EYE TRACKING THE USER EXPERIENCE

A Practical Guide to Research

by Aga Bojko Foreword by Steve Krug

Rosenfeld

CONTENTS

PART I

Why Eye Tracking?

CHAPTER 1

Eye Tracking: What's All the Hoopla?

Eye tracking, which is the process of identifying *where* someone is looking and *how*, has generated a great amount of interest in the user experience (UX) field since the beginning of the twenty-first century when the technology started becoming more widely accessible. Once a novel addition to the UX research toolbox, used by only a handful of early adopters, eye tracking is now frequently employed to help evaluate and improve designs (from websites to product packaging) at various stages of the development cycle.

Because it captures behaviors that are not easily controllable (by study participants) or observable (by researchers), eye tracking has been perceived as both more scientific and more "magical" than conventional usability testing methods. Initially, this perception resulted in eye tracking frequently being used for its own sake, regardless of study objectives. The common belief was that any study would produce better insight if accompanied by eye tracking.

When I started applying eye tracking to UX research in 2003, the typical approach in the field seemed to be "ready, fire, aim" or "track now, think later," as we used to fondly call it. Practitioners would often turn on their newly acquired eye trackers and collect eye movement data with no consideration for the study design or the outcome. They would then embark on a fishing expedition, looking for data that might address their questions, failing to realize that they should have structured their study differently to obtain meaningful results.

While the mentality of eye tracking as the "be-all end-all" and the "track-now-think-later" approach still exist to some extent, more and more practitioners realize that in order to learn something useful from eye tracking, more emphasis must be placed on science and less on magic. They recognize the importance of being aware of both the capabilities and limitations of eye tracking, knowing how to properly incorporate it into UX research, and learning how to interpret and communicate eye tracking findings. This book covers all these topics, but before we start diving into deep waters, let's first examine the concept behind eye tracking.

A Quick Look Back

Eye tracking as a technique originated in reading research. Researchers in the late 1800s realized that people's eyes didn't move as smoothly through text as it had always been assumed. This (unaided) observation prompted researchers to develop technology to measure eye movements in an effort to better understand how people read.

The first eye tracking devices appeared in the early 1900s. These eye trackers were intrusive because they relied on electrodes mounted on the skin around the eye or on the use of large, uncomfortable contact lenses that study participants had to wear. Non-intrusive eye tracking techniques started emerging shortly thereafter. They involved recording light that was reflected on the eye or filming the eyes directly.

The advances in eye tracking technology since then have focused on reducing the constraints posed by eye trackers on research participants, while increasing the precision and accuracy of these devices, as well as making data analysis easier. At the same time, eye tracking research has deepened researchers' understanding of the relationship between the different aspects of eye movements and the human cognitive processes.

The first application of eye tracking to UX-related research dates back to 1947, when Paul Fitts and his colleagues investigated how pilots used the information provided by instruments in the cockpit to land a plane.[1] At the time, however, eye tracking was still primarily used by academic and medical researchers. It wasn't until the late 1900s and early 2000s when the technology, mostly due to its improved affordability and usability, became more widespread among practitioners.

To learn more about the evolution of the eye tracking technology and details of how it works, you should check out Duchowski's *Eye Tracking Methodology*.[2]

---

1 Paul M. Fitts, Richard E. Jones, and John L. Milton, "The Movements of Aircraft Pilots During Instrument-Landing Approaches," *Aeronautical Engineering Review* 9, no. 2 (1950): 24–29.

2 Andrew T. Duchowski, *Eye Tracking Methodology: Theory and Practice*, Second Edition (London: Springer-Verlag, 2007).

What Is Eye Tracking, Anyway?

You are hopefully reading this book not because you want to build an eye tracker, but because you want to make use of eye tracking in your research. If that is the case, you do not need to know exactly how the hardware works to be successful in using it, just like you do not need to know what is under the hood of your car to be a good driver. However, as a professional, you should be at least somewhat well versed on the topic.

If you are already involved in eye tracking research, then you probably know what I mean. I am often asked about how eye tracking works by research stakeholders, other UX practitioners, study participants, and even my friends. And how can I blame them for their curiosity? Eye tracking is indeed fascinating.

Imagine that someone at a party overhears you mentioning eye tracking. Let's call him John.

> **JOHN** (*wrinkling his forehead*): Eye tracking? What is that?
>
> **YOU:** Eye tracking is the process of determining where someone is looking. It can also measure the characteristics of eye movements and the eye itself, such as the size of the pupil. To conduct eye tracking, you need special equipment called an *eye tracker.*
>
> **JOHN:** An eye tracker?
>
> **YOU:** Yes, an eye tracker. It's a piece of hardware that records your eye movements as you look at a computer screen, a physical object, or even your surroundings in general. Some eye trackers are affixed to a pair of glasses or a special hat you can wear. Others can be placed in front of you, like those that are attached to computer monitors.
>
> **JOHN:** This sounds pretty cool. But how does it work?
>
> **YOU:** The eye tracker shines infrared light onto your face, and then it records two things: the reflection of the infrared light from the retina, which helps find the center of your pupil, and the reflection of the infrared light from the cornea, which is called *corneal reflection.*
>
> **JOHN:** Retina? Pupil? Cornea? You kind of lost me there.
>
> **YOU:** The retina, pupil, and cornea are parts of the eye. Let me show you the eye diagram that I carry in my wallet for occasions such as this one (*proudly taking the eye diagram from your wallet*

[see Figure 1.1]). The retina is a light-sensitive tissue in the back of the eye. The pupil is a black-looking opening that allows light to enter the retina. The cornea is the transparent front part of the eye.

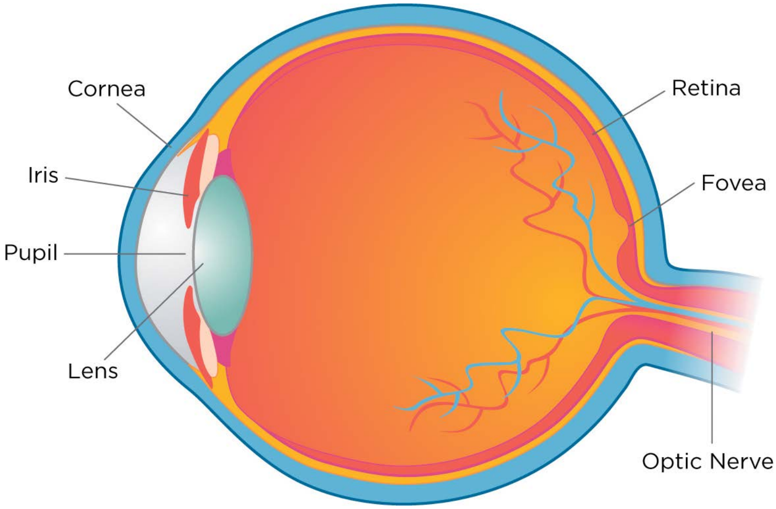

FIGURE 1.1
The human eye.

JOHN (*nodding*): Uh-huh.

YOU: If you look at my eyes right now, you will see the corneal reflection of the light in this room in each of them. If I keep my head still and look to the left, to the right, up, and down *(demonstrating),* the corneal reflection doesn't move—only the pupil does. You can see that the relationship between the pupil center and corneal reflection changes (see Figure 1.2).

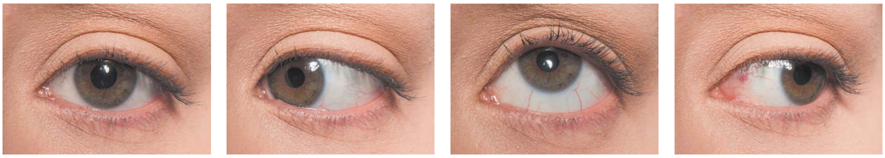

FIGURE 1.2
The relative position of the pupil and corneal reflection changes when the eye rotates but the head remains still.

JOHN: So where you are looking can be determined from the location of the pupil center relative to the corneal reflection.

YOU: Exactly. Now, if I move my head slightly while looking at the same spot *(demonstrating)*, the relationship between the pupil center and corneal reflection remains the same (see Figure 1.3). Even though I'm moving, the eye tracker would know I'm looking at the same spot.

FIGURE 1.3
The relative position of the pupil and corneal reflection does not change when the head moves but the person is looking at the same spot.

JOHN: So what's inside of the eye tracker that allows it to do something like that?

YOU: Modern commercial eye trackers consist of two main components. The first one, a source of near-infrared light, creates the reflection in the eye. The second component is a video camera sensitive to near-infrared light. The camera is focused on the eye and records the reflection. The software then figures out the location of the gaze and superimposes it onto an image of what you were looking at, such as a Web page.

JOHN: Why is infrared light needed? Wouldn't regular light work?

YOU: The trick is to use a wavelength that is invisible to people, and thus not distracting, yet reflected by the eye.

JOHN: But isn't infrared light dangerous?

YOU: Any light wavelength—ultraviolet, visible, and infrared—can be harmful in high intensities, but the exposure from the eye tracker is just a tiny fraction of the maximum exposure allowed by safety guidelines. There is no danger, even if I were to track your eyes for hours.

This is when you and John realize that everyone else who was initially listening to your conversation has already walked away, and you decide to rejoin the party.

Webcam Eye Tracking

While most commercial eye trackers are based on the infrared illumination approach described in this chapter, it is important to mention the recently evolving appearance-based systems. Instead of relying on infrared light, these low-cost solutions use off-the-shelf webcams to extract and track eye features on the face. Webcam eye tracking is most often employed in remote testing, during which participants use their computers at home or at work without having to come to a lab.

One of the current constraints of webcam eye tracking is poorer accuracy as compared to the standard infrared devices. The accuracy decreases even further when participants move around or move their computer—something that's difficult to control in a remote session (see Figure 1.4). In addition, the rate at which the gaze location is sampled by webcams is relatively low, which greatly limits data analysis.

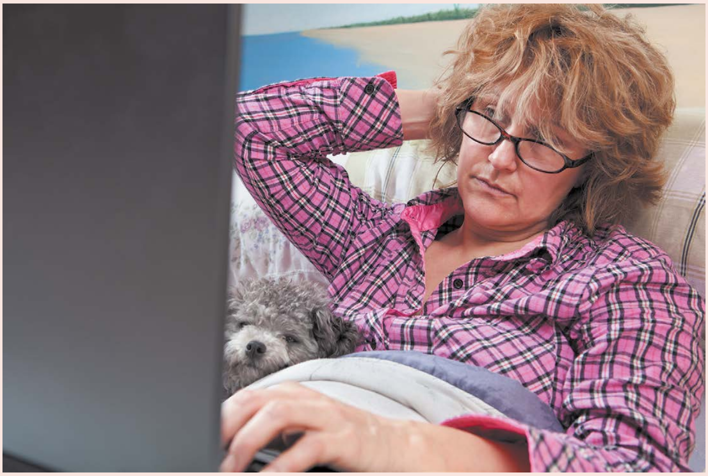

FIGURE 1.4
Some of the challenges of remote research with webcam eye tracking stem from the researcher's inability to control the test environment.

Why Do the Eyes Move?

If you take a step back just for a bit, you'll realize that when people talk about eye trackers recording eye movements, they usually take it for granted that the eyes move. Out of the hundreds of conversations I've had with people new (and not so new) to eye tracking, not once has anyone (not even John) questioned why the eyes move. They just do, right?

Human eyes, without rotating, cover a visual field of about 180 degrees horizontally (90 degrees to the left and 90 degrees to the right) and 90 degrees vertically (see Figure 1.5). Any time your eyes are open, the image of what you see is projected onto the retina. The retinal cells convert that image into signals, which are then transmitted to the brain. The cells responsible for high visual acuity are clustered in the center of the retina, which is called the *fovea* (refer to Figure 1.1). When you are looking at something directly, its image falls upon your fovea, and thus it is much sharper and more colorful than images that fall outside of the fovea.

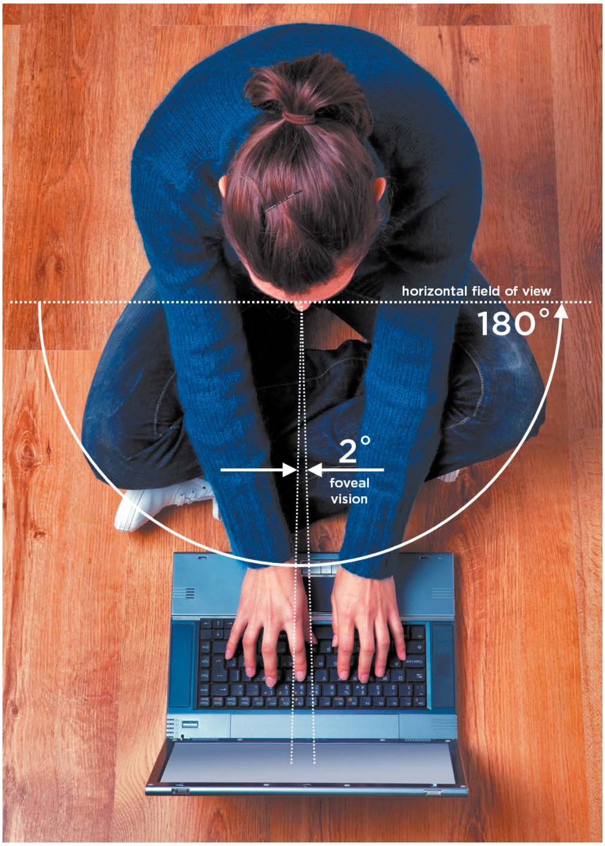

FIGURE 1.5
When looking straight ahead, humans have a visual field of about 180 degrees but only 2 degrees of it belongs to sharp, foveal vision.

The foveal area is quite small—it spans only two degrees, which is often compared to the size of a thumbnail at arm's length. Even though you typically do not realize it, the image becomes blurry right outside of the fovea in the area called the *parafovea* (2–5 degrees) and even more blurry in the periphery (see Figure 1.6). Therefore, eye movements are necessary

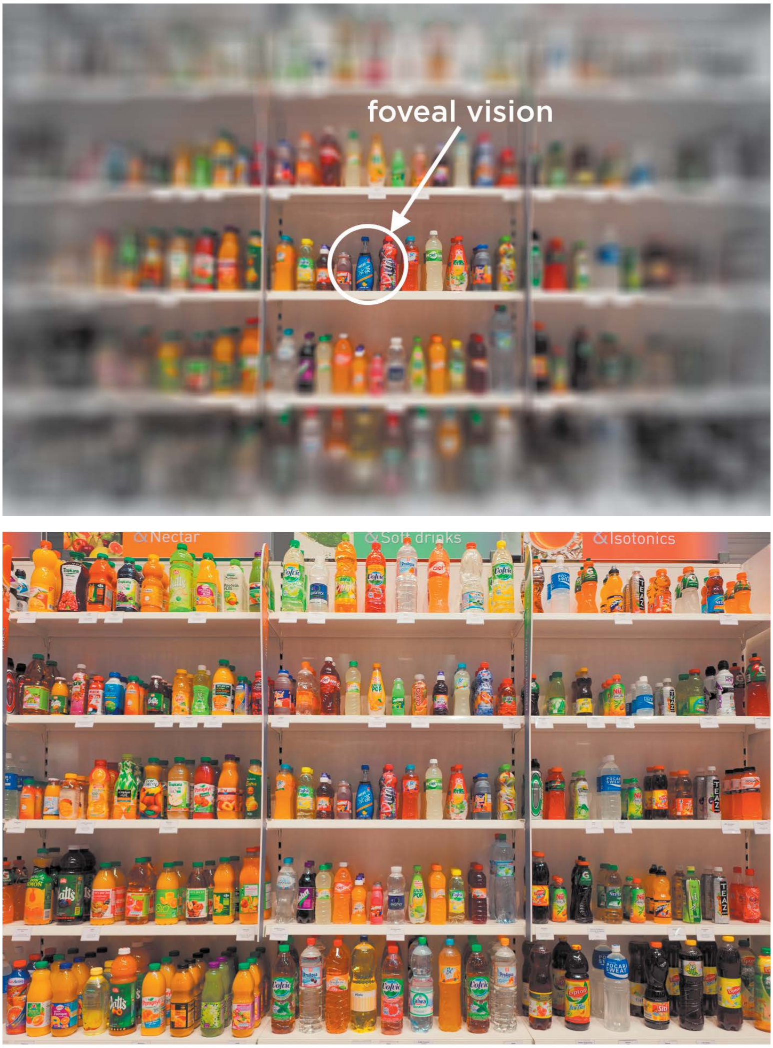

FIGURE 1.6
Top: The area that is in focus represents your foveal vision; the farther away from the fovea, the less detailed the image is. Bottom: Eye movements allow you to focus on multiple areas, giving the impression that you can see everything clearly.

to bring things into focus. This is an important, information-filtering mechanism—if everything were in focus all at once, your brain would be overloaded with information!

How Do the Eyes Move?

Your eyes jump from place to place a few times per second (three to four times, on average). These rapid movements, called *saccades*, are the fastest movements produced by an external part of the human body. To prevent blurring, your vision is mostly suppressed during saccades. Visual information is only extracted during fixations, which is when the eyes are relatively motionless and are focusing on something (see Figure 1.7). Fixations tend to last between one-tenth and one-half of a second, after which the eye moves (via a saccade) to the next part of the visual field. Although there are a few other types of eye movements, saccadic eye movements, consisting of saccades and fixations, are most common and of the greatest interest to UX research.

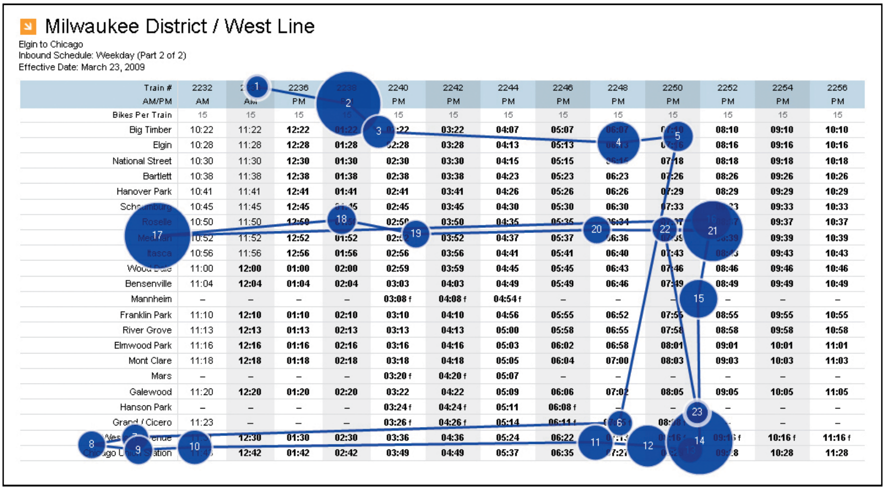

FIGURE 1.7
Gaze plot representing eye movements of a person looking up a train schedule in a timetable. Fixations are represented as dots and saccades are shown as lines connecting the dots. The size of the dot is proportional to the fixation duration.

Why Should You Care Where People Look?

A great deal of research has established that where you place your gaze is typically associated with what you pay attention to and think about,[3] especially when looking at something with a goal in mind. This is called the *eye-mind hypothesis*.

Yet, there are skeptics out there who do not think that knowing where people look can be meaningful in any way. The argument is usually, “I don’t have to look at something in order to see it,” which tends to be followed by, “I’m looking at your face right now, but I can still see the color of your sweater” or something of that nature.

You certainly could direct your attention to the periphery of your visual field. But if you wanted to see what color sweater someone was wearing, you would look directly at it for two reasons: (1) you can see things much more clearly when looking directly at them; and (2) paying attention to something and trying not to look directly at it is unnatural and requires conscious effort. Humans prefer moving their eyes when shifting visual attention, focusing on what they are trying to see. However, when people do not look at something directly, you cannot say for sure that they did not see it. Eye tracking only captures foveal vision, yielding no information about what was noticed peripherally. This is one of the limitations of eye tracking.

Another argument against eye tracking might be this: “People can look at something but not necessarily ‘see’ it.” Yes, that can happen. Close your eyes after you have been talking to someone face to face for a while and ask that person what color your eyes are. Many people will not know, although they have been looking at you (and presumably glancing at your eyes) for a while, and maybe even have known you for years. This is just one example of how you can look at an object but not necessarily register everything about it. Sometimes, you can even miss the entire object itself.

3 James E. Hoffman, “Visual Attention and Eye Movements,” in *Attention*, ed. Harold Pashler (London: University College London Press, 1998), 119–154.

To sum up this discussion, a lack of fixation does not always mean a lack of attention, and fixation does not always indicate attention, but fixation and attention coincide a whole lot. Attention is actually slightly ahead of the eyes because it plans their next destination. Once the eyes move there, attention helps allocate the processing resources to what is being fixated upon. Knowing where users' attention is directed helps the researcher evaluate and improve products, which is the focus of Chapter 2, "To Track or Not to Track."

Why Do People Look at What They Look At?

Your visual behavior is influenced by anything that *makes* you look (bottom-up attention), as well as your voluntary intent to look at something (top-down attention). Bottom-up attention is stimulus-driven. Attention is involuntarily shifted to objects that contrast with their surroundings in some way. For example, bright colors and movement can make you look at something. Things that are new and unexpected in a familiar environment can grab your attention, too.

If bottom-up factors were the only ones influencing people's attention, everyone would look at the world in the same way, regardless of what they knew and what they were trying to accomplish. This consistency would certainly make your research easier, wouldn't it? Studying different user groups and multiple tasks would no longer be necessary.

Unfortunately (but also more interestingly), this is not the case, due to the involvement of top-down factors. Top-down attention is knowledge-driven and relies on your previous experience and expectations. You intentionally choose to look at information that you consider relevant to your goals.

You have probably already heard that eye movements are task-dependent. What this means is the same person will look at the same object differently if given a different task. For example, someone looking at mobile phone packaging will generate a different gaze pattern when trying to determine the brand of the phone than when trying to find out if the phone will allow him to browse the Web (see Figure 1.8). It is the top-down attention that is responsible for these differences.

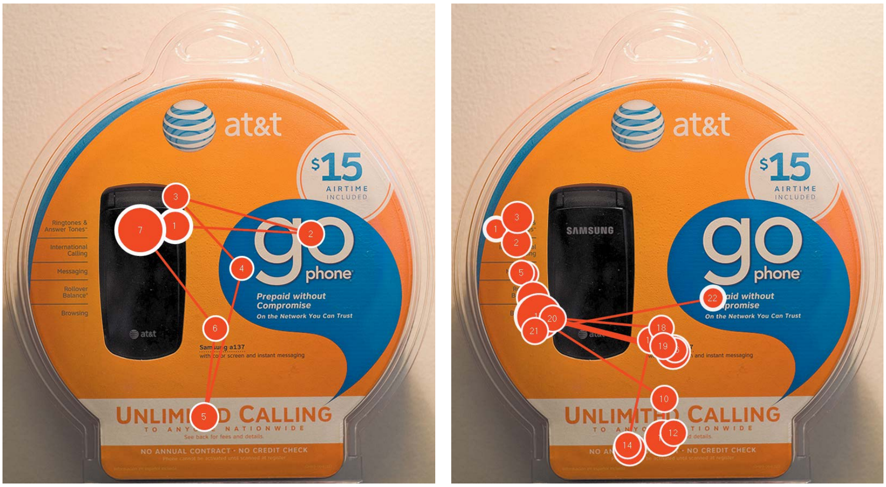

FIGURE 1.8
Left: Gaze plot of a person looking for the brand of the phone. Right: Gaze plot of the same person trying to find out if the phone offers Internet access. Notice how the same person with the same package—but a different task—produced different fixation patterns.

Applications for Eye Tracking

There are two main applications for eye tracking: as a research technique and an input device. As an input device, eye movements become control signals for a computer system, either instead of or in addition to a mouse and keyboard. People with disabilities such as ALS and cerebral palsy use gaze-controlled applications to help them communicate. Gaze interaction is also used in entertainment (for example, gaming) and is making its way into mainstream mobile applications.

The rest of this book focuses on how eye tracking can be applied to research, specifically UX research—an investigation of how people experience products, interfaces, services, or even their surroundings. UX research can be divided along two dimensions: the target of the research and the scope of the research (see Table 1.1). This breakdown results in four categories: Engineering Psychology Research, Design Research, User Research, and Design Evaluation. Eye tracking can be used in all four types of research, but this book is primarily about eye tracking for formative (or diagnostic) and summative design evaluation (the shaded bottom-right quadrant of the table).

TABLE 1.1 BREAKDOWN OF UX RESEARCH

Scope of the Research	Target of the Research: Users	Target of the Research: Design
Generalizable across products (theory-driven and often academic)	**Engineering Psychology Research** *Goal:* To understand the capabilities and limitations of human perception, cognition, and movement control. *Sample research questions:* What is the effect of noise on human performance? What factors impact how well people can switch between different tasks? How do visual scan strategies of expert pilots differ from those of novice pilots?	**Design Research** *Goal:* To understand how design types or elements impact the user experience. *Sample research questions:* Which layout for an online form is most efficient? How do people view search results pages? Do icon labels improve user performance?
Product-specific (atheoretical and doesn't generalize beyond one product)	**User Research** *Goal:* To learn more about the needs, preferences, motivations, and processes of the users or potential users of a particular product. *Sample research questions:* What are the current practices and preferences of patients learning to use a self-injection device? Why do users of a particular mobile phone go to the customer support website and does the site meet their needs?	**Design Evaluation** *Goal:* To evaluate a particular product based on the user experience it creates. *Sample research questions:* How can this product be improved to provide a better user experience? (formative research) How does the new interface compare to the old one? (summative research)

Tool or Method?

When I was learning how to drive, a friend took me to a large parking lot by a mall. It was a late evening, so the lot was almost empty. There he taught me how to start and stop the car, how to speed up, slow down, steer, and use the turn signal and mirrors. By the end of the night, I was pretty good at it. However, before I could take my Dodge Intrepid on real streets (and unleash its true potential), I had to learn how to drive according to the rules of the road, both written and unwritten.

By the same token, just because you can operate an eye tracker does not mean you can or should use eye tracking in your research. Knowing how to turn on the system and where to click to start a recording or obtain a heatmap is not sufficient (see Figure 1.9). You need to know how to apply

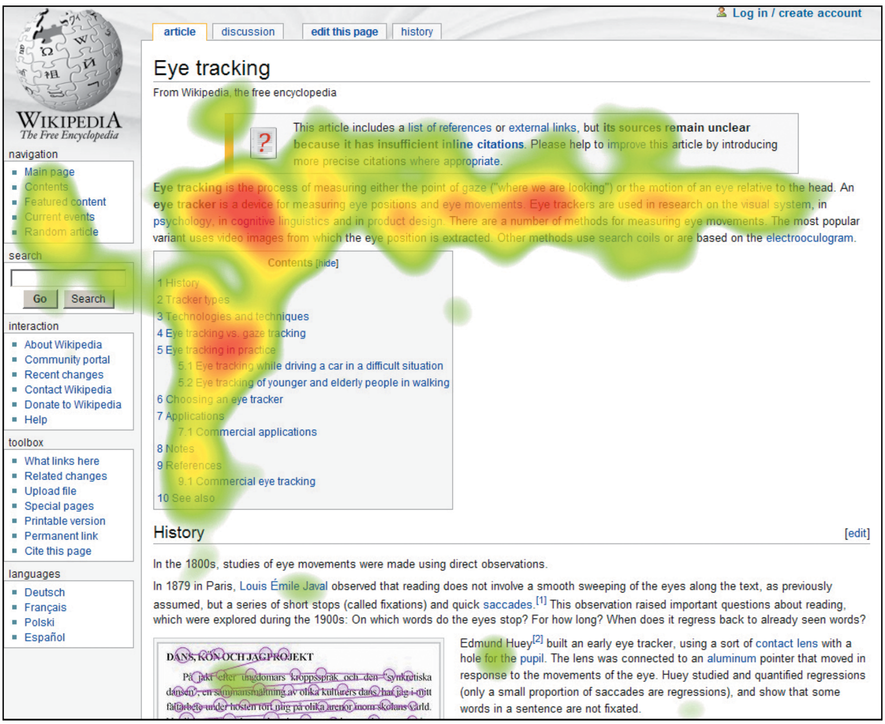

FIGURE 1.9

A heatmap is a commonly used (and misused) eye tracking data visualization. It represents the values of a variable (for example, gaze duration) as colors, where the amount of "heat" is proportional to the level of the represented variable.

eye tracking in a way that yields valid and useful insights. This knowledge encompasses planning, preparing, and conducting studies, as well as analyzing and interpreting the collected eye tracking data. Without a sound method, you are in danger of producing meaningless information and attractive but likely misleading visualizations.

When you think about eye tracking, you should consider both the tool—the eye tracker—and the methods that allow you to put the tool to good use. Learning how to operate the tool is not difficult because modern eye trackers are much easier to use than those manufactured even a couple of decades ago. Each eye tracker is accompanied by a detailed manual, and additional training is often provided by the manufacturer.

Information on methods, on the other hand, is not as readily accessible. There are several articles and book chapters available, but no one source that consolidates all the knowledge. What motivated me to write this book was the high demand for sound eye tracking methods specific to the field of user experience. But before we get to the methods, let's first discuss how to decide whether or not to use eye tracking in the first place. And that's what Chapter 2 is all about.

Summary

- Eye tracking is the process of determining where someone is looking. It can also measure the characteristics of eye movements and of the eye itself.
- Eye tracking is usually conducted with the help of a device called an eye tracker. Most commercial eye trackers work by emitting near-infrared light to determine gaze location based on the relative position of the pupil center and corneal reflection.
- Human eyes jump from place to place a few times per second. The purpose of these jumps, also known as saccades, is to bring visual stimuli into the fovea (a small area of the highest visual acuity on the retina) and thus into focus. Information is extracted during fixations, which are short pauses in-between the saccades.
- Foveal vision spans only the central two degrees of the visual field. The farther away from the fovea, the more blurry and colorless the image becomes.
- Even though eye tracking only captures foveal vision (what we are fixating/focusing on directly), it provides useful information about visual attention because, in most cases, fixation coincides with attention.
- Saccade destinations (i.e., where the next fixation is going to occur) are selected based on a combination of bottom-up and top-down cognitive processes. In other words, where you look depends on the properties of what you are looking at, as well as your goals, experiences, and expectations.
- This book focuses on the application of eye tracking for UX research (design evaluation in particular) rather than for human-computer interaction. Both the device and method aspects of eye tracking are discussed, with a special emphasis on the methods, which help put the eye tracking devices to good use, leading to valid and useful results.