Is an Eye Tracker Really Useful in Usability Testing?

Introduction

Eye trackers have come a long way in their ability to track eye movement. Many units are passive devices that no longer require the wearing of awkward headpieces. They are getting almost affordable for many smaller labs. People can be trained on setup and use in a short period of time. Though some technical issues been resolved, most eye trackers provide only an estimation of actual eye movement.

Despite these limitation issues, eye trackers are increasingly finding their way into the usability lab. Vendors of eye trackers claim they provide “clearer results” and that they “unveil information normally missed” [1]. They claim that the eye tracker will tell you “in what order information is absorbed and processed” [2]. Users of eye trackers often create “heat maps” (a type of visual summary of the places where the participants looked) that are shown to eager clients, despite working with sample sizes too small to be generalized. But do eye trackers really provide the insight claimed? Unfortunately, the answer is mostly no.
Humans, like machines, can be conceptualized as having three elements: input, processing, and output [3]. Assuming that an eye tracker is fully capable of monitoring human eye movement, and sensitive enough to detect even quick glances at an object, what the eye tracker is actually showing us is the input to the participant at the sensor level. What the eye tracker fails to take into account is the processing that occurs to this input. This processing determines what happens to this input, i.e., whether it results in any change to us cognitively or behaviorally. Also, this processing is rather sophisticated in the case of humans. This lack of accounting for how visual information is processed results in both false positive and false negative assumptions.

Issue 1: Limitation in Technology

Some Our eyes briefly fixate to process information but are constantly moving. These movements are known saccades. Constant input into a photo receptor diminished quickly. It is through these saccades that we are able to maintain input from a static signal. In addition, the density of photoreceptors in the eye gives us a equivalent native resolution of about one megapixel. It is through small saccades that our perception is a high-resolution image. These saccades are below the detection of most eye trackers and are of particular concern. However, our eyes are also capable of movement at 900 degrees per second. We use these wider eye movement to maintain awareness outside of the man focal area. The ability to detect these very high-speed movements are important since they do tell us when a person is checking the peripheral areas. Detecting this high speed movement is also beyond the ability to most eye trackers. A high-speed eye tracker that operates as greater than 250 Hz only offers accuracy (low systematic errors) for sentences, words, or (depending on the actual refresh rate) characters. Most commercially available eye tracker run well below even this this rate. Lacking true movement input, software computational definitions associate cluster of gaze coordinates within a specified range in space and time as a single fixation.

Even assuming we have a high-speed eye tracker, the most reliable tracking occurs when the participant refrains from head or body movement. In laboratory research on eye movement, head movements are typically minimized by the use of a chin and/or head rest. This is unnatural for a typical usability study and participants often move about both naturally and particularly when engaged or confused by interactions with a product. This further corrupts the data available to the software.

Issue 2: We See What We Look At

Ignoring the technical limitations, people viewing eye tracking data typically assume that the images that are fixate on is that we see. This is, in fact, false to a large extent. Our perceptual system is not like a camera; it does not dutifully record what is in front of us. The translation from what is in front of us to what we perceive is influenced at every step in the process from detection through processing and storage. The input signal is filtered by limits in attention, and then modified by being evaluated and processed through our emotional and cognitive systems. A gist of the event is stored in memory (our memories are not complete images or recordings of events despite what is often called “snapshot memories”) that deteriorate over time unless recalled, but corrupted (modified and added to) each time we recall them.

Consider some of the facts about basic process involved in “seeing” something. The human visual system is a complex combination of signal detection sensors, pre-processing some aspects of information within the eye, and processing other aspects at several locations within the human brain. The human visual system attempts to make logic and order out of what is put in front of us. The pre-processing of information is reasonably easy to demonstrate. And it occurs outside of our awareness and outside of our control.

Consider the two pairs of objects in Figure 1 below. The first image is a combination of a plus sign and a dot. Close your left eye and look directly at the plus sign. Then slowly move closer to the image (or print the page out and move it closer to your face). At a point a few inches away from your face, the dot on the right-hand side will disappear. Note that this will only work if you keep your main focus on the plus sign.

Figure 1: The Demonstration of Visual Data “Fill In”

This phenomenon occurs because this procedure places the dot within the optic nerve of the eye—an area without any photoreceptors. But our visual processing system does not allow us to “see” a hole in our vision. We “see” the visual information that our processing system assumes should be there—in this case a white field and not a lone dot floating in space [4]. In other words, our visual processing system fills in with white space, and what we “see” is not really what is in front of us.

Now repeat the same experiment with the image on the right. In this case, at the same distance away as in the first experiment, the two lines will suddenly become joined. Our visual processing system will insert a line segment between the two line segments. “Seeing” is the unconscious processing of data that enters (or fails to enter) our eye.

In addition, the relationship of other objects in the area in which we look has a profound effect on how we consciously perceive these objects. This phenomenon was studied as far back as 1889 by Müller-Lyer [5]. Consider Figure 2 below [6]. The figure shows a pair of blocks either side-by-side or on top of each other. One block is light in color and the other block is dark in color—or so it would appear. You may be surprised to learn that the two blocks on top of each other in the image below are actually the same shade of grey. To prove this to yourself, place a finger across the edge where the two blocks meet.

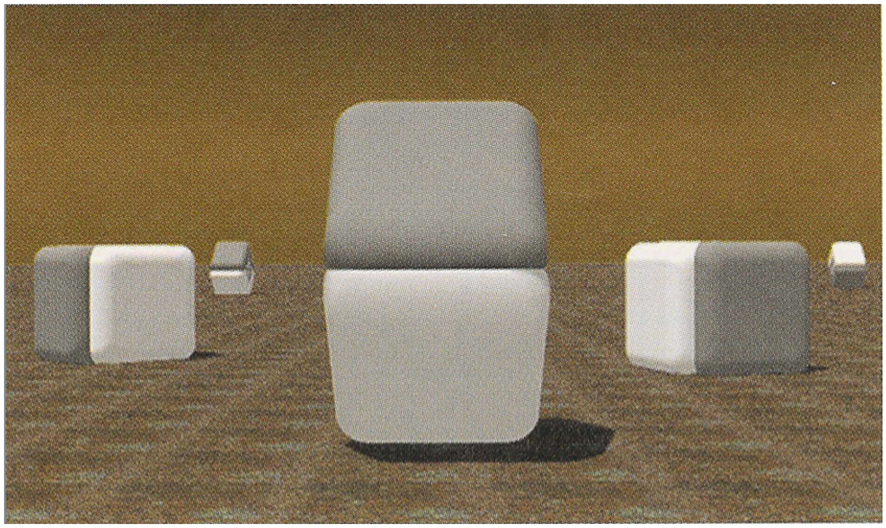

Figure 2: The Demonstration of Relative Perception

The gradient that ends the bottom of the top block and the reverse gradient that ends the top of the bottom block are incorporated into our perceptual system to determine what we believe is in front of us, even though it is not an accurate perception of what is actually there.

Another example of our inability to accurately perceive the information right in front of us is shown in Figure 3 below, originally developed by Roger Shepherd [8, 9]. This figure shows two apparently different coffee tables. In actuality, the two tabletops are exactly the same size, just rotated by ninety degrees. (You'll probably need a ruler to prove this to yourself since just knowing this information will not change your visual perception.)

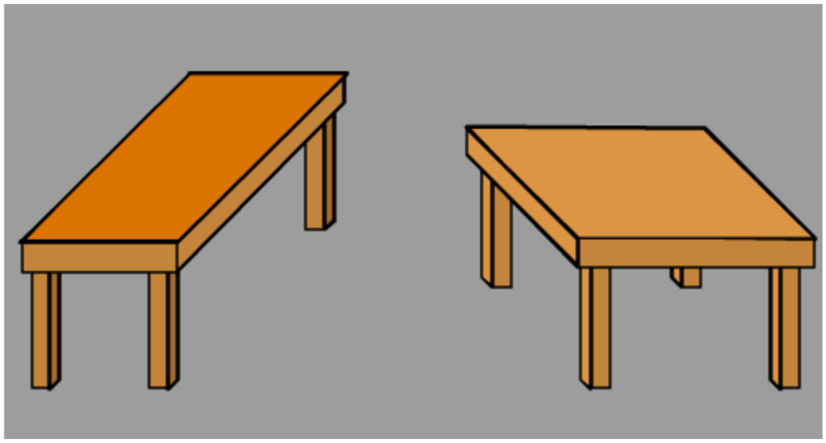

Figure 3: Roger Shepherd's "Turning the Tables"

Another effect that needs to be considered in how we process visual information is attentional filtering. At any given moment, our visual field is being bombarded by a vast amount of information. Except in very limited cases, the amount of information presented to us is more than our brains can possibly handle [10]. As a result, the human mind makes rapid decisions about what information within our visual field should be attended to and what should be ignored.

This effect has been researched for many years by Chabris and Simons [11]. Multiple experiments have been conducted to demonstrate that we cannot process all of the visual information presented to us, particularly when we are trying to perform a conscious task (a condition inherent in a usability study). Chabris and Simons coined the term "inattentional blindness" to describe this phenomenon. One version of the original experiments showing this effect is available on YouTube (http://youtu.be/Ahg6qcgoay4), though many other examples can be easily found on the web.

Inattentional blindness and its close relative, change blindness, are not random events. Though they are unconscious processes, they can be shown to be purposeful in nature. We process the information we see in front of us and make a decision about whether it fits or not into what we are thinking. If it doesn't fit, we often perceive it to be different from what it actually is in order to make it fit our decision, or we ignore it.

The Chabris and Simons experiment discussed above (which used a gorilla instead of a bear) has been studied with an eye tracker. According to Chabris and Simons on their research into this phenomenon:

> “Sports scientist Daniel Memmert of Heidelberg University ran our gorilla experiment using his eye tracker and found that subjects who fail to notice the gorilla had spent, on average, a full second looking right at it—the same amount of time as those who did see it!” [12]

In other words, the people who fail to “see” the moon walking bear fixated on it but, through an unconscious decision-making process, decided not to process the image into memory since it did not fit within the context of counting passes and the time required to process this image would have interfered with their conscious task.

This author was involved with research into the use of subliminal messages in commercials. Subliminal messages are generally one or two frames within a film intended to send a message to the viewer without them noticing it. However, in one case the “subliminal message” was a full second of video—enough visual information that it should have been perceived by all observers. The video, which happened to be of a naked woman fondling her breasts, was placed within a commercial for a bank. Most observers failed to notice the presence of the woman at all. It had no logical place in a bank commercial, so it was ignored (more accurately it was dismissed unconsciously). Those observers that did notice the woman may have experienced it a memory and not a visual experience. However, when told the image of the woman was real, people did notice it but explained the woman’s behavior by stating that she was taking a shower—which was, in fact, not the case. For these observers, the presence of the image was modified to make it more acceptable. No observers observed the visual event accurately.

Knowing that a person's eyes were directed at a spot on a screen for a specific period of time (a "fixation") does not tell us: (1) whether or not they processed what their eyes fixated on, (2) how accurately they processed it (assuming they processed it at all), or (3) how that processed information was integrated into memory (if it was).

Issue 3: We Need to Look at Something to See It

This other issue with reviewing eye tracking data is the assumption that we only see what we fixate on. This is also untrue.

The human visual system is made up of two types of vision: foveal vision and peripheral vision. Within the approximate center of each eye is a small (about 5mm), yellowish region called the macula. Contained within the macula is the fovea—an area of highly packed photoreceptor cells. This area of packed cells (generally known as our central vision) provides us with high acuity color information. The central vision provides us with a visual field of approximately 13°.

To "see" an object with our fovea does require us to look directly at the object. And this means that the central vision *usually* corresponds to our attentional focus, since we usually purposefully move our eyes to cause the object of our attention to fall within this region of our optical system. This is what an eye tracker is tracking.

However, outside the macula is a much larger area of cells that comprise our peripheral vision. This area of the eye, having far fewer photoreceptors, produces information of significantly lower acuity but is much better at detecting black and white information (which means it works well even in low lighting conditions). It also contains a number of specialized cells for detecting things such as horizontal and vertical structures, edges, and movement. The peripheral vision extends the field of view from approximately 13° to approximately 180°. Information presented to our peripheral system is processed, even at the lower visual acuity, for a different purpose than information presented to our central vision.

These two types of vision work together. Researcher Hans-Werner Hunziker [14] identified three primary functions of our peripheral vision: (1) the identification of similar forms and movements, (2) recognition of well-known structures, and (3) the delivery of sensations that form the background of our detailed vision. From a design perspective, the most significant element of our peripheral vision is how we use the information obtained from it to determine how much attention (if any) is needed to better classify an object that is detected in our periphery. In other words, objects detected within our visual field are processed (1) to determine if they are already of a readily identifiable form without requiring direct focus, or (2) to direct our focus on them to get more information. If we believe that we know what the object in our peripheral vision is, we continue to focus on the objects currently within our central vision and do not redirect our eye's central focus to that peripheral object.[1]

Many people are trained to consciously process the objects within their peripheral system. Astronomers are taught to use this area of their vision to detect faint objects more clearly. A dim star, for instance, is best seen when your

---

[1] If this function was not performed by our peripheral vision, we could not maintain focus on an object for very long, since we would have to continuously look at every object that appears anywhere in our total visual field. We would, in effect, have no ability to fixate.

eyes are not aimed directly at it.[2] Anyone who plays basketball, soccer, hockey, or any other sport involving passing between its team members learns to make a “blind pass” by passing to a player they are not looking at directly. (Otherwise, the pass is known as being “telegraphed.”) But the bulk of peripheral vision processing is done at an unconscious level.

An example of this phenomenon is common to most people who have conducted a usability evaluation and is often described as “banner blindness”—the tendency for people to ignore items in their peripheral vision that look like ads or banners [15]. However, this is often mistakenly described as users failing to “look at” objects in their peripheral vision. When asked why they did not look at the object, respondents often state that they “recognized” them as an advertisement or some other information that was not pertinent to the task at hand. (This is often a false assumption about the object, since designers often put objects in the user’s peripheral vision that are not advertising and are supposed to be used for the task.) This demonstrates that these objects are seen, processed, and affect the behavior of people without ever having to look directly at them.

Since an eye tracker is associated only with central vision, eye tracking data ignores the information that is processed in our peripheral vision (which, based on visual field size, is approximately 72% of our total visual field).

Issue 3: As Observers, Seeing is Believing

Discussions about the limitation or value of eye trackers are not new. Professional forum discussions occur regularly with strong opinions on both sides. Some researchers warn about the limitations and about making false inferences, but people who have attended usability sessions where eye trackers are used are

---

[2] It is believed that Aristotle detected one of the nebulae he discovered using this technique.

quite certain of their value and of the validity of the data they are seeing. They state they have seen eye trackers in use and *know* they work. This may be the most concerning issue of the use of eye trackers.

Presented with the information that an eye tracker *can* produce, limited though it may be, observers are able to provide rational explanations for what they are seeing. But humans are inherently wired to try to determine cause and effect relationships between observations. This tendency creates a set of reasoning biases, including the post hoc or false cause bias (a tendency to assume cause and effect relationships between two elements even if that relationship does not exist), the confirmation bias (a tendency for people to identify information that confirms preconceptions regardless of whether the information is true), and other biases [16, 17]. These tendencies form a powerful basis for rationalization—beliefs that can be explained in a rational and logical way but may often be untrue.

Since rationalization is internal (and therefore internally validated), these beliefs are powerful and difficult to challenge, even if incorrect. Given the limitation of the actual data present, coupled with the certainty with which observers believe what they see can be explained, observers can arrive at false but powerful conclusions.

Even if people could avoid these biases (which is highly improbable), watching the real-time display of eye tracking data during the test (a common way in which eye trackers are being integrated into usability testing) tends to distract the observer and perhaps even the facilitator. It distracts them from “being in the moment” with the participant and therefore being able to detect and understand the participant’s reactions and behaviors. Such events in testing are vastly better indicators of how “usable” a product is compared to participants’ eye movements.

Observers, also human beings, have limited attentional focus, and movement is a strong distractor. Observers tend to follow the eye tracking data and are unable to process other information, such as tone, intonation, body movement, facial expressions, and other elements that provide valuable data about the user experience—data arguably more valuable than what is obtained from an eye tracker.

Conclusion

Eye trackers are not bad things. There are specific cases where an eye tracker can provide useful information. These tend to be in controlled search settings where eye tracking behaviors are studied directly. But even in tracking research, eye tracking data is often compared to other data collected, and there are specific cases where an eye tracker can provide some useful information in a usability study. For example, when testing alternate visual layouts (where the participant does not have to engage with it beyond looking at it), viewing eye tracking data can give some insight into which elements attract the user's eye (though additional investigation is required to determine whether and how any fixations were processed). But even in this case, it is only possible to investigate fixations that are processed at a conscious level. If a fixation was processed at a subconscious level (where most psychologists believe that approximately 98% of our decision-making occurs), it is, by definition, impossible to discuss with the participant how it was processed and how they perceived it.[3]

Another useful application of an eye tracker in a usability study is to observe how people read a specific block of text. In this case, as in the case of static screen eye tracking, the data from eye tracking can help reveal specific words or phrases that

---

[3] There are numerous research studies that study subconscious influences that could possibly be used in these cases, but the methodology is far beyond the scope of a usability evaluation.

require repeated attentional focus—an indication that these phrases are either unclear, intriguing, or are understood and verified by subsequent text. In this case, though participants are often aware of this repeated reference back to troublesome words and phrases, the repeated reference observable in their eye tracking data may be unconscious.

However, despite having specific instances where it can be useful, the general use of an eye tracker in a usability evaluation is questionable. An eye tracker might provide real time feedback that helps identify an potential event to follow up on, not as conclusive data. The display of real time eye tracking data to other observers, who are unlikely to understand the psychological and physiological limitations of eye tracking data, is likely to compel observers to draw incorrect conclusions. And, as described above, eye tracking can distract from the observation and processing of other, more valuable data. Therefore, care should be taken to properly inform observers, or the real time data should not be present in the observation room. Finally, summary data (e.g., heat maps) should be avoided, since the data is both inconclusive and from a small sample.

Practitioners would be better served if the use of an eye tracker was associated with evaluating static page layouts, evaluating blocks of text, or determining when a participant should be interrupted to clarify a behavior. Practitioners should not assume that usability or performance data can be obtained just from tracking how a person's eye is moving and where it fixates.

References

[1]: http://www.interactivevideo.com/eye_tracking.html - IVP Eyetracking Labs, accessed 8/12/11

[2] http://www.schlesingerassociates.com/qualitative_solutions/technologies/eye_tracking.aspx - Schlesinger Associates, accessed 8/12/11

[3]: Meister, D. (1987). Systems design, development and testing. In *Handbook of Human Factors*, G. Salvendy, Ed., Wiley, New York, 17-42.

[4]: Ramachandran, V. S. & Gregory, R. L. (1991). *Perceptual filling in of artificially induced scotomas in human vision*. Nature, 350, 699-702.

[5]: Muller-Lyer F. C., (1889). *Optische Urteilstäuschungen*. Arch. Anat . Physiol., Physiol . Abt . 2, 263–270.

[6]: Muller-Lyer brick graphic

[7]: http://www.michaelbach.de/ot/sze_shepardTables/index.html - Shepard's rotated tables, accessed 8/16/11

[8]: Shepard, R. N. (1981). *Psychological complementarity*. In Kubovy, M. & Pomerantz, J. R. (eds.) Perceptual organization, 279-342. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
[9]: Shepard, R. N. (1990). *Mind sights: Original visual illusions, ambiguities, and other anomalies*. New York: WH Freeman and Company.

[10:] Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, New Jersey: Prentice Hall.

[11]: Simons, D. J. & Chabris, C. F. (1999). *Gorillas in our midst: Sustained inattentional blindness for dynamic events*. Perception, 28, 1059-1074.

[12]: Simons, D. J. & Chabris, C. F. (2010). *The invisible gorilla: And other ways our intuitions deceive us*. New York, New York: Crown Archetype.

[13]: Christman, R. J. (1971). *Sensory experience*. International Textbook Company: New York, NY.

[14]: Hunziker. H. W. (2006). *Im Auge des Lesers: Foveale und periphere Wahrnehmung - vom Buchstabieren zur Lesefreude [In the eye of the reader: Foveal and peripheral perception - from letter recognition to the joy of reading]*. Transmedia Stäubli Verlag: Zürich, Germany.

[15]: Benway, J. P. & Lane, D. M. (1998). *Banner blindness: Web searchers often miss 'obvious' links*. Internet Technical Group, Rice University.

[16] Plous, S. (1993). *The Psychology of Judgment and Decision Making*, McGraw-Hill.

[17] Bazeerman, M (1998). *Judgment in Managerial Decision Making*, 4th Edition, John Wiley & Sons.

About the Author

Bill Killam is a Human Factors Engineer with over 40 years of experience. He holds professional certification in Human Factors Engineering from the Board of Certification in Professional Ergonomics and is an active member of the American Psychological Association, the Human Factors and Ergonomic Society, the Usability Professionals Association, and the Association of Computing Machinery. He is the president and principle consultant at User-Centered Design, Inc., a Human Factors Engineering consultancy in Northern Virginia. He is also a professor in the College of Systems Engineering and Operations Research at George Mason University, the College of Information Studies at the University of Maryland, and the School of Library and Information Science at Catholic University of America.