Visual Imagery, Metadata, and Multimodal Literacies Across the Curriculum

Anita August

Sacred Heart University, USA
ABSTRACT

This research used eye movement tracking to study the ways adolescents decode and comprehend multimodal texts. The focus is on eye movement regression between text and picture to investigate how participants use the two-stage information-processing model when attempting to comprehend visual texts. Ten adolescents ranging in age from 14 to 19 were asked to read a series of different types of visually based texts. Specifically, they read six short excerpts from graphic novels that varied widely in the complexity of both textual and visual features. The narratives included graphic novels, graphic retellings of canonical texts, and wordless visual texts.

Visually based texts have become a dominant form of literacy in the multimedia environment of the digital age. Today’s educators must devise ways of teaching visual literacy, and to do so, they must themselves understand the processes underlying comprehension of visual texts. This project seeks to deepen this understanding, using eye movement tracking to study the ways adolescents decode and comprehend visual texts. Although eye movement tracking is well established as a technique for investigating the comprehension of written texts, we are aware of no studies that have mapped the eye movements of readers of visual texts (Paulson & Freeman, 2003). The findings provide important insights into the design, process, and practice of literacy instruction for adolescents in the new generation, laying a foundation for new theoretical and practical understandings of the teaching of visual texts.

DOI: 10.4018/978-1-5225-2808-1.ch013
THEORETICAL FRAMEWORK

This research utilized the two-stage theory of information processing developed by Carroll, Young, and Guertin (1992). Originally proposed to account for eye movements during reading performance, this theory posits that readers use one of two types—or stages—of reading processes, depending on the difficulty of the text. The theory further proposes that the two stages can be distinguished by different patterns of eye movements. Stage 1 processing, involves a simple, literal interpretation of text that is relatively easy to comprehend and is characterized by brief fixations and few inter-sentential regressions. Stage 2 processing, which involves solving problems encountered with increased text complexity, is characterized by long deep-processing fixations and more inter-sentential regressions. In this research, the focus will be on eye movement regression between text and picture to investigate how participants use the two-stage information-processing model when attempting to comprehend visual texts.

RESEARCH QUESTIONS

The central research question is: When participants are asked to comprehend more and less complex texts, how do their eye movements reflect how text and visuals interact during this process? The current project drew its inspiration from research on the relationship between eye movements and information processing during reading. The study of eye movement has a long history (O’Regan, 1990), and there is a substantial body of research on eye movements while viewers are engaged in reading (Huey, 1908/1969; Just & Carpenter, 1980; Justice & Landford, 2009; Rayner, 1978, 1998). One consistent claim in this field is that eye movements are influenced by textual variables (Rayner, 1998). Research has shown that as text becomes conceptually more difficult, fixation duration and regression frequency increase, while saccade length decreases (Rayner & Pollatsek, 1989). Carroll and Slowiaczek (1987) proposed that there are two modes of eye movement while people read extended texts: a normal processing mode and a search and problem-solving mode. In the first, eye movements are under the control of an automatic lexical encoding process, with few inter-word regressions. In the second, movements are under the control of top-down processes, with longer fixations and more regressive fixations.

The measurement of eye movements has been found to be a useful tool for studying the implicit processes involved in information processing. In particular, visually guided scan paths mirror the different levels of processing (Zihl, 2008). However, in the literature on eye movement during reading, we have found few studies that map the eye movements of readers of visual texts. The data resulting from this project shows how readers comprehend visual text and will provide information to teachers regarding what parts of visual reading must be scaffolded.

MODES OF INQUIRY AND DATA SOURCES

The work of Carroll et al. (1992), discussed earlier, showed that the more complex the task, the more likely readers are to move from basic literal interpretation into deep problem-solving processing modes. This finding raises the question of what relationship exists between complexity and comprehension in the reading of visual texts. For example, do readers use visuals to support their comprehension of text
Using Eye Movement to Study Adolescents’ Comprehension of Visual Texts

embedded in graphic novels? Should we expect to find that more difficult visual texts are associated with more or less searching for nonliteral interpretations of elements such as symbolism, metaphor, and foreshadowing?

To investigate these questions, the researcher undertook a series of data collections focused on eye movements during reading performance tasks. The participants, 10 adolescents ranging in age from 14 to 19 were asked to read a series of different types of visually based texts. Specifically, they read six short excerpts from graphic novels that varied widely in the complexity of both textual and visual features. The narratives included graphic novels, graphic retellings of canonical texts and wordless visual texts.

The participants’ eye movements were tracked using the ASL DH640 eye tracking machine. Each page of the graphic novel excerpt was scanned into a computer attached to the infrared camera. The resulting information is displayed as fixations (places where the eyes stopped) and saccades (patterns of eye sweeps across the page). After recording the reading, the participant’s eye movements are displayed as lines indicating saccades and dots indicating fixations. Each reader read approximately 20 pages of text from graphic novels. Participants were asked to read each page silently and then indicate when they were finished. Each page was recorded as a separate event by the ASL software. The readers were told to read normally and were informed that they would be asked to retell after finishing all the pages for a particular novel. For the most difficult text, Othello, participants were asked to read one page using only the pictures and not the text bubbles.

As stated earlier, the graphic novels chosen varied widely in their complexity of both the textual and visual features. After reading each of the excerpts the participant. Retellings were coded by two independent coders using a tool developed by miscue analysis researchers (Wilde, 2000). For participants, three of the graphic excerpts were judged easy to comprehend. These were, The Arrival, Garage Band, and Pitch Black. The average comprehension scores for these three novels was 75%. One excerpt was judged to be of average difficulty, American Born Chinese. One of the texts, Othello, was judged difficult to comprehend. Ease of comprehension was estimated based on amount and difficulty of text as well as completeness of retell. In other words, when the reader could more completely retell the text, it was judged to be easier to comprehend.

The process of reading visuals was then analyzed through the use of a cumulative region reading time analysis (Brysbaert & Mitchell, 1996)). In this analysis, readings of text and visuals were compared based on the amount of time spent in one area. This analysis was conducted through using a feature of the ASL software called AOI or area of interest analysis. Particular portions of the graphic novel panels were identified and data for only that area are displayed. In this way the researcher was able to determine the amount of time spent viewing one particular area (amount of time spent on visuals versus text) and could also determine the fixation duration, saccade length, and order of fixations within a particular panel or series of panels in the graphic novel. AOI’s were created for text bubbles to determine fixation duration and number of fixations on text and were also created for visuals separate from text bubbles.

Results

This data shows that fixation duration increases with difficult level of the text for both text and visual areas within the graphic novel panel. In other words, when the text was more difficult, as with the novel Othello, readers spent longer times fixating within text bubbles. This was also true when visual became more complex. Visuals that were less dense or held less information produced shorter fixation durations.
Using Eye Movement to Study Adolescents' Comprehension of Visual Texts

Table 1. Average results

<table>
<thead>
<tr>
<th>Average Results:</th>
<th>Easy to Comprehend</th>
<th>Difficult to Comprehend</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixation duration</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Text</td>
<td>.26*</td>
<td>.33</td>
</tr>
<tr>
<td>Visual</td>
<td>.22</td>
<td>.30</td>
</tr>
<tr>
<td>b</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Text</td>
<td>53</td>
<td>189</td>
</tr>
<tr>
<td>Visual</td>
<td>19</td>
<td>69</td>
</tr>
<tr>
<td>Picture only reading</td>
<td></td>
<td>168</td>
</tr>
</tbody>
</table>

*Time is measured in second

Patterns of Reading Easily Comprehended Texts

By far the most common pattern for participants was a large percentage of their eye fixations were on the text in the graphic novels. Across all of the readings, 81% of all fixations were on the text bubbles embedded in the graphic novel. It was, in fact, rare for participants to fixate on any aspect of the pictures when they found the text easy to comprehend. The percentages changed when readers encountered more difficult texts. As can be seen the above table, the percentage of fixations on visuals increased to almost 40% when participants read the more difficult texts, such as Othello. This may indicate that the readers found it necessary to use the visuals more when reading more difficult text.

The pattern of reading easily comprehended texts was characterized by brief fixations on text bubbles and few inter-modal regressions (or eye sweeps between text and pictures). In other words the readers fixated for brief periods on instances of text and rarely looked at the pictures during the reading of a panel. Readers tended to move their eyes in a smooth linear left-to-right path through the text bubbles embedded in the graphic novels. This pattern was not entirely true however. For those students who had

Table 2. Fixation duration

<table>
<thead>
<tr>
<th>Fixation Duration</th>
<th>Experience Reading</th>
<th>No Experience Reading</th>
</tr>
</thead>
<tbody>
<tr>
<td>Easy Text</td>
<td>.22*</td>
<td>.28</td>
</tr>
<tr>
<td>Difficult Text</td>
<td>.30</td>
<td>.35</td>
</tr>
<tr>
<td>Easy Visual</td>
<td>.29</td>
<td>.24</td>
</tr>
<tr>
<td>Difficult Visual</td>
<td>.34</td>
<td>.28</td>
</tr>
<tr>
<td>Number of fixations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Easy Text</td>
<td>57</td>
<td>43</td>
</tr>
<tr>
<td>Difficult Text</td>
<td>177</td>
<td>193</td>
</tr>
<tr>
<td>Easy Visual</td>
<td>14</td>
<td>11</td>
</tr>
<tr>
<td>Difficult Visual</td>
<td>53</td>
<td>21</td>
</tr>
<tr>
<td>Picture only</td>
<td>99</td>
<td>70</td>
</tr>
</tbody>
</table>

• Time measured in seconds
more experience reading graphic texts, there were more inter-modal saccades even with easily comprehended text. An example of this can be seen in the two readings illustrated in Figure 1.

The first reading is from a female participant who had less experience reading graphic novels. She indicated that she had read some Japanese Manga but had not read other types of graphic novels. The second reading comes from a boy who said in his interview that he was an avid reader of graphic novels. The first reading shows that the reader ignored the visuals entirely as she read through the text. The second shows long saccades and short fixations on both the visuals and text in the novel. It is important to note that the second reader gave a more complex retell of the story when he had finished reading. In his retell he said, “I noticed that the bird flying up was similar to the pattern of the father flying the plane in the next panel. I think it is about like freedom for the son, that he sees the bird and the remote control plane and thinks about his own freedom.” The first reader gave a much more simplistic retell, “The story was about this kid who heard some dogs and then went to see his dad.” This seems to indicate that a focus on visuals, rather than making the reading less complex, actually allowed the participants to think more deeply about the abstract literary elements of the text, including metaphor and symbolism. When the second reader indicated that the bird’s flight path was repeated in a subsequent visual and then understood that repetition to be important, as representative of freedom, he shows that a concentration of visuals significantly expanded his understanding on the story compared to the first reader.

When reading easy and average text, the readers fixated 81% of the time on the text bubbles in the graphic novels. Only 10% of readers’ eye fixations were on visual features (the other 5% of fixations were off the page as the reader moved from page to page while the other 3% can be accounted for by calibration error within the machine). This number was consistent for each of the easily comprehended texts. The readers fixated on pictures primarily before and after they had finished reading the text bubbles rather than regressing to pictures while reading text. This indicated that the readers were not using pictures to help them make sense of the text while reading.

Figure 1. Two readings of a page from Garage Band
This held less true for readers who had experience reading graphic novels. These participants, especially when reading more difficult texts showed more fixations on graphic elements. This may indicate that these readers recognize the value of the visuals in aiding comprehension. This can be seen in Figure 2.

For the most part this reader’s eyes fixated on the text bubbles. The first panel shows that the reader did follow the vertical movement of the visuals as she followed a waterfall down the mountain. In panel one there is a significant fixation on the waterfall at the bottom of the panel The reader’s eyes then swept to the right, paused for one brief and one longer fixation, then went to the top of the second panel. This pattern shows that the reader used the visuals to help her understand that she was not reading the panels in the proper order. After moving to panel two, her eyes fixated on the text bubbles for the next three frames. It is interesting to note that the reader did not fixate on the text bubble in the last panel. This may indicate that she felt she had already “seen” the last panel as, when reading graphic novels, the reader sees the visuals and/or text peripherally. It is possible that the reader had perceived the final panel initially and there did not need a second fixation.

Figure 2 also shows that the reader fixated on visuals only in order to give her information regarding reading order. A non-fixation on graphics involves an efficient process of finding the most informative aspects of the texts. In other words, her fixations indicate that the pictures provided less of the information necessary to comprehend the text and were therefore viewed only as part of saccades or sweeps across the page.

However, the reader’s eyes did move from visual (panel 5) back to text (panel 3). This eye sweep can be considered an intermodal regression, or a movement between text mode and pictures while reading. Regressions are important as they are immediate indicators of comprehension processes used during reading (Paulson & Freeman). In the case of easy to comprehend texts readers displayed few intermodal

*Figure 2. Reader’s use of graphics to assist comprehension*
regressions during the readings. Across the fifteen pages of text read in the less difficult novels, there was an average of twenty intermodal regressions counted. The intermodal regression that did occur tended to focus on unusual or unexpected objects in the visuals. This can be seen in Figure 3.

As can be seen from the above reading, the participant initially focused on the faces of the characters then swept quickly to the shadow of the tail above them. He then returned to look again at the characters. This pattern of focusing first on characters, and particularly on faces, was common. Readers then tended to look for distinctive or unexpected features of the visuals. Figure 8 shows the reader looking at the shadow of a dragon’s tail. Every reader of this text focused on this detail and indicated they did so because it didn’t overtly fit with the rest of the story. Inconsistent objects (those not overtly mentioned in the text or not expected in the context of the story) were fixated earlier and longer - along with more important objects. There were longer fixations on low-frequency images versus high frequency images and this seems to indicate that longer fixations are required to make sense of what initially does not appear to belong in the scene. For these participants what seemed to be most informative in the visuals were distinctive or unexpected visuals.

This same pattern can be seen in Figure 4.

The Figure 4 graphic shows that the reader spent more time looking at the “unexpected” or anomalous faces in the party scene. You can see that she spends more time fixating on the bird figure and the three male gods depicted at the top, middle and bottom of the picture than she does on the female characters who look less unexpected. She later indicated that the strangeness of the characters cued her that this

Figure 3. Pattern of fixation in easy to comprehend visuals
was a fantasy story. This may indicate that using unexpected visuals may give students a great deal of information regarding genre, etc. It should be noted that fixation duration on unexpected objects decreased as students read within the same text. For example, in *The Arrival*, the tail of the dragon shown in figure 8 appears several times in the novel and fixation duration on the tail decreased as students saw it more times.

Overall, most participants shared patterns when reading easy to comprehend texts. They initially made many long saccades to fixate on informative parts of the scene. Less complex visuals resulted in short fixations and fixation duration increased as visuals became more difficult. However, fixation duration on visuals decreased as readers became more familiar with the story. Even with easy to comprehend texts, retelling accuracy increased with viewing duration and number of fixations. This seems to indicate that, when reading graphic novels, meaningful information can be extracted more visuals than from text.

**Patterns of Reading for Difficult Texts**

Those readers who scored lowest on the comprehension task for the difficult text reading were those that seemed to ignore the visuals almost completely.

This reader gave a very brief and somewhat inaccurate retell of the passage. She said, “I’m not sure what is happening, he is talking about money and trying to convince the guy to make money, I’m not really sure what is happening.” Interestingly, the participant in Figure 6 was an older reader who had read Shakespeare before. She had not however read graphic novels before. For the entire group of participants, those who spent more time looking at visuals and who had more intermodal regressions while reading scored higher on the retell task for the difficult text, whether or not they had significant
experience reading Shakespeare. This may indicate that the visuals provide much more information to readers than does text that is extremely difficult to comprehend.

The Figure 6 reading excerpt comes from a reader who scored highly on the comprehension of the most difficult text. This reader said, “Iago is trying to convince himself that the court of Othello is out of balance and that he can bring it back into harmony and, that if he does that and makes money, there is nothing wrong with that. But you can tell by the fact that he keeps drinking wine that he is becoming more evil, he is building up his courage. The angels in the first picture are symbolic of the battle within him, to betray his friend or make money for himself.” The overtly non-linear movement of the reader’s eyes in the figure above seems to show that frequent intermodal regressions from text to visuals allow for more complex retelling. By using the visuals during the reading process, the readers who were less tied to the text showed higher comprehension.

Researchers have argued that fixation time is related to the readers’ cognitive load (Rayner, 1998). This is not entirely true for the reading of multimodal texts. For the participants of this research, as the difficulty of the text increased and when they were asked to read using primarily the visuals, saccade length increased as their eyes seemed to “bounce” around the visual until they identified important features. Only then did the length of fixations increase. Since the eye movements in this case were more erratic than in linear text reading, those saccades that showed movement between text and visuals were classified as intermodal regressions.

The importance of regressions only became evident when students were asked to do text/visual combination reading. In this case it became clear that, unlike text-only reading, the higher frequency of regressions from text to picture was associated with higher levels of comprehension. This can be seen from the Figure 7.
This reader moves from text bubble to visuals during the reading process. In fact it was this pattern that allowed for the most comprehensive retells from readers. The initial long fixation on the painting in the background indicates that the reader is trying to incorporate the meaning of the ruined painting into his understanding. During his retell this reader said, “Iago is saying that he has decided to betray Othello and he just has to bring his monster to light. The monkey has to indicate that monster or that evil. The monkey isn’t mentioned in the text so he must be symbolic or something so I thought since he is sitting on Iago’s head that indicates that the evil has won out and defeated the good. I noticed that the monkey had thrown the apple or fruit at the painting of the angels and that also shows that he is trying to destroy the good. Iago looks defeated like that monster won.” Clearly this reader’s ability to use intermodal regression significantly helped his comprehension of the text.

DISCUSSION

Text in graphic novels tends to have more abbreviated syntax, less sentence complexity, and is often written as dialogue between characters. This suggests that concentrating on text alone when reading graphic novels may not offer students the kind of experiences that will increase their reading abilities. If the words in graphic novels are less important and less sophisticated than in traditional text, the question then becomes, what is the value of using graphic novels in the literacy classroom? The complexity of graphic novels only emerges when the text and visuals are read in conjunction. In order for this truly
multimodal reading to happen, it may be necessary for the teacher to ask students to significantly alter how they think about reading.

The data reported in this study shows that the process of reading graphic novels is very different than the process of reading traditional text. Readers who scored highest in comprehension used shorter fixations on text than did the readers who received lower scores. They also had longer and more fixations on visual features than did the other group. The highest scores on comprehension were also associated with longer, less linear saccades across the page and with more intermodal regressions. In other words, those readers who scored better were more likely to look at the visuals and to look at them during their reading of text. This may indicate that these readers were using visuals to initially establish the “gist” of the scene and would integrate the text into their understanding of the visuals.

This process involves two main factors regarding eye movement. First is fixation duration. Fixation duration increased as the text and visuals become more complex. However, even with less complex texts, increased fixation duration on visuals led to higher retell scores. This seems to indicate that reading ability for graphic novels has much more to do with experience reading multimodal texts than it does with overall reading ability. All of the participants shared a pattern of increased fixation duration on unexpected objects and decreased duration on visuals in general as the text became more familiar. These common patterns in the retells show that the readers understood that paying attention to unexpected objects would give them vital information in understanding the text as a whole. Once this symbolic or metaphorical understanding is established then less time spent fixating on the object is necessary.
The second main factor regarding eye movement was fixation pattern or the order with which readers looked at texts and visuals. Once again, less experienced readers of graphic novels tended to look at visuals only before or after they had read text bubbles. More experienced readers and those who had higher retell scores were more likely to look at visuals while they were reading text. In other words, intermodal regressions were related with higher retell scores. These participants were also more likely to look more often at visuals generally. However, this difference became most pronounced when the difficulty of the text and visuals increased. For the task of reading the excerpt from *Othello*, the students with the highest retell scores were more likely to focus more often on visuals than on text.

While the link between eye fixations and cognitive processes might not be as tight as they are in reading text-only, this research does indicate that more complex visual texts are more easily understood when fixation time increases. In other words, it is not simply seeing the visuals, but using them to make sense of story that characterizes good readers of graphic texts.

**IMPLICATIONS AND CONCLUSIONS**

Eye movements vary between readers and between texts. The same reader will use visuals differently depending on the complexity of the text. Also, different readers use visuals in different ways. This may have significant implications for teaching visual texts.

First, teachers must model the transmediation process. This means that adolescents need to be explicitly shown how visual elements and textual elements are related as a coherent whole. Teachers need to model basic visual interpretation skills (what do we pay attention to) as well as modeling how those skills can suggest more complex meanings in the text. For example, when an unexpected visual appears (such as the tail of a dragon) this is often an indication that the readers should interpret the visual metaphorically.

Second, teachers should ensure that collaborative discussions and support materials focus equally on visual and textual meanings. This may mean that teachers provide materials to encourage students to discuss the visual aspects of the texts. Assignments related to readings should also include reflections on the visual modes in the text. By offering similar weight to work around visual and textual elements, teacher can illustrate that important meanings are held in the visual mode.

Third, this research suggests that graphic novels offer easy and accessible opportunities to “fix-up” misunderstandings (through looking at visuals). As such, graphic novels could allow struggling readers more opportunities to practice using comprehension as the primary tool to get through difficult readings. By allowing for more sustained comprehension, visual texts may be useful as a transition tool to encourage students to read longer texts. The greatest value in transmediation may be realized when trying to teach students to read very difficult texts (such as Shakespeare).

According to Hassett (2001), the texts students encounter today embody cues for reading that do not primarily rely on alphabetic print and extend beyond letters and words on the page, requiring readers to actively focus on textual elements beyond print. This research provides important data on how both the complexity of visual texts and the individuals’ previous experiences influence reading behavior. Through the simple mapping of eye fixations and regression patterns, the work offers the first empirical data on the processes involved in reading visual texts.
The two stage reading process of literal interpretation versus problem solving mode has proven to be too simplistic to account for the variety of participant’s eye movements. Rather, comprehension processes do play a major role in influencing when and where the eyes move. Readers pick and choose what they look at; they sample the text and visuals in order to establish meaning. This seems to indicate that eye movements when reading graphic novels are closely related to the individual backgrounds of the readers. Research with multimodal texts seems to offer the potential for providing insights for teachers. With this initial understanding of adolescents’ reading behaviors while engaging with multimodal from this research, we hope that educators will see the value of graphic novels and be more confident integrating the visual in their literacy instruction.

REFERENCES


Using Eye Movement to Study Adolescents' Comprehension of Visual Texts


**KEY TERMS AND DEFINITIONS**

**Eye Tracking:** Eye tracking is process through which infrared cameras track eye movements of participants as they read texts.

**Fixations:** Fixation refers to the eyes pausing on a word or group of words.

**Saccades:** Saccades are voluntary movements of the eye from one fixation point to the next.

**Transmediation:** Transmediation is the translation of ideas between two media (text and visual).

**Two Stage Theory:** This theory posits that readers use one of two types—or stages—of reading processes, depending on the difficulty of the text. The theory further proposes that the two stages can be distinguished by different patterns of eye movements.

**Visual Texts:** Visual texts are texts where the primary meaning is held in the visuals. This includes graphic novels, illustrated plays, and wordless books.

**ENDNOTES**

1. *Fixation* refers to the eyes pausing on a word or group of words.

2. *Regressions* are right-to-left movements along a line of text or movements back to previously read lines (Rayner, 1998).

3. *Saccades* are voluntary movements of the eye from one fixation point to the next.