Best practices in eye tracking research

Benjamin T. Carter, Steven G. Luke⁎

Brigham Young University, United States of America

A R T I C L E   I N F O

Keywords:
Eye tracking
Eye movements
Best practices
Open Science
Pupillometry

A B S T R A C T

This guide describes best practices in using eye tracking technology for research in a variety of disciplines. A basic outline of the anatomy and physiology of the eyes and of eye movements is provided, along with a description of the sorts of research questions eye tracking can address. We then explain how eye tracking technology works and what sorts of data it generates, and provide guidance on how to select and use an eye tracker as well as selecting appropriate eye tracking measures. Challenges to the validity of eye tracking studies are described, along with recommendations for overcoming these challenges. We then outline correct reporting standards for eye tracking studies.

The present paper is as a basic primer on effective, replicable eye tracking research for new researchers employing the method, as well as editors and reviewers who are unfamiliar with it. Eye tracking is a rich experimental method that has seen a surge in use in recent years (see Fig. 1). With this proliferation comes an increased potential for misuse; lack of information and poor training can result in poor study design, inappropriate analysis, and inadequate reporting. In recent years a reproductibility "crisis" has gripped psychology and many other scientific disciplines. According to one estimate only 37% of studies have replicable results (Open Science Collaboration, 2015). Eye tracking studies are not exempt from this trend and can produce non-replicable results for a number of different reasons. This work hopes to prevent some of the more avoidable mistakes by providing an adequate introduction to basic eye-tracking theory and practice. This will benefit new researchers by facilitating proper experimental design, execution, and documentation. This paper will also aid reviewers and editors by providing a condensed reference discussing the essential elements of an eye tracking study, their importance, as well as common pitfalls that must be avoided.

To accomplish these goals, the paper contains the following sections: What is eye tracking?, a brief definition of eye tracking, its history and development, and its usefulness; Basics of eye anatomy and eye movements, a discussion of the fundamentals of the biology of vision and eye movements (For those researchers who want to use eye tracking in clinical or neuroscience research, Appendix A provides more in-depth information about the physiology of eye movements); Eye trackers, a guide on how to select and use an eye tracker, including ensuring data quality and selecting appropriate eye tracking measures; Creating a valid eye tracking study, advice on how to avoid the most common threats to validity in eye tracking; and Reporting eye tracking research, a guide on how to write the methods section of an eye tracking paper. Finally, there are sections with additional information for readers interested in Pupillometry in which the use of eye trackers to measure changes in pupil size is discussed, or who want to know more about how eye tracking is used in tandem with EEG or fMRI.

It is impossible to cover each of these topics in detail in a single article. Throughout the paper, readers are referred to one or more of the following book-length works for more comprehensive guidance: Bergstrom and Schall (2014), Bojko (2013), Conklin et al. (2018), Duchowski (2017), Godfroid (2020), Holmqvist et al. (2011), Klein and Ettinger (2019), or Wade and Tatler (2005). Some of these are general introductions to eye tracking. Others are focused on a specific discipline but should contain valuable information and guidance for any researcher. There are also books and published papers that illustrate how eye tracking has been or can be used or provide guidance for using eye tracking in specific disciplines or topics, including reading (Rayner, 2009; Schroder et al., 2015), economics (Lahey and Oxley, 2016), infancy and developmental research (Feng, 2011; Gredebäck et al., 2009; Hessels and Hooge, 2019; Oakes, 2012), learning (Alemdag and Cagiltay, 2018; Conklin and Pellicer-Sánchez, 2016; Godfroid et al., 2013; Hyönä, 2010; Lai et al., 2013), memory (Hannula et al., 2010), affective disorders (Armstrong and Olatunji, 2012), autism (Chita-Tegmark, 2016; Falck-Ytter et al., 2013; Papagiannopoulou et al., 2014), diagnosis (Brunyé et al., 2019), decision making (Fiedler et al., 2019; Orquin and Loose, 2013), neurological and neuropsychiatric disorders (Anderson and MacAskill, 2013; Itti, 2015; Molitor et al.,...
1. What is eye tracking?

Eye tracking is an experimental method of recording eye motion and gaze location across time and task. It is a common method for observing the allocation of visual attention. A thorough history of eye tracking research is provided by Wade and Tatler (2005), which we summarize briefly here. The origins of eye tracking can be traced to Charles Bell, who first ascribed eye movement control to the brain, classified eye movements, and described the effect of eye movement on visual orientation (Bell, 1823). This defined a physiological connection between the eyes and the nervous system, connecting their motion to neurological and cognitive processes and thereby opening a potential window into the inner workings of the mind. Over the next century, various methods were developed to enable the objective measurement of eye movements. For example, Delabarre (1898) developed a method of recording eye movements via pins mechanically linked to a plaster-of-Paris ring placed upon the cornea. Buswell (1935) developed a method of capturing the corneal reflection of a beam of light onto film through the use of prisms. Yarbus (1967) developed a corneal lens that would attach via suction, which was used in his classic experiments on perception. No matter the method, historically, eye tracking was expensive and effortful, requiring a researcher to directly observe and catalogue individual participant behavior. This served as a barrier to many researchers and limited the rate of research.

Thankfully, improvements in eye tracking technology have made eye tracking more affordable and user friendly for both participant and researcher. Video-based eye trackers can determine the direction of gaze with a high degree of accuracy by measuring the position of the corneal reflection of an infrared light relative to the pupil. These can be found in both table and head mounted configurations and allow for eye tracking in real time, enabling a much wider range of experimentation than was previously possible. The development of better and more adaptable methods of eye tracking has enabled more and more researchers to conduct eye tracking research. As a result, the use of eye tracking in research has exploded over the past 20 years across several disciplines (see Fig. 1).

A majority of the publications represented in Fig. 1 are within the field of psychology (58.13%), primarily experimental psychology (21.72%). However, eye tracking technology is used across a wide variety of disciplines, including medicine and health care (32.75%), neuroscience (17.86%), mathematics and computer science (13.51%), other social sciences and education (12.94%), engineering and technology (11.89%), linguistics (9.39%), biology and agriculture (6.52%), physics and chemistry (3.32%), business and law (2.87%), and environmental science (1.83%).

1.1. Why track the eyes?

An eye tracker measures where, how and in what order gaze is being directed during a specific task. The structure of the eye limits high acuity vision to a small portion of the visual field (called the fovea; see Anatomy of the eyes section, below). As a result, there is a strong motivation to move the eyes so that the fovea is pointed at whatever stimulus we are currently thinking about or processing. This is known as the eye-mind link (Just and Carpenter, 1980; Rayner, 2009; Rayner and Reingold, 2015), and makes eye tracking a reliable tool for exploring questions concerning the allocation of visual attention.

Where we look, and for how long, is influenced by cognitive processes beyond attention, such as perception, memory, language, and decision making. While the link between eye and mind is not absolute (Anderson et al., 2004; Murray et al., 2013; Pickering et al., 2004; Reichele et al., 2010; Steindorf and Rummel, 2019), it is generally true that the eyes reflect mental processing of whatever we are looking at in any given moment. This makes eye tracking broadly applicable to most research that explores mental processes. Because of its high temporal sensitivity, eye tracking can provide a moment-by-moment insight into unfolding cognition rather than simply revealing the final outcome. Further, eye movements are largely outside of conscious control, i.e. while individuals may choose what to look at and when, the finer details of that movement are largely reflexive; individuals are routinely poor at remembering specifically where they looked (Clarke et al., 2017; Kok et al., 2017). This means that eye tracking can tap into non-conscious processing.

Because eye movements are controlled by an extensive and distributed system (see Fig. 3 in Appendix A), they can become perturbed when the brain is damaged or disordered (Castellanos et al., 2000; Dong et al., 2013; Huettig and Brouwer, 2015; Motic et al., 2015; Roberts et al., 2012; Samadani et al., 2015; Wang et al., 2015). Perturbations in the eye movement control network of the brain produce unique and measurable signs, making eye movements useful for characterizing lesion location and breadth (Leigh and Zee, 2015). As such, eye movements have the potential to be used as diagnostic criteria (Carter and Luke, 2018; Itti, 2015). Researchers interested in using eye tracking in clinical populations should make themselves aware of the physiological bases of eye movement control (see Appendix A).

2. Basics of eye anatomy and eye movements

A basic understanding of the physiology of the eye and of how the eye moves is useful for running a successful eye tracking study. The section that follows provides basic information about the structure and movement of the eyes that is important for any user of eye tracking technology to know. Researchers who desire a more thorough description of eye movement physiology and classification, including the neural and motor systems underlying vision and eye movement control, should see Appendix A.

2.1. Anatomy of the eyes

The eye functions to gather, focus and transduce light. In many respects it resembles a camera with an aperture, lens, and photosensitive
region. Light enters the eye via the pupil, whose diameter controls the amount of light entering the eye and the resulting image intensity. The cornea and lens focus this light on the retina (Platter, 1583; Wade and Tatler, 2005), forming an inverted but clear image (Kepler, 1604, 1611; Wade and Tatler, 2005). Fine detail and color vision are primarily a product of a depressed structure in the center of the retina called the fovea centralis, where color sensitive photoreceptors—cones—are most heavily concentrated. Light that falls on the fovea forms the center of the visual field and the densely packed receptors within this region encode a high level of detail. In other words, the fovea captures a detailed image of whatever the eyes are currently pointed at. The fovea is quite small at 1.5 mm in diameter (Kolb, 1995), with the most sensitive region only 250–300 μm in diameter. This represents only 1°20′ of visual angle (Hendrickson, 2009) of the available 140′ per eye (Clark and Kruse, 1990). This is approximately the size of a thumbnail at arm’s length. Outside of the fovea is a region called the parafovea, where acuity is weaker but some information can be gleaned. Outside of this is the periphery, which detects only low-frequency visual information (Rayner, 2009; Rayner et al., 1981).

Because the fovea is so small, the eye must move across the visual environment to gather information about it. This movement is coordinated by a complex network of structures found throughout the cortex and brainstem (see Appendix A, below). Once gathered, visual information is conveyed to the central nervous system via the optic nerve (CN II), which then passes this information on to the thalamus, and then the occipital cortex. At this stage, visual information is assembled into an image, which is then passed on to other cortical centers within the parietal, temporal and frontal lobes for interpretation and reaction.

2.2. Eye movements: fixations and saccades

A fixation is a period of time during which the eyes are fixed on a visual target, perception is stable, and the eyes are taking in visual information (Rayner, 2009). As noted above, the fovea is small. The eye is unable to acquire high quality information from the entire visual field in a single fixation, so it is necessary for the eyes to move frequently. Most fixations are relatively short as a result. Fixations vary in length depending on a variety of factors, such as the nature of the visual stimuli, the task’s purpose and complexity, and the skill and attention of the individual, but they generally last 180–330 milliseconds (Rayner, 2009). Saccades are ballistic movements of the eye from one fixation to the next (Rayner, 2009; see Fig. 2 for examples in reading and scene viewing). During saccades, visual input is suppressed, so that when our eyes are making a saccade we are effectively blind (Burr et al., 1994; Castet et al., 2002; Rolfs, 2015). Saccade velocity and duration are a direct function of the distance traveled (Bahill et al., 1975). Saccades vary in size and duration according to the task at hand. A typical reading saccade is small (a 2′ rotation) and lasts about 30 milliseconds, while saccades in scene perception are generally larger (about 5 degrees of rotation) and last 40 to 50 milliseconds (Abrams et al., 1989; Rayner, 1978).

2.3. Other types of eye movements

While saccades are the most commonly tracked form of eye movement, there are other forms of movement that researchers are sometimes interested in. These include some movements that, like saccades, can be made deliberately, such as smooth pursuit (following a moving visual target) and vergence (bringing the eyes together or apart as a visual target moves closer or further from the participant). Other ocular motion is not subject to voluntary control. For example, pupil diameter is modulated by the antagonism of the parasympathetic and sympathetic nervous systems (see Pupillometry section for more information). Other reflexive movements include the optokinetic response (the smooth pursuit of an object as it travels through the environment followed by an immediate return of the eye to its original position; Distler and Hoffmann, 2011) and vestibulo-ocular reflex (the movement of the eye to maintain a stable retinal image due to vestibular activation; Hess, 2011). Even during a fixation, when perception appears stable, the eye continues to move, having both tremor, drift, and microsaccades (Duchowski, 2017; Krauzlis et al., 2017).

3. Eye trackers

This section is focused on issues surrounding eye tracking equipment, with particular attention paid to selecting and using this technology.

3.1. How do eye trackers work?

Most modern eye trackers are video-based. They shine some light source into the eye, usually an infrared light that is invisible to humans. This light produces a reflection on the cornea (see Anatomy of the eyes section, above) that is identified by the eye tracking software. The center of the pupil is also identified by the software. Then a calibration is performed, where the participant is instructed to look at a series of points at known locations on the screen. This calibration is tested in a validation stage. If the calibration is good, the point of gaze (where the participant is looking) can then be estimated with a high degree of accuracy from the relative positions of the pupil and corneal reflection.

3.2. Selecting an eye tracker

There are a wide variety of commercially available eye trackers. The two most prominent manufacturers are Tobii (https://www.tobii.com/) and SR Research (https://www.sr-research.com/), and researchers are in the process of developing open access software to turn webcams into low resolution eye trackers (Semmelmann and Weigelt, 2018).

Before deciding on an eye tracker, it is important to consider what the tracker will be used for, as different systems are more suitable for some uses than others. Trackers vary in their speed of data acquisition. The sampling rate of an eye tracker is measured in Hertz (Hz). The fastest commercial eye trackers record eye position up to 2000 times per second (2000 Hz), while wearable eye tracking glasses might only sample 50 times per second (50 Hz). If millisecond accuracy is needed (as when looking for a temporal effect with a small effect size, or when the experiment involves gaze contingent display changes), a higher sampling rate is preferable; when the sampling rate is lower, more data must be collected to average out temporal sampling error (Andersson et al., 2010). If a user is mostly interested in recording where participants looked, a lower sampling rate is usually acceptable.

Some trackers require that the head be stabilized via a chin rest, while for others the head is unsupported. A chin rest increases accuracy from the relative positions of the pupil and corneal reflection. Some trackers make the head stabilized via a chin rest, while for others the head is unsupported. A chin rest increases accuracy of measurement and should be used when knowing the precise gaze position is important for the study (e.g. when studying small eye movements, or in reading studies when knowing which word was fixated is important). Trackers that do not rely on a chin rest still provide acceptable levels of accuracy for most purposes (see Niehorster et al., 2018 for some recommendations on using eye trackers without a chin rest), and chin rests can be problematic in some situations, such as when working with infants or children (see Schlegelmilch and Wertz, 2019, for other considerations on using eye tracking with infants).

Many eye trackers are stationary, which works well for laboratory use, but some are portable (such as the SR Research Eyelink Portable Duo) while others are designed as mobile eye trackers that can be worn while participants engage in everyday tasks (such as the Tobii Pro Glasses; see Lappi (2015) and Niehorster et al. (2020b) for some recommendations on how to use mobile eye trackers). Typically, increased mobility means decreased precision and accuracy (see Ensuring data quality section, below). Also, some trackers follow both eyes, while others track only one (and still others can be configured to track one or two eyes). As the eyes move together under most circumstances, tracking both eyes is typically not essential. See Hooje et al. (2018) for a comparison of monocular and binocular tracking.
3.3. Using an eye tracker

Eye tracking can be employed in two different ways: in a diagnostic or interactive manner (Duchowski, 2017). Diagnostic use refers to simply recording eye position throughout an experiment to determine where a participant looked, for how long and in what order. This can be done with faces, scenes, text, video, web pages or any other visual stimuli. Most of the time, researchers will use eye trackers in this mode. Eye tracking can also be used interactively. The term “interactive” brings to mind the use of eye tracking devices by quadriplegics to control a cursor. While researchers will not often use an eye tracker in a purely interactive mode, they can take advantage of the high temporal and spatial sensitivity of (certain) eye trackers to design studies that use participants' gaze position to trigger preprogrammed responses by the experimental paradigm. Changes to the display that are triggered by eye movements in this way are commonly referred to as gaze contingent display changes. A simple example of a gaze contingent display change would be only revealing a picture on the screen after the participant has focused on a fixation cross for 500 ms at the start of the trial. A more complex interactive design commonly used for research purposes is the moving window technique (McConkie and Rayner, 1975; Schotter et al., 2012). In this paradigm, a “window” is created around wherever the eyes are currently fixated so that items falling within the center of gaze are visible but those outside the center are obscured or changed, preventing the participant from gathering useful information from the periphery. As the eyes move, the window moves with it, so that readers can only see a certain number of letters at any given time. By varying the size of this window and measuring any disruption in eye moves, this method has been used to explore the perceptual span, the area of a stimulus from which useful information is obtained. This is typically not limited to the fovea, but can include the parafovea and even some of the periphery. Similar methods can also be applied to pictures of visual scenes (Nuthmann, 2013).

Fig. 2. Basic eye movements are known as saccades (blue arrows in A, yellow arrows in B). Saccades are quick jumping movements the eye makes as it traverses from one location to the next. In between these saccades are fixations (represented by red circles). Fixations are periods of time when the eye is focused on a single point, such as a word in a sentence or object in a scene. These pauses allow the eyes to take in visual information. A. When reading, fixations progress from left to right (in English). Some words are skipped. A regression in reading is an eye movement to a previous region of the text. B. When viewing a scene, fixations generally focus on meaningful or visually salient parts of the image, and there is more variability in the direction and amplitude of saccades. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
3.4. Ensuring data quality

Eye tracker data quality is usually described in terms of its accuracy and precision. Eye tracking data is accurate if the measured eye position corresponds to the actual eye position, while eye tracking data is precise if it provides consistent measurements of eye position (Reingold, 2014). These two terms are closely analogous to the broader concepts of validity and reliability, respectively. Eye tracker manufacturers provide accuracy and precision information about their devices. However, these values represent a best case scenario, and the accuracy and precision of the data provided by an eye tracker can vary significantly depending on a number of factors beyond the type of eye tracker being used, including the eye tracker setup, experimental procedures, and the behavior and physiology of the participant, among others (Blignaut and Wiium, 2014; Ehinger et al., 2019; Hessels et al., 2015; Hutton, 2019; Nyström et al., 2013).

At the beginning of every eye tracking session, a calibration is conducted (see How do eye trackers work? section). This is the stage where the experimenter has the greatest control over the quality of the eye tracking data, and extra care taken to ensure that the calibration achieves acceptable levels of accuracy and precision will pay dividends later. Because data quality can degrade over time during an eye tracking session, it is good practice to build multiple calibrations into the session. If greater accuracy is needed for a particular study (because regions of interest are small or because the study is gaze contingent), recalibrations should occur more often.

Knowing how eye trackers work makes it easier to ensure good calibrations that permit the collection of accurate, precise data. For example, if the infrared light is not getting to the eye in sufficient quantity, tracking will not work well (or at all, in some cases). This can occur if a participant is wearing eyeglasses with a strong prescription or if the glasses are dirty or have a tint or an anti-glare coating. It also occurs if the participant is too far away or the infrared light point is sourced in the wrong direction. Since point of gaze estimation is difficult or impossible without the pupil, factors that make the pupil hard to identify will reduce data quality. Such factors could include partially occluded pupils (as in sleepy participants) or other areas of darkness around the eye, such as dark eye lashes (or eye lashes darkened by makeup), that the camera might mistakenly assign as part of the pupil. Furthermore, anything that produces other reflections, such as glare on glasses or a reflection on the cornea from some extraneous light source, will obscure the true corneal reflection and also interfere with tracking. If the position of the corneal reflection is unpredictable variable, which happens as a result of gas-permeable (hard) contact lenses shifting in the eye or from bifocal or multifocal glasses or contacts, tracking will be impossible. For more information on factors that influence eye tracking data quality see Blignaut and Wiium (2014), Dalrymple et al. (2018), Ehinger et al. (2019), Hessels et al. (2015), Hutton (2019), and Nyström et al. (2013).

3.5. Eye tracking data

In its raw form, eye tracking data is a series of samples. Each sample contains the point of gaze estimate for one or both eyes as an x and y screen position in pixels. Other information might also be included, depending on the tracker used and the experimental design. The number of samples per second depends on the sampling rate. For some research (e.g. measuring pupil size or exploring smooth pursuit eye movements) it is necessary to work with these raw sample data, but under most circumstances it is neither necessary nor desirable to do so. Instead, the raw sample data is processed to identify fixations, saccades, blinks, and lost data. During this processing, an individual sample will be assigned to a fixation if it belongs to a group of samples that are relatively spatially close to each other. A sample becomes part of a saccade if temporally adjacent samples are farther apart spatially, indicating that the eye was moving with some velocity. Commercially available eye tracking software will usually do this processing automatically. There is a multitude of open access software packages designed to process eye tracking data as well (de Uraibain et al., 2015; Frame et al., 2019; Hessels et al., 2017; Houpt et al., 2018; Komogortsev and Karpov, 2013; Leppänen et al., 2015; Niehorster et al., 2020a; Pedrotti et al., 2011; Scurr et al., 2014; Sogo, 2013; Špakov et al., 2019; Startsev et al., 2019; van Renswoude et al., 2018; Wass et al., 2013; Weber et al., 2018; Zemblay et al., 2019; Zhang and Hornof, 2011). Some of these are designed for general use, while others are tailored to specific purposes. Every software package, whether commercial or open access, parses the data using a different algorithm, so it is important to clearly report what software was used and how the parameters were set (see Results section, below).

3.5.1. Regions of interest

In many eye tracking studies, researchers want to know how long or how often participants looked at a particular part of a stimulus, such as a particular word in a sentence, an object in a scene, or the eyes of a face. When this is the goal, researchers should create a region of interest (also known as an area of interest or interest area) that encompasses part(s) of a stimulus. Most eye tracking software allows the user to predefine regions of interest. After the data is collected, the software further processes the eye tracking data to provide a description of how each participant interacted with these regions of interest, including variables such as first fixation time and duration, number of fixations, number of visits, total time spent, and others. The way that regions of interest are defined can have consequences for the outcome of an experiment (see Hessels et al., 2016).

3.6. Eye tracking measures

Eye tracking can provide a multitude of different dependent measures for analysis (see Holmqvist et al., 2011, for an exhaustive list). This makes eye tracking a highly flexible technology that can be applied to many different research questions and experimental tasks, but the wealth of possible measures can also lead to problems (see Selecting measures section). In this section, we describe some common measures derived from eye tracking data and when to use them. Keep in mind that not all researchers label these measures in the same way, so be sure to check the definitions of the measures provided in a study. Also note that the descriptions below are meant to apply broadly across many types of stimuli. When studying reading, there are several reading-specific measures that should be used (Rayner, 2009).

As noted above, fixations and saccades are the basic unit of data for most analyses. If we are comparing how participants view different images or read different texts, we can focus our analyses on aggregate measures of the duration and location of fixations (average fixation duration, number of fixations) and saccades (average saccade amplitude). These measures can reveal how participants interacted with the stimuli on a global level. A few examples from different domains will illustrate this. In their meta-analysis of eye movements by experts, Brams et al. (2019) found that when experts perform visual scanning tasks, they have a consistently higher average fixation duration than do novices; their pauses are longer. Brams et al. interpreted this (in combination with other measures) as indicating that experts spend more time focused on the relevant parts of the stimulus, while novices spread their attention out across multiple fixations. As another example, in reading Rayner et al. (2006) observed that when letters within the words of a text are jumbled, readers on average make longer fixations and more of them. The increase in fixation durations and number of fixations together paint a picture of slower, more laborious reading, indicating increased processing effort when reading texts with jumbled letters. In a study of facial emotion recognition in individuals with Anorexia Nervosa, Phillipou et al. (2015) observed that the Anorexia group made more and shorter fixations when viewing faces than did the control group. This pattern of behavior represents a series of brief glances at the faces, a strategy dubbed hyperscanning (Horley et al.,
that suggests avoidance and is sometimes observed in clinically anxious populations. Castellano et al. (2009) observed that participants made more fixations when memorizing a scene than when searching for an object in that scene, but their average fixation duration and saccade amplitude did not differ across these two tasks. This pattern indicates that viewing task affects where participants look in a scene, with fewer fixations in search indicating that less of the scene is explored, but the lack of difference in fixation durations and saccade amplitudes suggests that the moment-by-moment process of interpreting the fixated portion of the scene does not differ much between tasks.

If one or more regions of interest (see Regions of interest section) are defined for a stimulus, this introduces more possible measures. These measures can be broadly categorized temporally, according to the stage of processing they index: early or late measures. Early measures tap into the initial stages of processing, and include measures such as first region of interest fixated, time to first fixation (how long before a region of interest is first fixated during a trial), and duration of first fixation. For example, in a study where participants looked for changes in faces, Thompson et al. (2019) found that participants’ first fixation was usually to the eyes, even when they were informed beforehand that they should look at the mouth to complete their task. They interpret this result as indicating a strong attentional bias toward the eyes when viewing faces. Time to first fixation is commonly used in visual search tasks. For example, Russell et al. (2019) found that when searching for targets in pictures of real-world scenes, individuals with autism had longer time to first fixation than did typically developing controls, indicating that they took more time to find the search targets.

Later measures include measures such as dwell time (the total amount of time spent fixating a region of interest), number of fixations in a region of interest, or proportion of fixations within a region of interest. For example, Kellough et al. (2008) observed that depressed individuals had longer dwell times on negative images than on positive ones, indicating a strong and stable attentional bias toward unpleasant stimuli in depression. In a study of shared storybook reading, Luke and Asplund (2018) defined regions of interest around the illustrations and the text of the storybook. A much higher proportion of the prereaders’ fixations fell on the illustrations than would be expected by chance (i.e. the illustrations took up 68% of the screen area but received 92% of the fixations), and the proportion of fixations on the text was less than chance would predict, indicating that children likely do not develop print awareness during shared storybook reading because they rarely look at the words. When analyzing late eye tracking measures, it is important to keep in mind that they are usually not independent of early measures; a late measure such as dwell time will include early measures such as first fixation time.

Often, studies will analyze multiple region-of-interest-based measures in order to explore the time course of attentional allocation. For example, Werthmann et al. (2011) used eye tracking to explore attentional bias toward food in overweight individuals. They used three measures, two early (first region of interest fixated, duration of the first fixation) and one late (dwell time). The overweight participants were more likely to look first to the food image, and their first fixation durations were longer, but they did not show a bias toward the food image in dwell time. They interpreted this pattern of findings as indicating that food captured but did not maintain the attention of the overweight individuals in their study. Võ and Henderson (2009) observed that when participants viewed scenes, such as an image of a kitchen, they had longer dwell times and made more fixations on regions of interest containing unexpected objects (such as a printer) or objects in unexpected locations (such as a pot hovering over the stove). However, the time to first fixation did not differ, suggesting that objects which violate expectations maintain attention but do not necessarily capture it.

This distinction between early and late measures is especially prominent in reading studies. In reading, early measures include word skipping probability, first fixation duration, and gaze duration (the sum of all fixations on a word the first time it is encountered). Late measures include measures such as regression probability, dwell time (also called total reading time or total time), and total fixation count. As an example, a study by Knickerbocker et al. (2019) found that emotion-laden words (e.g. birthday, funeral) have a processing advantage over neutral words (i.e. they are read faster), but the time course of this advantage depended on the valence of the word (positive or negative). Specifically, positive emotion-laden words showed advantages in the early measures of skipping probability and first fixation duration as well as in the late measures, while negative emotion-laden words did not show an advantage over neutral words until the later measures of total time and regressions in. This difference indicates that the effect for negative words arises later in the time course of reading. See Conklin et al. (2018) for a thorough discussion of eye tracking measures in reading.

4. Creating a valid eye tracking study

When steps are taken to ensure data quality (see Ensuring data quality section), eye trackers provide highly reliable data (Carter and Luke, 2018; Henderson and Luke, 2014). While reliability is a prerequisite for good data, it is useless without validity. Ensuring that eye tracking data is valid requires careful study design and appropriate analysis. There are many considerations that go into designing a good eye tracking experiment. In this section, we will cover the two most common challenges that new eye tracking users face: making appropriate comparisons and selecting measures. These considerations are important in any experiment, but are especially relevant to eye tracking studies. Readers interested in a more thorough discussion should consult Orquin and Holmqvist (2018), who cover these challenges in detail and also list other challenges to the validity of eye tracking studies that are important to consider.

4.1. Making appropriate comparisons

Eye movements (where people look, and for how long) are known to be influenced by a variety of factors (for reviews, see Henderson, 2003, 2011; Rayner, 1998, 2009). Some of these are visual, and include the complexity and salience of the stimulus or its parts (Baddeley and Tucker, 2006; Itti and Koch, 2001; Nuthmann, 2017; Parkhurst et al., 2002; Torralba et al., 2006). In other words, stimuli that are more visually complex will receive more attention, and stimuli that are more eye-catching, because of bright colors or large objects, will attract attention. The quality of the visual stimulus also matters; stimuli that are degraded, blurry, dark, or otherwise more difficult to see will require more viewing time (Henderson et al., 2013b; Henderson et al., 2014; Mannan et al., 1997; White and Staub, 2012). And of course larger stimuli will attract more fixations than smaller stimuli. A larger stimulus is more salient, easier to see, presumably more important, and requires more fixations to take in visually. Even if fixations were distributed at random across the screen a larger stimulus would receive more of them. This effect of size applies to longer words and sentences as well as larger images and regions of interest. So, it is important to control the complexity, salience, quality, and size of visual stimuli and of the regions of interest within these stimuli.

There are a number of cognitive factors that influence eye movements as well. Stimuli that are less familiar (such as an infrequent word or a face you have only seen once) will require longer viewing times before recognition (Joseph et al., 2013; Just and Carpenter, 1980; Kliegl et al., 2006). Stimuli that are less expected are also viewed longer (Henderson et al., 1999; Luke and Christianson, 2016; Staub, 2015; Võ and Henderson, 2009). Any stimulus that is meaningful to the participant (such as words they know, a human face, or a picture of a house) will elicit a different viewing pattern than a meaningless stimulus (such as a word-like shape, a jumbled face, or an abstract painting) (Henderson and Luke, 2012; Luke and Henderson, 2013, 2016; Rayner and Fischer, 1996; Vitu et al., 1995). Even within an image, the more meaningful parts draw our attention (Henderson and Hayes, 2017;
Peacock et al., 2019). Similarly, the emotional content of words and images also affect eye movements; emotional stimuli are processed more quickly and attract more attention, at least in typically developing individuals (Knickerbocker et al., 2019; Scott et al., 2012; Stephenson et al., 2019). And even if the stimuli are the same, when participants are performing different tasks, their eye movement behavior changes; the location and duration of people’s eye movements change significantly when they receive different instructions (Henderson et al., 2013a; Kardan et al., 2015; Kardan et al., 2016; Luke et al., 2013; Navalpakkam and Itti, 2005; Nuthmann et al., 2010; Yarbus, 1967).

If the stimuli in two experimental conditions unintentionally differ on any of these visual or cognitive factors, then eye movements in these different conditions will also differ, creating a confound. An example given by Orquin and Holmqvist (2018) illustrates this problem. They cited a study that examined how readers viewed bar graphs as compared to neuroimages (pictures of the brain). An analysis of eye tracking data revealed that bar graphs received more fixations. While eye tracking could determine which stimuli received more attention, it gave little insight into why. This is because neuroimages and bar graphs differ in terms of a variety of factors, including visual complexity, salience, familiarity, and meaningfulness. When confounding factors are not controlled, one type of representation could receive more fixations than the other simply because it is more interesting or novel to the viewer. It could also receive more fixations due to difficulty of interpretation or even to differences in image size. Studies that do not carefully control for these confounds leave readers with more questions than answers. Unless the study is deliberately manipulating one of the above-mentioned factors, these factors should be controlled by the experimenter.

### 4.2. Selecting measures

An unwieldy number of measures can be derived from eye tracking data. For example, a fixation report produced by DataViewer software from SR Research ([https://www.sr-research.com/data-viewer/](https://www.sr-research.com/data-viewer/)), which describes the fixations and saccades made by each participant in the study, can include 77 different variables. These include fixation duration, number of the fixation in the trial, fixation location, saccade velocity, saccade duration, saccade amplitude, saccade latency, and saccade direction. When using regions of interest, even more measures are available, such as time of first fixation, dwell time, number of fixations within the region of interest, and number of visits to the region of interest.

The wealth of measures produced by eye tracking is one of the reasons that eye tracking is such a versatile method. At the same time, it poses a significant problem. The ready availability of so many potential dependent variables from which to choose increases the temptation to engage in “fishing.” This will likely be unintentional; novice researchers might be tempted to analyze every variable that their eye tracker can provide, not realizing that they are increasing the risk of Type I error significantly by doing so. This risk is compounded by the fact that many eye tracking metrics are highly correlated and are not independent of each other. For example, number of fixations is positively correlated with total dwell time; the more often a participant looked at a region of interest, the more time the participant spent looking at that region of interest. Likewise, dwell time is correlated with first fixation duration because dwell time includes the duration of the first fixation. The same is true for other variables such as saccade amplitude, saccade duration, and saccade velocity; larger saccades last longer and move faster.

Given the large number of potential variables, it is essential that researchers choose which variables they will analyze before the study is conducted. Such a choice should be motivated by the research questions being asked, as different measures are more appropriate for different questions. For example, if the research question is “what part of the face attracts attention first” then the early measure of time to first fixation in each region of interest would be the best variable to analyze. On the other hand, if the question is “what part of the face receives the most attention”, then the late measure of dwell time in each region of interest would be a better variable. If multiple variables are chosen for analysis, they should not be redundant; each variable should answer a different question or provide additional information about the time course of processing. Researchers who are unsure which variables to select should consult eye tracking experts and reference guides or be guided by existing research, both in and out of their field (See **Eye tracking measures** section, for more information on selecting measures). If more than one dependent variable is analyzed, as is often the case in eye tracking research, appropriate corrections for multiple comparisons (such as Bonferroni) should be performed even if the variables are independent (von der Malsburg and Angele, 2017).

The best way to restrict analyses to an appropriate set of variables is through preregistration. Preregistration, or publicly defining research questions and analysis plans before observing outcomes (Nosek et al., 2018) is a well-established practice in clinical research but has yet to become regular practice among many non-clinical scientists. Preregistration offers several benefits, including ensuring that the hypotheses, measures, and analytical methods are predefined. This helps readers to distinguish between predictive and post hoc analyses made by a study (Nosek et al., 2018). It also combats other threats to research reproducibility, such as experimenter bias, poor quality control, low statistical power, P-hacking, and HARKing (i.e. hypothesizing after results are known; this often takes the form of framing a post-hoc analysis as a prior). Preregistration is even more important for an eye tracking study than for other types of studies because of the ready availability of multiple analyzable variables. A good preregistration would specify which variables will be analyzed and provide a justification for each variable chosen. Additional discussion of preregistration can be found in Krypotos et al. (2019); Nosek et al. (2018); Pu et al. (2019).

### 5. Reporting eye tracking research

When reporting the results of an eye tracking study, some unique information is required. Many eye tracking studies fail to report enough information to ensure reproducibility (Fiedler et al., 2019). In this section, we describe some of the most essential pieces of information that must be provided, focusing on eye-tracking-specific information. A more complete list is provided by Fiedler et al. (2019). The information below is organized according to the sections where it would normally be presented in an APA-style paper. We also provide recommendations for code and data transparency.

#### 5.1. Methods

Additional information needed in the Methods section is described below, broken up into the different subsections where it will typically be found. Not all papers will include this information in the same subsection, but it should be included somewhere.

##### 5.1.1. Participants

In addition to the usual information, the participants section should contain a description of any visual inclusion/exclusion criteria. The most common such criteria include requiring normal or corrected-to-normal vision and excluding colorblind participants, if relevant. In any study, some percentage of the participants will not provide usable eye tracking data and should be excluded. This percentage should be reported, along with a justification for the exclusion (i.e. the quality threshold for data exclusion).

##### 5.1.2. Apparatus

Most methods sections in eye tracking studies include an apparatus subsection. In this section, the eye tracker setup is described. Below is a list of information this section should contain:

- The make and model of eye tracker (e.g. SR Research EyeLink 1000 Plus, Tobii Pro Spectrum).
The sample rate of the eye tracker in Hz (e.g. 120 Hz, 1000 Hz).

The accuracy and precision of the eye tracker.

Whether one or both eyes were tracked. If only one eye was tracked, which one (left, right, dominant)?

What was the eye tracking set up (desktop mount, remote tracking, head-mounted)? Was a chin/head rest used?

The make and model of the monitor.

The size and resolution of the monitor.

Viewing distance from the eyes to the monitor.

The stimulus size(s). The size of a stimulus's image on the retina is determined both by its actual size and its distance from the eye. For this reason, stimulus sizes are expressed not just in absolute terms (i.e. inches or pixels) but also in degrees of visual angle. Imagine a triangle, with the peak starting in the eye and the two legs extending to the sides of the stimulus (the stimulus forms the base of the triangle). The visual angle is the size of the angle formed at the eye by these two legs. For images (i.e. scenes, faces), the horizontal and vertical sizes should be reported (e.g. “all images subtended 30 by 25° of visual angle”). Image size in pixels should be reported in addition to the visual angle, although this often appears in the Materials subsection. For text, it is most common to report the number of letters that can fit into a single degree of visual angle (e.g. “one degree of visual angle included four characters”), along with the font type and size.

The software that was used for stimulus presentation and data acquisition.

5.1.3. Materials

In this section, the selection and manipulation of materials should be described. How were the different materials matched in terms of visual and cognitive factors that might affect eye movements (see Making appropriate comparisons section)? What are the sizes of the materials, in absolute terms (pixels, font size) and relative terms (visual angle)? How were the different stimuli positioned on the screen? What regions of interest were created, and what were their absolute and relative sizes? How were different regions of interest matched, and what other considerations went into their creation (Hessels et al., 2016)?

5.1.4. Procedure

In this section, authors should describe the calibration procedure and other steps taken to ensure accuracy and precision of data. How many points of calibration were there? What was the acceptance criterion for the calibration (e.g. “Calibration was accepted if average error was < 0.30° of visual angle (corresponding to approximately 1 character) and maximum error was < 0.50°”)? How often was the calibration conducted?

Authors should also describe how participants progressed through the experiment. Many eye tracking studies have a drift check between trials; participants must look at a particular region of the screen to proceed. What was the location of this gaze trigger on the screen? How often and how long did it appear? What triggered the start of the trial (was it automatic, or did the participant or experimenter control it)? What steps were taken to ensure data quality during the study (see Ensuring data quality section)?

5.2. Results section

Prior to reporting of results, a description of any data processing that occurred is required. Most commonly, the eye tracker software will segment the data into fixations and saccades automatically. The software used to do this should be identified and the criteria for identifying saccades described (e.g. “fixations and saccades were segmented in DataViewer with EyeLink's standard algorithm using velocity and acceleration thresholds (30°/s and 8000°/s)”). This information can usually be located in the user's manual for the software. If some other software is used, the software should be cited and the segmentation thresholds described. Additionally, any data cleaning, such as removal of outlier fixations and/or saccades or trials/fixations disrupted by blinks, should be described and justified and the percentage of data removed should be reported. If any spatial adjustments to fixations are made (i.e. minor vertical adjustments to fixation positions to account for small drift in a single line reading study), this should be reported as well. A list of the dependent variables selected for analysis, and a justification for that selection, should also be provided (see Selecting measures section).

5.3. Improving transparency

Eye tracking studies are sometimes reliant on custom software or scripts that are modified by or completely created by the researcher for both the experiment and the processing and analysis of the data. Whenever this is the case, code and scripts should be publicly available. Doing so ensures that all relevant information is reported. The data itself should also be shared whenever possible; eye tracking data is relatively more compact than other types of data, such as EEG or MRI data, and so is ideal for sharing publicly. Additional documentation related to the study should be released as well, such as prescreen information, data from pilot studies, researcher notes, scripted dialogue with participants, in short anything that could aid in reproducing the study as long as it does not divulge protected information (e.g. participant information protected by law, documents protected by copyright). For an example of an OSF page where data and documentation is shared see Carter and Luke (2019b) or Luke and Christianson (2018). For an example of a Github repository where code is posted publicly, see Carter et al. (2019).

6. Pupillometry

As noted earlier, modern video eye trackers track the eyes by identifying the pupil. This means that many eye trackers can also be used effectively for pupillometry. Pupillometry is a technique that records changes in the diameter of the pupil (for reviews, see Hartmann and Fischer, 2014; Laeng and Alnaes, 2019; Laeng et al., 2012; Mathôt, 2018; Sirros and Brisson, 2014). The size of the pupil changes in response to changes in luminance, a change that has a latency of approximately 200 ms (Ellis, 1981). The pupils also dilate in response to internal cognitive/affective processes such as shifts of attention, motivation, mental effort, and cognitive load (Beatty and Lucero-Wagoner, 2000; Einhäuser, 2017; Laeng and Alnaes, 2019; Laeng et al., 2012; Mathôt, 2018). Most psychological research focuses on these cognitive/affective distortions, which are much smaller in magnitude than pupil size changes driven by luminance (Mathôt, 2018).

Like eye tracking generally, pupillometry has become increasingly popular in recent years (see Kret and Sjak-Shie, 2019, Fig. 1). Pupilometry is used to study a wide variety of topics, including perception (Laeng and Sulutvedt, 2014), development (Eckstein et al., 2017; Hepach and Westermann, 2016), language (Engelhardt et al., 2010; Fernandez et al., 2018; Scheepers et al., 2013; Schmidtke, 2018), emotion (Bradley et al., 2008; Schmidtke, 2018), social support (Graff et al., 2019), attention and mind wandering (Franklin et al., 2013; Konishi et al., 2017; Mathôt et al., 2016; Unsworth and Robison, 2018a), mental effort and memory load (Hess and Poll, 1960, 1964; Kahneinan and Beatty, 1966; Unsworth and Robison, 2018b), memory (Bergt et al., 2018; Goldinger and Papesh, 2012) and decision making (Cavanagh et al., 2014; de Gee et al., 2014).

Because eye trackers can provide pupil size estimates along with gaze location, it can be tempting to collect and examine both sets of data in a single experiment. Generally speaking, this is not advisable unless the research requires it. Systematic errors in pupil size estimation can occur when the eyes move (Brisson et al., 2013; Gagl et al., 2011; Hayes and Petrov, 2016). Additionally, when the eyes are exploring some stimulus, such as an image, the pupils adjust to the luminance of different regions of the image. This change in pupil size begins before the eye movement is executed (Mathôt et al., 2015) and can even occur...
for covert shifts of attention that do not involve an eye movement (Binda and Murray, 2015; Mathôt et al., 2016). This means that eye movements introduce noise in the pupillometry data that can easily swamp small changes in pupil size arising from cognitive or affective factors. More complex visual stimuli of the sort used in many eye tracking studies are of non-uniform luminance and will introduce significant noise to the pupillometry data even if the eyes are not moving. For these reasons, researchers who want to use eye trackers as pupillometers are advised to instruct participants to maintain a fixed gaze in the center of the screen. Auditory stimuli are ideal for pupillometry studies because the visual input is unchanging, but visual stimuli can be used if they are simple, small enough to fit in foveal vision, and of consistent luminance within and across stimuli.

When reporting the results of pupillometry studies, the same information about the eye tracker apparatus and setup should be provided as for other eye tracking studies (see Methods section). In addition, information about how luminance levels were controlled should be provided. It is essential to describe how different stimuli were matched for luminance, especially across different conditions, as any difference in luminance across conditions represents a major confound. Luminance levels in the room where the study is conducted, and steps taken to keep these levels constant for all participants, should also be reported.

Reliable pupillometry data can be obtained from less expensive eye trackers (Titz et al., 2018). In fact, because pupillometry requires working with individual samples rather than aggregated fixation and saccade data, eye trackers with higher sampling rates provide an unyielding amount of data. Even with lower sampling rates, analyzing pupillometry data requires significant computing power. When analyzing pupillometry data, there are unique considerations (see Kret and Sjåk-Shie, 2019; Sirois and Brisson, 2014 for walkthroughs of the data analysis process). Pupillometry data usually requires baseline correction, where the pupil size is expressed as a difference from some pretrial baseline value (Mathôt et al., 2018; Reilly et al., 2019). Pupillometry data also requires significant cleaning before analysis, but fortunately some researchers have made open-source scripts available for this purpose (Hershman et al., 2019; Kret and Sjåk-Shie, 2019).

7. Eye tracking with EEG or fMRI

Eye tracking is widely used as a stand-alone research method. In combination with other technologies, eye tracking can be an even more powerful research tool. Most technologies blend seamlessly with eye tracking. In the last few years, researchers have been working on combining eye tracking with both EEG and fMRI. The marriage of eye tracking with EEG and fMRI is complicated but potentially highly rewarding. Some basic information about how these technologies are being used in unison with eye tracking is provided below for interested readers.

7.1. EEG and eye tracking

While eye tracking reveals where a person is focusing their visual attention, electroencephalography (EEG) provides a record of the neural response to that information. Putting these two technologies together, then, gives a more complete picture than either could alone. In addition, allowing participants to move their eyes freely adds ecological validity to EEG studies. However, combining eye tracking with EEG during free viewing of stimuli presents significant technological and analytic challenges (Dimigen et al., 2011; Henderson et al., 2013c; Nikolaev et al., 2016; Plochl et al., 2012; Touryan et al., 2017). Notable among these are, first, the numerous large eye-movement-related artifacts present in the EEG data and, second, the significant (and varied) overlap of brain responses arising from successive fixations.

Nevertheless, progress has been made, and many studies have used potentials time-locked to fixation onsets (fixation related potentials, or FRPs) to examine a variety of topics, including visual search (Hiebel et al., 2018; Kamienkowski et al., 2012), face processing (Buonocore et al., 2019; Guérin-Dugué et al., 2018), reading (Degno et al., 2019; Dimigen et al., 2011; Frey et al., 2018; Henderson et al., 2013c, Luke et al., 2013; Kornrumpf et al., 2016), change detection and memory (Nikolaev et al., 2013; Nikolaev et al., 2011), natural scene viewing (Dandekar et al., 2012; Fischer et al., 2013; Simola et al., 2015), and aesthetic judgments of art (Fudali-Czyż et al., 2018). Open-source code for the combined analysis of eye movement and EEG data is available (EYE-EEG; Dimigen et al., 2011).

7.2. fMRI and eye tracking

Eye tracking has also been combined with functional magnetic resonance imaging (fMRI). Like the pairing with EEG, the motivation for combining eye tracking and fMRI is that the two technologies are complementary; eye tracking provides a record of what is being looked at and fMRI records the brain’s response to that visual information. Also, the addition of eye tracking can add significant ecological validity to fMRI studies. In some ways, eye tracking and fMRI make a less obvious pairing than eye tracking and EEG. While eye tracking and EEG are both highly temporally sensitive measures, fMRI lags far behind both. This mismatch in temporal sensitivity does present some analytic challenges when combining eye tracking and fMRI. On the other hand, allowing participants to freely view a stimulus is much less disruptive to fMRI data than to EEG data.

In this technique, individual fixations are treated as events, and BOLD activation related to these fixation events is analyzed (Carter and Luke, 2019a; Henderson and Choi, 2015; Henderson et al., 2015; Himmelstoss et al., 2019; Marsman et al., 2012; Richlan et al., 2014). This technique has been applied to reading (Carter et al., 2019; Carter and Luke, 2019a; Desai et al., 2018; Henderson et al., 2016; Henderson et al., 2015; Schuster et al., 2019; Schuster et al., 2016; Schuster et al., 2015) and to scene processing (Henderson and Choi, 2015; Kuniecki et al., 2017), but researchers are beginning to apply the technique to other topics (Jiang et al., 2017).

8. Conclusion

Eye tracking is a powerful tool that can be applied to a wide variety of research questions across many different disciplines. Technological advances have made eye tracking more affordable and accessible to many researchers. With this increased accessibility comes increased risk of incorrect use. The present paper is a cursory overview of the use of eye tracking for research. Descriptions of relevant eye anatomy and basics about eye tracking technology were provided, as were recommendations for constructing valid, reproducible studies and reporting them completely. References to other eye tracking guides and many different studies and reviews were provided for researchers interested in learning more about eye tracking.

We see the explosion of eye tracking research as a positive development that will lead to significant scientific breakthroughs in many disciplines. But as with any technology or methodology, an eye tracker cannot reveal anything unless it is employed correctly by careful and deliberate users.

Appendix A. Supplemental Information for clinical and neuroscience research using eye tracking

In certain circumstances, a more in-depth knowledge of the neurological system that directs eye movements is important. This allows for appropriate participant screening and study design when a clinical population is the subject of study. Differences in ocular and neurological health and development can have a significant effect on findings in an eye tracking study, so a basic understanding of the underlying anatomy and
Fig. 3. The oculomotor network. Voluntary motor plans created by cortical regions are relayed to intermediate structures in the brainstem. These structures then coordinate and relay these plans to the oculomotor nuclei in the midbrain and pons which directly innervate the extraocular muscles. Cortical regions are as follows: Frontal eye fields (FEF), dorsolateral prefrontal cortex (DLPFC), cingulate eye fields (CEF), supplementary eye fields (SEF), medial temporal lobe (MT), medial superior temporal (MST) lobe, parietal eye field (PEF), and posterior parietal cortex (PPC). Intermediate regions are as follows: Superior colliculus (SC), nucleus of the posterior commissure (nPC), prepositus hypoglossi nucleus (PHN), rostral interstitial nucleus of the medial longitudinal fasciculus (Rimlf), interstitial nucleus of Cajal (IC), paramedian pontine reticular formation (ppRF), and central mesencephalic reticular formation (cmRF). Oculomotor nuclei: Oculomotor nerve (CN III), trochlear nerve (CN IV), abducens nerve (CN VI). Extraocular muscles: (1) Superior rectus, (2) superior oblique, (3) medial rectus, (4) inferior oblique, (5) inferior rectus, and (6) lateral rectus.
physiology of eye movements is necessary to define and justify inclusion/exclusion criteria, inform the formation of hypotheses, and guide study design and interpretation of findings. Given this, researchers interested in using eye tracking with clinical populations or in combination with other neuroscience techniques will benefit from a more in-depth discussion of the physiology of eye movements. For even more information see Duchowski (2017); Eckstein et al. (2017); Horn and Adamczyk (2012); Liversedge et al. (2011); Paxinos and Mai (2012); Wade and Tatler (2005).

In this appendix, we provide information for researchers interested in using eye tracking in clinical or neuroscience research. Below, we describe the cortical and motor processes that control eye movements in more detail.

A.1. A deeper dive into the physiology of eye movements

For researchers who are interested, the following section provides a more detailed description of how the eye moves and of the extensive neural network that controls these movements.

**Ocular action.** Eye position within the orbit is controlled by six extraocular muscles arranged in opposing pairs (Netter, 2017), see Fig. 3. Superior and inferior rectus enable elevation and depression (looking up or down). Lateral and medial rectus enable abduction and adduction (looking away from or toward midline). Superior and inferior oblique enable incyclotorsion (rotation of the top of the eye toward midline) and excyclotorsion (rotation of the top of the eye away from midline). These six muscles are innervated by the oculomotor, trochlear and abducens nerves (see Fig. 3).

The pupil acts as a simple aperture, controlling the amount of light entering the eye. Pupil diameter is a function of two muscles within the iris—the dilator pupillae and sphincter pupillae. The dilator pupillae is innervated by the long ciliary nerve and is under sympathetic control. Activation will increase pupil diameter. The sphincter pupillae is innervated by the short ciliary nerve and is under parasympathetic control. Activation will constrict pupil diameter. The resulting autonomic antagonism between the dilator pupillae and sphincter pupillae makes pupil diameter a reasonable measure of sympathetic or parasympathetic dominance.

Lens shape is a product of tension between the ciliary muscle, zonular fibers, and the inherent elasticity of the lens. As the ciliary muscle constricts, the lens becomes round, resulting in increased magnifying power. This allows the eye to bring nearer objects into focus. When the ciliary muscle relaxes, the zonular fibers stretch the lens, flattening it, and allowing more distant objects to come into focus.

**Oculomotor control.** In general, oculomotor plans are generated and modified by several different regions of the cortex known as eye fields. These include the frontal eye fields (FEF), cingulate eye fields (CEF), parietal eye fields (PEF), supplementary temporal visual area (MT), medial superior temporal visual areas (MST), and the dorsolateral prefrontal cortex (DLPFC). Motor plans are communicated to intermediate structures within the brainstem (primarily the superior colliculus and reticular formation) that coordinate these plans with the vestibular system and the cerebellum. These intermediate structures then activate three pairs of alpha motor nuclei (oculomotor nucleus, trochlear nucleus, and abducens nucleus) which innervate the six oculomotor muscles to effect eye movement (see Fig. 3).

The oculomotor network spans most of the brain, including the frontal, parietal, occipital, temporal lobes, brainstem, cerebellum, basal ganglia, thalamus, cranial nerves, and tracts. Particular oculomotor tasks will often involve additional brain regions. For example, reading involves cortical language areas (Henderson et al., 2015). This wide distribution results in a high likelihood of involvement in neurological disease pathology, so an understanding of the network and affected structures is necessary when studying patient populations. Lesions within this network will produce specific and unique symptoms. For example, damage to the abducens nucleus or nerve will impair the ability of the participant to turn their eye outward, while lesions of the DLPFC result in increased errors in antisaccade tasks (an experimental paradigm in which participants are instructed to look away from a defined point upon the presentation of a cue) but no apparent change in ocular mobility (Müri and Nyffeler, 2008). Other neurological pathologies produce more subtle changes. Parkinson’s disease, for example, has been shown to reduce ocular tremor (Gitchel et al., 2012), a nearly invisible change.

References


A.1. A deeper dive into the physiology of eye movements

Bell, C., 1823. On the motions of the eye, in illustration of the uses of the muscles and nerves of the orbit. Philos. Trans. R. Soc. Lond. 113, 166-186.


Platter, F., 1583. De corporis humani structura et usu.


