More than Just Eye Candy

Top Ten Misconceptions about Eye Tracking

By Agnieszka (Aga) Bojko and Kristin A. Adamczyk

Eye tracking is no longer a novel addition to the user experience (UX) research toolbox, used by only a few specialists. As more UX professionals incorporate eye tracking into their studies, many misconceptions are being created and perpetuated. These false beliefs and questionable practices have given eye tracking an undeserved bad reputation. It is time to start the process of change. This article describes common misconceptions about eye tracking in the UX field and attempts to clarify its proper application.

MISCONCEPTION #1

All usability research can benefit from eye tracking.

One common belief is that any usability study will provide better insight if accompanied by eye tracking. However, a simple cost-benefit analysis of the insight gained versus the amount of additional time and resources that eye tracking requires will show that eye tracking is not always appropriate.

In formative usability testing, the ratio of insight to cost for eye tracking is very small; most usability issues can be uncovered with traditional usability research. For example, if a study goal is to improve the overall user experience on a new website or to make instructions for a medical device easier to understand, eye tracking may be of little value.

However, eye tracking will benefit studies that aim to answer specific questions that arose from previous testing (for example, Why are users struggling to find drug storage instructions on a package insert?). Eye tracking can also be useful in summative testing as it provides additional measures to help quantify the user experience.

MISCONCEPTION #2

Eye tracking is all about heatmaps.

Some UX researchers may not be aware that visual attention can be quantified in a variety of ways, and seem to believe that the goal of eye tracking is simply to produce heatmaps, such as the one in Figure 1. This is why many stakeholders do not consider eye tracking to be a valuable tool in the field. It is hard to blame them—all they have seen as output are pictures that provide no definitive answers.

Even practitioners who try to quantify attention often use just one or two measures, not realizing that there are a multitude of measures to choose from, each providing insight into different cognitive processes.

It is, therefore, important to analyze the measures that address the study questions. For example, to understand the interest that a particular object creates, the number of fixations on the object should be used instead of the average fixation duration (which indicates information processing difficulty, among other things) or the time to first fixation on the object (which indicates its discoverability).

MISCONCEPTION #3

Eye tracking results are easy to interpret.

"We just want to know where they are looking," is often mentioned as a research objective. The people saying that seem to believe that if they knew where users were looking, they would know how the interface should be improved. Many stakeholders will be quite happy to receive heatmaps and/or a description of where the study participants looked. However, some will eventually realize that there is no easy translation between attention distribution and design recommendations, which will lead to the infamous "so what?" question. It is often overlooked that eye tracking results reveal the "what," but not the "why."

"Even practitioners who try to quantify attention often use just one or two measures, not realizing that there are a multitude of measures to choose from, each providing insight into different cognitive processes."

For example, the DVD New Release promo for the movie *21* in the center of the Blockbuster homepage shown in Figure 1 attracted significantly more fixations than the other four promos in the carousel. Did the promo for *21* get more attention because the movie received more hype than the other three movies and was thus of more interest to the participants? Did it get more attention because the promo had human faces in it while the others did not? Were participants who came to the website to look for movies to rent simply more interested in seeing

Figure 1. Heatmap representing the aggregate number of fixations across eighteen participants looking for movies they would be interested in renting (red = 10+ fixations).

DVD releases than what was in theaters or on TV? In studies with no systematic variable manipulation, the "why" is very difficult to tease out and the researcher can only speculate.

MISCONCEPTION #4

There is *only one* way to look at something.

Not all practitioners realize that context (task, background information, time allotted, etc.) can significantly change the way a stimulus is viewed. People do not look at anything in only one way. Asking someone to find a specific product on a web page will produce completely different results than having him look for the company contact information. If participants are not given specific tasks, it will be difficult or impossible to interpret the results in a meaningful way.

Also, results (both quantitative data and their visualizations) must be presented with the proper context or they will be meaningless. Heatmaps accompanied by statements such as, "This is how participants looked at the page," must all beg the question, "While doing what?"

MISCONCEPTION #5

Let's track and see what comes out!

Some studies are conducted without a clear understanding of how eye tracking data will address the research objectives. The researcher collects data, and then goes on a "fishing expedition" trying to retrofit results to objectives. Let's say that the goal of a study is to determine which of two designs is "better." Unless "better" is defined prior to data collection, the researcher may look to see if there are *any* differences between the designs, which introduces bias.

A lack of preparation for eye tracking studies can also result in "could have/should have" moments during analysis. Not thinking ahead may cause the need for a follow-up study. In addition, a poorly prepared study makes analysis more difficult and time consuming, especially if the recorded eye movements are not divided by task. Prior to the study, the researcher must select a clear start and end point for every task, and set up the study in a way that makes tasks easy to identify and compare later on.

MISCONCEPTION #6

There is a magic sample size for all eye tracking UX studies.

We often hear claims that you need thirty participants to conduct an eye tracking study. However, this is an oversimplification. There is no one sample size appropriate for all eye tracking studies. As in any other type of study, the sample size depends on multiple factors including research objectives and study design. Before the sample size can be decided, the researchers should determine if they are comparing different conditions or trying to generalize a particular score to the population. If two or more conditions are being compared, will each participant experience all of them (within-subject design) or just one (between-subjects design)? Also, what effect size should the study be able to detect?

Thirty participants are more than sufficient for a qualitative study in which the eye movement data are used to illustrate certain usability findings. Thirty should also be enough for a within-subjects study with a large expected effect size. However, thirty is not sufficient if five different designs are tested in a between-subjects study and definitive (that is, statistically significant) rather than directional results are required.

"There is no one sample size appropriate for all eye tracking studies. As in any other type of study, the sample size depends on multiple factors including research objectives and study design."

MISCONCEPTION #7

Eye movement analysis can be done by watching gaze videos in real time.

Some practitioners believe that they do not need to conduct a formal data analysis. Instead, they base their findings on what they saw while watching the participants' gaze point, typically denoted by a moving crosshair or dot overlay on the stimuli.

However, people make several eye fixations

Figure 2. Frames from a recording of a participant examining baby monitor packaging. The crosshair indicates his point of gaze.

per second. One minute of eye tracking can result in 200-300 data points! Therefore, it is impossible for a human to process and remember the data for one participant, let alone reliably aggregate data across several participants, by watching videos of their eye movements in real time.

Further exacerbating the problem is researcher bias. Since researchers know the objectives of the study, they may overemphasize the amount of attention on areas of interest to the study. They may also prioritize gaze patterns that make sense to them while downplaying gaze patterns that may be more representative of the study's participants.

Real-time videos are good for illustration of findings but not for analysis. The data needs to be systematically aggregated before the results can be objectively determined.

MISCONCEPTION #8

The dot indicates exactly where a person looked and what they saw.

When viewing gaze replays or fixation plots, some practitioners incorrectly believe that the dot (or crosshair) indicates exactly where the participant was looking.

First of all, the eye trackers typically used in the UX field are less accurate than practitioners may assume. It is not uncommon for there to be a difference of up to a centimeter (on the recording) between the recorded gaze point and what the participant was actually focusing on. This difference tends to increase as the eye tracking session progresses due to changes in the relative position between the participant and the infrared cameras that capture the image of the eye. This can happen, for example, when a participant starts slouching while sitting in front of a remote eye tracker, or when a wearable eye tracker moves on the participant's head. Figure 2 shows a few frames from a recording made with a wearable eye tracker. The participant seems to be reading the text on the baby monitor packaging even though the crosshair indicating his gaze point is sometimes outside of the box.

Second, the area of foveal vision (where people have the highest acuity) is often larger than the mark indicating a fixation. For example, when looking at a computer screen, the area a person sees with high resolution could be twice the size of their thumbnail.

Both the imperfect tracking accuracy and the properties of human vision should be taken into account during data analysis. One way to do this is by making areas of interest large enough to ensure that all relevant fixations are captured. When areas of interest are very close together, it is important to recognize that fixations captured

Figure 3. Gaze plot of a participant trying to determine if the prepaid phone was web-enabled. The red circles indicate fixations—where the eyes were relatively still and the person was focusing on a particular spot on the packaging.

within one area could belong to the other one.

Another reason there is no perfect relationship between the recorded gaze point and what people see is parafoveal and peripheral vision. Though objects outside of the fovea appear as blurry, they can be registered if they are large and/or familiar enough. For example, in Figure 3, it can be assumed with a high degree of certainty that the participant saw the phone in the packaging even though she did not look directly at it. Therefore, we should never say that someone did not see something, only that they did not directly fixate on it.

MISCONCEPTION #9

All data collected should be analyzed.

Eye tracking data is no different from the data obtained with other research methods in the way it should be treated. However, data cleansing procedures do not seem to be commonly used in eye tracking. The inclusion of outliers and participants who did not calibrate or track sufficiently well in the analysis may lead to an inaccurate representation of the data.

Just like researchers may exclude usability test data for a participant who took ten minutes to complete a task while everyone else took less than a minute, they should consider excluding participants with unusual (for example, three standard deviations from the mean) eye movement data. These data may indicate that task instructions were not followed.

Including data for participants who tracked only part of the time may also confound the results because a lack of attention can be meaningful. Without carefully examining gaze replays, it is impossible to know if these participants did not have any fixations on a particular area due to lost tracking or because the area did not attract their attention.

MISCONCEPTION #10

Anyone can do eye tracking.

Manufacturers of eye trackers do their best to make their hardware and software easy to use. After all, if anyone can do eye tracking, more systems will be sold. There is nothing wrong with tools that are easy to use and make researchers more efficient. However, just because someone can operate an eye tracker does not mean they can do eye tracking. Quality eye tracking research involves a lot more than clicking on buttons. In addition to common sense, it involves some knowledge of how the eye works, visual perception processes, previous research, research methodology, and statistics, among other things.

By applying more rigorous practices that are often used with other research techniques, UX professionals can ensure that eye tracking does not become a circus sideshow, but continues to generate valuable insights into the cognitive processes of the users, thus helping improve the user experience.UX

About the Authors

Aga Bojko is an associate director at User Centric, Inc. She has more than ten years experience with various types of eye trackers that she has used to conduct UX research with software applications, websites, instructional material, pharmaceutical labels, and product packaging. Aga holds an MS in Human-Computer Interaction from DePaul University and an MS in Human Factors from the University of Illinois.

Kristin A. Adamczyk is a user experience specialist at User Centric, Inc., where she focuses on combining eye tracking with more traditional user research methodologies. Since 2006, she has collected data from hundreds of participants using remote and wearable eye trackers. Kristin has an MS in Human-Computer Interaction from DePaul University.