Eye movements and cognitive psychology: How can eye movements work as a window on mental processes?

Carrick C. Williams
Mississippi State University
Abstract

Eye movements have a long history of being used to study cognition. Because eye movements are tied to attention, they provide a window into on-going cognitive processes. The field has developed specific measures from the eye movement record that give an indication of different aspects of cognition. In addition, the field has evolved from simply observing eye movements to using eye movements to trigger changes in stimuli to test cognitive theories. Finally, I will present a study that was done at Mississippi State University using eye movements to study climate change education.
Psychology in general and cognitive psychology in particular is interested in the study of mental processes that for the most part cannot be observed directly. Within the field of cognitive psychology, we study processes such as attention, memory, language, problem solving, decision making, etc., by examining behavioral changes when an individual is performing a task. We also employ many techniques such as neuroimaging (e.g., fMRI), electrophysiology (e.g., EEG), eye tracking, reaction times, accuracy measures, and surveys (to name a few) in an attempt to measure a behavioral consequence of some mental process. For example in fMRI studies of cognition, researchers observe that blood flow to a particular place in the brain is tied to a particular aspect of a task and therefore, they infer that the cognitive process in that task uses that portion of the brain. Although that may be true, one is not directly observing the mental task, but rather the consequence of the mental operation in the brain. Unfortunately, the ability to make direct observations of mental processes is difficult if not impossible.

Although it is nearly impossible to observe mental processes directly, the researcher can “spy on” the ongoing cognitive processes through various techniques. In essence, this is the goal of neuroimaging and electrophysiological measures of brain activity. However, a common and simpler method for “spying” on an ongoing cognitive process is to use eye tracking. The first question of importance is why study eye movements at all? The field of Cognitive Psychology has been at the forefront of using this technology to study the processing that occurs in both text and pictures. The primary reason that cognitive psychologists study eye movements
is that they provide an observable/measurable marker for what a person is processing at a particular moment. Although this type of measure is not perfect, it does allow researchers to use a behavioral response that is linked to a cognitive mechanism. In addition, we make eye movements all of the time with little awareness that they are occurring, and thus, eye movements provide an unobtrusive (mostly) measure that does not influence an on-going task. The person viewing the display will make the eye movements without prompting; all the researcher does is measure them.

The next reasonable question then is, why do humans make eye movements in the first place? Although we have the impression that the visual world is equally clear at all points, this is an illusion. The eye is anatomically arranged so that only 2 degrees of visual angle are seen with the highest level of acuity (a degree of visual angle is equivalent to 1 cm at a distance of 57 cm from the eye; Rayner 2009). The area of the visual field corresponds to the anatomical feature of the fovea in the retina. Around the central 2 degrees and extending approximately 5 degrees out, is an area of reasonably high acuity (called the parafovea; Rayner, 2009). As one moves farther from the center of vision, the visual details become more and more blurred. Because only the central 2 degrees to 5 degrees of the visual field are seen with sufficient clarity to do tasks such as read, it is necessary to focus the light from different parts of the world on this small part of the visual field in order to have the light fall on the high acuity part of the retina. By pointing the eye at multiple parts of a graph, for example, one can clearly read the axes, the legend, and the data, but this
processing must occur in serial. Thus, one of the primary functions of eye
movements is clarity of the visual input.

However, from a psychological point of view, the functionality of eye
movements extends beyond the acuity issue because eye movements are linked to
what one is attending in the environment. William James (1890) claimed that
everyone knew what attention was. To paraphrase his claim, attention is the
selection of part of the world to the exclusion of other parts for enhanced processing.
Although people have an intuitive notion of what attention is, attention as a
psychological construct is difficult to study because selecting parts of the world is
not necessarily observable and the selection is under the control of the viewer
(mostly). One can easily choose to attend to a word on this page, attend to the page
number, or some other aspect without any apparent difficulty or noticeable
behavior. For psychologists who want to study the psychological construct of
attention, it became necessary to find a measurable action that was tied to the
selection process.

Generally, visual attention involves the selection of one area or object in the
visual world for enhanced processing. Under normal circumstances, attentional
selection and where the eyes are pointed are linked. In other words, where one is
looking (or is about it look) tends to be the focus of attention. In fact, visual
attention will obligatorily shift to the location that the eyes are moving to (e.g.,
Hoffman & Subramaniam, 1995). In addition, where the eyes are currently pointed
tends to be the object/location that is currently being attended. Although it is
possible to stare at something and attend elsewhere in the visual field, normally, if
one wants to “look” at something, both attention and the eyes will be targeted in the same place. The fact that the eyes follow attentional patterns allows researchers to examine what is being selected for processing in a picture/sentence/face etc. by measuring where the eyes are currently pointing. In other words, eye location and movements provide an overtly measurable behavior that is tied to the mental selection process. From a psychological and educational point of view, the ability to have a physical marker of what someone is selecting for processing on a moment-to-moment basis is where the true power of eye tracking comes into play.

There are many types of eye movements that can occur, but the most critical for understanding what is being processed and attended in an image or text is the saccade. The term was coined in France in the late 1800s by Javal (as described in Huey, 1900) and describes observable eye movements that have a particular characteristic movement pattern. Regardless of whether one is reading a sentence, looking at a painting, searching for car keys, or examining a graph of temperature data, when one makes a saccade the eye will rotate exceptionally rapidly (in reading, almost all saccades are completed under 50 milliseconds) to point the eye toward the new point of interest and then stop for a period of time. Interestingly, while the eye is in motion, the visual input from the eye to the brain is suppressed (called saccadic suppression), and, means that humans are effectively blind during the saccade. Because the visual input is suppressed, we do not see the blur of colors that actually sweep across the retina while the eye is in motion. After the saccade has ended and the eye is pointing to a new location, the visual input to the brain is disinhibited, allowing the brain to resume processing the visual input. Following a
saccade, the eyes will pause on the new location for a brief amount of time (called a \textit{fixation}), and then the eyes will move again. Because vision is suppressed during a saccade, the only time that visual information is effectively being sent to the brain is during these fixations.

On average, a human will make around three saccades and fixations each second. Although, a person will make three fixations on average, the length of the fixation and its location are variable and indicate the relative difficulty or ease of processing at each moment. By examining which items (e.g., axes on a graph or an object in a picture) are fixated, how often they are fixated, and the amount of time that a person fixates a particular item, researchers are given a measure of the processing that occurred for that item. Vision researchers have attempted to break down the sequence of fixations and saccades into common measures that provide insight into the cognitive processes occurring. Although this is not a complete list of measures used in this type of research, the following list represents those that are common in not only reading research, but also that in scene viewing and visual search.

To appreciate the subtle differences in the various measures described below, it is easier to examine a sequence of fixations and saccades (called a \textit{scan pattern}) on an image. Figure 1 provides a cartoon of a scan pattern of an imaginary person viewing a rather boring scene. The fixations are represented by the solid black circles and are numbered in sequence from the first fixation on the screen (Fixation #1 on a central fixation cross) to final fixation on the screen (Fixation #7 on the
square). The fixations are connected by straight black lines that represent the saccadic eye movements that transition the eyes from one fixation to the next.

![Figure 1: A demonstration of a scan pattern on a simple scene. The numbered black circles represent fixations and the black lines represent saccades.](image)

**Time based measures:**

Time based measures can be thought of as processing time on a specific item or object in the image. Normally, an item must have received at least one fixation to be included in these analyses. Thus for these measures, in reference to Figure 1, the triangle would not be included in the time measurements.

**First Fixation Duration:** The duration of the first fixation on an object or word.

In Figure 1, the first fixation durations on the gray shapes would be the length of Fixation #2 (circle), #3 (square), and #5 (octagon). The triangle would not have a
first fixation duration because it was never fixated. The first fixation duration is typically interpreted as a measure of initial processing difficulty of a particular stimulus.

**First pass or Gaze Duration:** The sum of all of the fixations on an object or word the first time that an object or word is fixated. The gaze duration on the circle simply would be Fixation #2, and the gaze duration on the octagon would be the length of Fixation #5. The gaze duration for the square, however, would be the sum of the fixation durations of Fixation #3 and #4. First pass/gaze duration can be thought of as the length of time it is necessary to process a particular object or word.

**Total Fixation Time:** The sum of all of the fixations on an object or word regardless of when they occurred. The total time on the circle would be the sum of Fixations #2 and #6; the total time on the square would be the sum of Fixations #3, #4, and #7. Once again, the total time on the octagon would be the duration of Fixation #5. The total fixation time provides a marker for the time one has to process an object if the person double or triple checks the object.

**Average Fixation Duration:** The average duration of all of the fixations on the object of word regardless of when they occur. The average fixation duration on the circle would be the mean of Fixations #2 and #6, and the average fixation duration on the square would be the mean of Fixations #3, #4, and #7. Finally, the average fixation duration on the octagon would be the duration of Fixation #5. The average fixation duration is a marker for whether or not a particular object is harder to process on average than another object. In addition, average fixation
duration can be independent of the particular objects or words in a trial; returning to Figure 1, one could calculate the overall average fixation duration by averaging the durations of Fixations 1-7. It is not as common or a measure as the others in studies of the psychological processing of text or scenes because it tends to be too coarse of a measure and lacks the ability to measure the processing of individual items, but can give a general idea of the processing difficulty.

**Count based measures:** Because they are dealing with how often something is fixated or viewed, count based measures provide a more general indication of which objects were processed rather than for how long.

**Proportion Viewed/Skipped:** This is a measure of whether or not a particular object or word was fixated during a trial. In Figure 1, the circle, square, and octagon are all fixated and thus would have a proportion of 1 on this trial. Because the triangle was not fixated, it would have a proportion of 0 on this trial. Skipping is the exact opposite of the proportion viewed. This measure is perhaps the most coarsely defined measure of processing: was an object looked at or not?

**Number of Entries/separate views:** The number of separate times that an object/word is viewed during a trial. In Figure 1, both the circle and the square received 2 separate entries. The circle was viewed early in the trial (Fixation #2; the first pass) and later in the trial (Fixation #6). Because the eyes had visited other areas before returning, Fixation #6 is a new entry into the object. For the square, the initial viewing (Fixations #3 and #4) were followed by a much later fixation (Fixation #7) after the eyes had viewed other shapes. The number of entries count for the octagon would be 1 because it was only fixated with Fixation
#5 (the triangle would be 0). Although it is not as common as other measures, this measure provides measure of the number of times an object must be looked at to fully process it.

**Fixation Count:** Fixation count is the total number of fixations that occur on a particular object/word during a trial regardless of when they occur. In Figure 1, the fixation count for the circle would be 2, the square would be 3, the octagon would be 1, and the triangle would be 0. Fixation count provides an indication of the processing amount necessary to understand an object (similar to total time, but this measure is less influenced by very long or short fixations). This measure was claimed by Loftus (1972) to be the best predictor of memory for an object.

**Saccade Measures:** Measures of the saccadic eye movements themselves. These measures can provide an indication of the decision about when to move the eyes and the extent of the eye movements.

**Saccade initiation time or latency:** The time required to make an eye movement in response to a stimulus change. Typically latencies for eye movements are on the order of 150-200 milliseconds.

**Saccade Length:** How much distance is covered on average by an eye movement in a display. Typically this measurement will be given in degrees of visual angle. It can indicate how densely or sparcely packed relevant stimuli are.
### Table 1: Common Eye movement measures.

<table>
<thead>
<tr>
<th>Type of Measure</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Time-based measures</strong></td>
<td></td>
</tr>
<tr>
<td>First Fixation Duration</td>
<td>Initial processing difficulty</td>
</tr>
<tr>
<td>First Pass/Gaze Duration</td>
<td>Initial Integration and/or full access</td>
</tr>
<tr>
<td>Total Fixation Time</td>
<td>Total processing including reprocessing</td>
</tr>
<tr>
<td>Average Fixation Duration</td>
<td>Overall difficulty of processing possibly not tied to any particular object in a scene</td>
</tr>
<tr>
<td><strong>Count-based measures</strong></td>
<td></td>
</tr>
<tr>
<td>Proportion viewed/Skipping</td>
<td>Indication of the likelihood of fixation</td>
</tr>
<tr>
<td>Number of separate views/entries</td>
<td>Difficulty in processing or avoiding an item</td>
</tr>
<tr>
<td>Fixation count</td>
<td>Total processing including reprocessing</td>
</tr>
<tr>
<td><strong>Saccade Measures</strong></td>
<td></td>
</tr>
<tr>
<td>Saccade Initiation or Latency</td>
<td>Decision difficulty of when to move the eyes</td>
</tr>
<tr>
<td>Saccade length</td>
<td>Density of stimuli and exploration strategies</td>
</tr>
</tbody>
</table>

**Applications of eye movement measures within cognitive psychology**

Thousands of studies using eye tracking have been performed (over 22,000 articles in PSYCINFO have the term eye movements, PSYCINFO, 2013; See Rayner, 1979; 1998; 2009 for reviews of eye movements in cognitive psychology), and although
there are many differences in the equipment, there are two general techniques for collecting eye tracking data in cognitive psychology: observational techniques and eye movement contingent techniques. Observational techniques are straightforward – the participant views or interacts with a stimulus and the researcher measures where the eyes are looking and what eye movements are made. In the second technique, eye movement contingent technique, the researcher defines some area or marker so that when the eyes move or cross the boundary of the area, the information displayed to the participant changes.

Observational techniques date back to the earliest days of eye tracking (e.g., Huey, 1900; Buswell, 1936; Yarbus, 1967) and continue to be used frequently. The participant is shown a particular stimulus (text, picture, graph, etc.) and the researcher records the locations and times of the fixations on the stimulus. For example, the original studies of reading by Huey (1900) recorded the pattern of eye movements during reading text using a plaster cup attached both to the eye and to a pointer that marked on paper drum (somewhat like a seismograph). When the participant read the text, the paper drum provided a record of the eye movements with a time marker. Although normally the stimulus will appear on a computer monitor today and the recording is through reflecting lights off of the eye, the technique of monitoring the eye movements to collect additional dependent variables (like those described above) is the same. This technique can extend outside of the laboratory to real-world situations such as driving or walking. However, even in those cases, the goal is still the same, to determine what is being fixated.

In contrast to observational techniques, eye movement contingent techniques alter the display based on the location of the fixation. In other words, the information that is
available will change based on the participant’s behavior. The most dramatic of these
types of technique is the moving window technique. In the moving window, a portion of
a stimulus is visible dependent on the current fixation location and the remaining portions
of the stimulus are masked in some way (McConkie & Rayner, 1975). When the eye
moves, previously masked material will become visible and previously visible
information will be masked (see Figure 2 for an example of the moving window with an
image). In order to accomplish this technique, the changes in the image must happen
during the saccadic eye movement because saccadic suppression will prevent the eyes
from literally seeing the image change. Because saccadic eye movements are very quick,
this technique requires an high sampling rate eye tracker (to know as soon as possible
when the eye movement has been executed), a high precision eye tracker (in order to
know where the eyes are pointed so that the window is properly located), and rapid
monitors (once the eye movement has been detected, it is necessary redraw the stimuli on
the monitor before the eye movement is completed).

By manipulating what is available to be viewed at any one moment, researchers
can determine the relative contribution of areas outside the window to processing. One
place where the moving window made a big contribution early on was in the study of the
perceptual span of reading. The perceptual span is the size of window where normal
reading and reading with a window are the same. The window for reading English is
approximately 4 characters to the left of fixation and approximately 14 characters to the
right (See Rayner, 1998 for a review). If a reader were reading with a window of that size,
the reader would be able to read normally and be oblivious to the changing text that is
occurring outside the window.
A related eye movement contingent display change technique to the moving window is the boundary change technique. In this paradigm, the participant is looking at a scene or reading text, but when they cross an invisible boundary (e.g., the end of a particular word or the area around an object), the stimulus is changed. The requirements for eye tracking and monitors are the same as for the moving window paradigm. By changing the stimulus as the eyes move into an area (during the saccadic suppression period), the researcher can make relatively subtle changes to the stimulus to determine the effects on processing. The most prominent use of this technique has been in reading studies. In a study that my colleagues and I did looking at reading, we manipulated what was visible of a word prior to its fixation (called a parafoveal preview): sweet, sleet, or the nonword speet. When the eyes actually moved to the word, we detected the eye movement and replaced the text so that it said “sweet”. We could then determine if providing most of the letters with a real word, sleet, before the individual read the word “sweet” helped or hindered in comparison to reading a non-word speet (Williams, Perea, Pollatsek, & Rayner, 2006). The results indicated that it did help, but only under some conditions. Although frequently used in reading, this technique can also be used with images (e.g., Hollingworth, Williams, & Henderson, 2001).
Eye Movement Research

Eye movements as causal mechanism

The above techniques primarily focus on the eye movements as a dependent variable that can indicate attention and processing. However, there have been several attempts to ascribe causal mechanisms to eye movements. In other words, by forcing a particular set of eye movements, performance can be changed. Attempts of this nature have been difficult to see manifested. In reading, for example, it was discovered that poor readers have more erratic eye movements (i.e., more regressive or backward eye movements while reading) than better readers. Believing the cause of poor reading was the erratic eye movements, some researchers attempted to train poor readers to move their eyes like good readers (See Tinker, 1946; 1958 for reviews). Ultimately, these attempts failed because the root cause of the more erratic eye movements was poor comprehension and thus needing to regress in the text more to understand the text. Unfortunately, simply
having someone copy a successful individual’s eye movement pattern does not mean that
he or she will acquire the same information as the better performer.

Although it has failed as a mechanism for training better reading, there appear to be a couple of cases where there might be causal effects of eye movements. The first is an effect of bilateral horizontal eye movements on memory. There have been several studies claiming that by executing repeated left-right eye movements for approximately 30 seconds, participants’ memory will show improvement (e.g., Christman & Propper, 2001). One explanatory mechanism for this effect is that the repeated stimulation of both hemispheres in the brain serves to improve overall brain activation and thus leads to better memory. Although my lab has failed to replicate this particular effect when applied to individual items, the multiple reports of a causal effect of eye movements is interesting. A second instance where eye movements may have a causal aspect (though it is not known at this point) is the Looking at Nothing phenomenon. Participants tended to have better memory for a stimulus if they looked to the place that the stimulus was when it was studied (Ferreira, Apel, & Henderson, 2008). Although it is difficult to disentangle the causal mechanism in this case, it does appear that memory and the pattern of eye movements were correlated.

**Application of eye movements to the study of scientific information**

In relation to these measures, as part the Climate Literacy Partnership in the Southeast (CLiPSE) project, my colleagues and I have examined how participants (N=19) read a common pamphlet about climate change (*Climate literacy: The essential principles of climate science*, U.S. Global Change Research Program, 2009) while their eye movements were monitored. In addition, we examined the processing of climate-
related figures (CO2 emissions, global temperature data, etc.) both before and after reading the pamphlet to explore the impact of this type of education material on figure reading and interpretation. We were interested in the possibility that by providing climate-related information, we could alter the way that participants viewed scientific information. The presentation of the figures was counterbalanced so that one group of participants saw five figures before reading the pamphlet and another five after reading the pamphlet, whereas the other group of participants saw the figures in the other order. Finally, we asked participants a question about each figure immediately after viewing the image. The questions about the figures were answered in a 4-alternative multiple choice format, and the questions ranged in specificity from the scale that the figured used to more interpretative questions such as the relationship between two factors in the image.

With respect to the results, we found that participants were no better at answering our questions after reading the pamphlet than before. However, our greater interest was in the eye tracking patterns on both the figure and the pamphlet. Figure 2 shows scan patterns of a single participant reading page 4 from the Climate Literacy pamphlet. As is obvious from this image, the participant spent a vast majority of time looking at the text with an occasional look to the picture to the side. Figure 3 shows the overall percentage of fixations and of fixation time on the pages of the pamphlet that were on the text portion of the page versus the image portion.

Finally, we examined the how participants looked at the figures prior to reading the pamphlet and after reading the pamphlet. We were particularly interested in whether following the pamphlet, participants would look more to the y-axis, x-axis, and the data area of the figures (for those figures that had them). A critical factor in working with
images is defining the areas of interest for analysis. Figure 4 shows 3 versions of the figure viewed both before and after reading the pamphlet along with the areas of interest defined. In the end, we did not find differences in the proportions of fixations devoted to the x-axis, the y-axis, or the data between participants seeing the image before or after reading the pamphlet.

We can conclude from the eye tracking data that reading the pamphlet did not alter the way that scientific novices processed graphical information. This finding may not be surprising because the purpose of the *Climate Literacy* pamphlet is not to educate people on how to interpret scientific information, but rather to raise awareness of the issue. However, we did show that when asked to read the pamphlet, participants rarely looked to the part of the page that contained an image that was supposed to illustrate the concept. In other words, on many of the pages, half of the space was not really processed. This type of information could be used to redesign the pamphlet to more effectively integrate the text and images if the images are believed to be an important component. The one exception that we found was on a page where the image and the text were directly integrated because the image was a graphic description of the concept.

CONCLUSIONS

The use of eye tracking to determine what is being processed by people has exploded over the last three decades. As a technique for data collection, it provides a wealth of information about what the viewer is doing on a moment-to-moment basis that also marks cognitive processes. Its unique ability to pinpoint what in the environment is the current focus of attention allows researchers to know what draws attention and what does not. However, one has to be cautious about applying causal mechanisms to eye
movements. Eye movements can show large differences between novices and experts in reading or other tasks, but simply having novices imitate the eye movements of experts does not tend to yield expert-like skilled performance. Even with this limitation, the study of eye movements can give insight to the thought process and thus advance theories of education and mental processes.

Figure 3. The green dots represent locations of fixations when reading the page. The size of the dot represents the duration of the fixation. Saccadic eye movements are represented by the lines connecting the fixation dots.
Figure 4. Proportion of fixation count (left) and Proportion of Total fixation time across participants for each of the pages in the pamphlet (excluding pp. 1, 7 and 8 which were section breaks). Page 9 contained was the only page containing a graphic image directly related to the text.
Figure 5: Panel A shows the spatial extent of the y-axis region (red), the x-axis region (blue), and the data area of the temperature graph. Panel B shows one participant’s scan pattern during the pre-reading phase and panel C shows a different participant viewing the same image post-reading. As in previous figures, the green dots represent locations of fixations when reading the page. The size of the dot represents the duration of the fixation. Saccadic eye movements are represented by the lines connecting the fixation dots.
References


doi:10.1037/0894-4105.15.4.607


doi:10.3758/BF03196215


doi:10.3758/BF03203972


