disrupted plots (e.g., the rightmost panel of Figure 2). Nevertheless, all subjects—younger and older—performed the task with little difficulty.

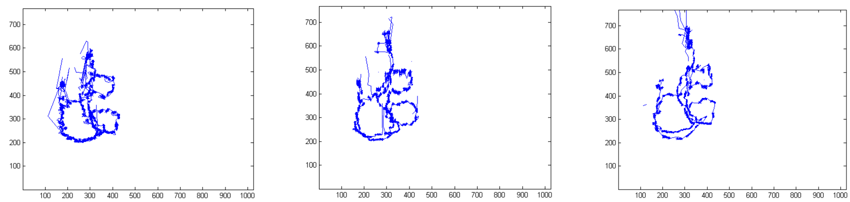

Fig. 1. Gaze traces of three younger participants on smooth pursuit

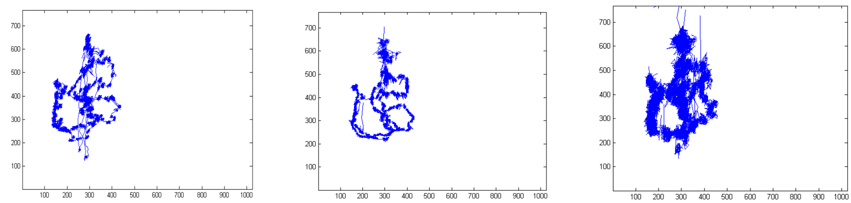

Fig. 2. Gaze traces of three older participants on smooth pursuit

5.2 Fixations

Physiological research suggests that few aging differences would be found in fixations, and the analyses here are consistent with that body of literature. Table 1 shows several fixation measures across the three studies.

Table 1. Number of Fixations, Mean Fixation Duration, and Task Duration for older and younger participants in three studies

		Study 1			Study 2			Study 3		
		Task 1	Task 2	Task 3	Search 1	Search 2	Search 3	Shop 1	Shop 2	Shop 3
No. Fixa-tions	Younger	31.1	68.1	72.9	na	na	na	na	na	na
	Older	48.1	77.7	82.8	na	na	na	na	na	na
Mean FixDur	Younger	2.41	0.69	0.55	0.23	0.27	0.24	0.26	0.24	0.24
	Older	1.27	0.53	0.48	0.18	0.22	0.21	0.22	0.24	0.21
Task Dura-tion	Younger	na	na	na	9.33	4.99	10.28	45.2	29.4	42.5
	Older	na	na	na	16.23	8.69	15.15	48.7	41.1	53.1

Task duration is not computed for fixed length tasks (Study 1), and number of fixations is not computed for open-ended tasks (Study 2 and Study 3). For all tasks in all studies, large variability was observed for both younger and older participants. Only one comparison in Table 1 is significant: the mean fixation duration is significantly larger for younger than older participants in Search 2. However, there are several general trends. For instance, older participants took longer than younger participants to complete all three searches and all three shopping experiences. Mean fixation duration is longer for younger participants than older participants across all tasks in all studies. And, older individuals produced more fixations during all three timed segments of Study 1.

5.3 Cognitive Workload and Pupil Size

Table 2 gives mean pupil size and mean Index of Cognitive Activity for both groups on all three studies. As expected, the pupil size of older participants was significantly smaller than younger participants for all three studies, with $F(1,16)=5.21$; $F(1,39)=12.25$; and $F(1,26)=6.26$. The pupil values are relatively stable across the three different samples for the two age groups.

Table 2. Mean pupil diameter and mean Index of Cognitive Activity (ICA)

		Study 1			Study 2			Study 3		
		Task 1	Task 2	Task 3	Search 1	Search 2	Search 3	Shop 1	Shop 2	Shop 3
Pupil Diameter	Younger	3.33	3.75	3.81	3.80	3.70	3.82	3.26	3.11	3.15
	Older	3.03	3.37	3.28	3.22	3.10	3.18	2.81	2.75	2.91
ICA	Younger	0.21	0.24	0.26	0.40	0.48	0.42	0.38	0.38	0.38
	Older	0.35	0.44	0.43	0.53	0.58	0.53	0.65	0.61	0.50

The Index of Cognitive Activity was significantly different for the two age groups on the three studies, with $F(1,16)=3.49$; $F(1,39)=7.24$; and $F(1,26)=6.21$. Older participants consistently showed higher workload than younger participants.

6 Discussion

This research confirms what has been previously found in clinical research and does so in the context of everyday tasks such as searching grocery images or shopping online. As expected, age differences were found in smooth pursuit, but performance was not affected. Older individuals took longer to perform visual searches than younger, but they nonetheless performed successfully. Fixation comparisons were generally not statistically significant, although older individuals seem to have shorter fixations, at least on these tasks. Older individuals also had significantly smaller pupil sizes. However, cognitive workload was larger for older individuals across all tasks in all studies.

The pupil and workload data are intriguing. On the one hand, pupil size in older individuals is smaller than in younger individuals, which is a finding that occurs repeatedly in the literature. Changes in pupil size are among the most stable results in aging studies about the eye. Because the Index of Cognitive Activity is based on the pupil, one might anticipate that it too would show a similar decrease with age. In fact, the opposite was true. On all tasks in all studies, older individuals had larger Index values, suggesting higher cognitive workload levels than those shown by younger individuals.

What do these findings suggest for interaction science? The primary implication is that these routine tasks do appear to be more difficult for the older individuals. We can turn back to some of the research cited previously to see why this might be. In the visual search task of grocery shelves, color and contrast are obvious features of the task as is visual acuity. If small print is hard to read, cognitive workload may increase. And many products on shelves have small print. Similarly, these same features may impact the online shopping tasks. Small pictures of images may be harder to resolve for older participants, and small fonts in item descriptions may again be a problem. Large numbers of items on the display or on the website might also be causing more distraction for the older viewer. All of these factors contribute to more effortful cognitive processing.

It is worth repeating again that both groups of individuals were generally successful on all tasks. But we saw significant differences in eye-movement measurements and cognitive workload between the mid-twenties and the mid-forties. One can expect that these differences will continue to grow as age increases and that performance will be negatively impacted at some point. One future goal for interaction science researchers will be to try to achieve display designs which simultaneously lower cognitive workload for older users but still attract younger users. It is outside the scope of this study to specify what these design changes might be, but it will be important to define and measure them.

Eye tracking can be a useful tool in this endeavor, providing both attentional evidence (i.e., where are users looking?) and workload evidence (i.e., how cognitively effortful is the experience?) Both types of evidence can be beneficial in assuring universal access that is satisfying to and productive for the aging user.

References

1. Meisami, E., Brown, C., Emerle, H.: Sensory systems: Normal aging, disorders, and treatments of vision and hearing in humans. In: Timiras, P. (ed.) Physiological Basis of Aging and Geriatrics, 4th edn., ch. 8, pp. 109–136 (2007)
2. Haegerstrom-Portnoy, G., Schneck, M., Brabyn, J.: Seeing into old age: Vision function beyond acuity. Optometry & Vision Science 76(3), 141–158 (1999)
3. Roth, D.: Electronic Reading Devices Increase Reading Speed and Comfort in Patients with Moderate Vision Loss. Presented at Annual Meeting of American Academy of Ophalmology, Chicago, November 11 (2012), http://newoldage.blogs.nytimes.com/2012/12/12/the-gift-of-reading/
4. Cuiffreda, K., Tannen, B.: Eye movement basics for the clinician. Mosby, St Louis (1995)
5. Marshall, S.: Identifying Cognitive State from Eye Metrics. Aviation, Space, and Environmental Medicine 78(5), 165–175 (2007)
6. Bartels, M., Marshall, S.: Measuring Cognitive Workload across Different Eye Trackers. In: Proceedings of 2012 Eye Tracking Research and Applications (ETRA) Symposium, pp. 161–164. ACM, New York (2012)