Examining the Effects of Text and Images on Story Comprehension: An Eye-Tracking Study of Reading in a Video Game and Comic Book

Charles K. Kinzer
Selen Turkay
Daniel L. Hoffman
Nilgun Gunbas
Pantiphar Chantes
Apichai Chaiwinij
Tatyana Dvorkin

Teachers College Columbia University

Reference:

Kinzer, C. K., Turkay, S., Hoffman, D. L., Gunbas, N., Chantes, P., Chaiwinij, A., & Dvorkin, T. (2012). *Examining the Effects of Text and Images on Story Comprehension: An Eye-Tracking Study of Reading in a Video Game and Comic Book*. In P. J. Dunston, S. K. Fullerton, C. C. Bates, K. Headley, & P. M. Stecker (Eds.), Literacy Research Association Yearbook 61 (pp. 259-275). LRA: Chicago, IL.

Examining the Effects of Text and Images on Story Comprehension: An Eye-Tracking Study of Reading in a Video Game and Comic Book

Charles K. Kinzer
Selen Turkay
Daniel L. Hoffman
Nilgun Gunbas
Pantiphar Chantes
Apichai Chaiwinij
Tatyana Dvorkin
Teachers College Columbia University

Understanding how students comprehend information in multimodal text can offer valuable insights into the relative importance of text and images in comprehension. To this end, this study investigates the effects of text and images on sixth graders' comprehension of narratives presented in two different formats: a comic book and a video game. Both formats are popular with today's adolescents and have been suggested as vehicles potentially useful for instructional purposes. Yet, despite their popularity, there is little research available on narrative comprehension as related to playing video games or reading comics, although researchers (e.g., Mayer, 2011, p. 298) emphasizes the usefulness of knowing whether people can learn as well from a computer game as from a more conventional medium.

Previous work found differences in students' comprehension of narrative when presented in a book, video game, and comic (Kinzer, Hoffman, Turkay, Gunbas, & Chantes, 2011); results showed that students who read the story in book form achieved higher literal comprehension scores than the other two groups, though there was no difference in inferential comprehension scores. Although this finding was consistent with early studies investigating text and images in reading comprehension (e.g., Miller, 1938), it was still surprising considering today's "massively visual society" (Emmison & Smith, 2000, p. viii) in which children are presented with multimodal information in almost every aspect of their lives. Given our image-laden culture, why did the more visual formats studied fare worse as measured by comprehension?

A possible explanation is that although today's children are saturated with images, they do not necessarily use these images to aid their reading comprehension. Kinzer et al.'s (2011) findings suggested that the images in the game format might distract readers from the story, or that the images, while containing information relevant to comprehending the narrative, were ignored in favor of the text. Another possibility is that schools focus on text-based reading and comprehension, with relatively little focus on understanding and interpreting images, which may have influenced the previous findings (L. Vasudevan, personal communication, March 8, 2011).

To address these issues, four research questions were proposed: (a) Do sixth graders differ in their comprehension of a narrative presented within a comic book versus an adventure video game?, (b) Do sixth graders differ in their comprehension of information that appears in text or graphics, when presented in a comic book or video game narrative format?, (c) Do sixth graders focus differently on text and images within a narrative that is presented as a comic book and a video game?, and (d) What can eye tracking tell us about how sixth graders focus on text and images within a narrative presented in a comic book or an adventure video game?

BACKGROUND

Early work on the impact of text and graphics on student learning provided mixed results. Work that aimed to determine the effect of illustrations on elementary-aged children's comprehension of stories found either no effect (e.g., Miller, 1938; Vernon, 1953; Weintraub, 1966) or a negative effect (Rose, 1986; Samuels, 1967; Watkins, Miller, & Brubaker, 2004). However, many of these studies were conducted in times when children were not exposed to interactive text and embedded images, audio, and video, to the degree common today.

More recent work has been influenced by Paivio's (1991) dual coding theory, which suggests that graphics and text are coded into memory differently. Paivio's theory implies that comprehension is enhanced when text and images are related and research following this line of reasoning has argued for a type of multimedia principle, positing that students learn better from text and pictures than from text alone (Anglin, Vaez, & Cunningham, 2004; Mayer, 2009). Variations in the effectiveness of the multimedia principle may be related to the domain being taught. Kress & van Leeuwen (2006) argue that in technical fields images have become the major means of curricular content, whereas in the more humanistic subjects images vary in their function between illustration, decoration, and information (p. 16).

For younger children, the presence of images alongside text seems to benefit comprehension and Peeck (1993) recommends teaching visual literacy in the context of teaching reading comprehension. Gambrell and Jawitz (1993) found that fourth graders' reading comprehension increased significantly when they read text with illustrations, and Fang (1996) argues that pictures contribute to children's development of literate behavior. Sipe (1998) suggests that children construct meaning through the interplay of text and image, while Levin, Anglin, and Carney (1987) argue that easy-to-follow texts are unlikely to yield any cognitive benefits from the inclusion of pictures as they are already concrete and engaging. In contrast, Jewitt and Oyama (2001) maintain that graphics in text are needed to enhance understanding.

Of course, the mere presence of images does not mean that students access them or know how to use the information they contain, and Carney and Levin (2002) raise an important question about the ability of today's students to process pictures and text when reading, asking if the "cyberstudents" of today will differ from the book-learned "liberstudents" of the past in their ability "to process picture and text information comprehensively, and with comprehension" (Wang & Newlin, 2000, p. 23).

Use of Comics for Literacy

Comics are multimodal stories told by using an almost seamless blend of text and image (Kuechenmeister, 2009). McCloud (1993) defines comics as "juxtaposed pictorial and other images in deliberate sequence, intended to convey information" (p. 9). Although comic books are sometimes criticized as being light reading, they are often part of children's pleasure reading, which has attracted the attention of educators who view comics as a way to help students engage in text (Eikmeier, 2008; Krashen, 2004; Tiemensma, 2009). Comics have been effectively used in language classrooms (Wright & Sherman, 1994; 1999) and in teaching foreign languages (Chun, 2009; Ranker, 2007; Liu, 2004). Sherman and Wright (1996) showed that newspaper comic strips can be used to promote higher-level thinking, and Liu (2004) found that students with low-level English proficiency benefited from comics significantly more than students with high-level English proficiency. In a study of young readers of Archie comics, Norton (2003) concluded that the pleasure children derive from comics is due largely to a sense of text ownership, which in turn provides confidence to engage "energetically and critically" (p. 145). Even in middle school, where boys sometimes "don't like to read, and often don't read very well" (Blair & Sanford, 2004, p. 453), research has found those who read comics read more in general and enjoy reading more compared with boys who do not read comics (Ujiie & Krashen, 1996).

Reading and Video Games

Video games can be considered an important medium for literacy learning, as they can provide authentic purposes to read text (Commeyras, 2009). The authentic use of text can increase students' interest in reading (Watson, 1989). Interest, in turn, can result in increased persistence leading to higher overall achievement (Moje, Overby, Tysvaer, & Morris, 2008). In addition to motivating students, video game text can help with decoding and word recognition practice.

A national survey found that school-age children devote about seven hours per week to playing video games (Woodard & Gridina, 2000), and reports show an increase in the time youth spend playing video games, with boys averaging 13 hours and girls averaging five hours weekly (Gentile, Lynch, Linder, & Walsh, 2004; Rideout, Foehr, & Roberts, 2010). Given the popularity of video games among adolescents, educators might consider them as texts to promote reading.

Gee (2003) notes that adolescent game players engage in multiple literacy practices. These practices have the potential to meet state and national standards in reading, writing and technology (Steinkuehler, 2007). Studies have shown that students engage in advanced reading comprehension when playing games (Martin & Steinkuehler, 2011) and also in game-related activities outside actual play by participating in forums and affinity groups (Steinkuehler, Compton-Lilly, & King, 2009; Turkay, Kinzer, Hoffman, Gunbas, & Nagle, 2010). Steinkuehler (2011) argues that videogames "support a complex textual ecology" that can spark "expansive reading practices" (p. 13).

Taken together, the mixed modal nature of comic books and video games and their potential as instructional tools highlight the need to better understand how students process text and images in rich narrative settings. Does format alone influence how the text and images are processed? Have students developed varied strategies to decipher multimodal texts depending on form or do they treat all text and images the same? To address such questions, it is important to think about a research design that can couple measures of comprehension with measures of eye movement in response to the visual stimuli of text and images. Together, such data can help educators tease apart the relationship between what students attend to in multimodal texts, and what they subsequently comprehend about the narrative.

Eye-Tracking

Due to advances in technology, the use of eye-tracking equipment has become a more feasible way to investigate people's visual actions and attention in various contexts. Eye-tracking has been used by researchers in many areas to examine aspects of learning, skill level, cognitive processing, and emotion (Rayner, 1998; Goldberg & Kotval, 1999; Coan & Allen, 2007). This technology produces a number of objective metrics that can be used to better understand what readers look at in a stimulus, how long they look, and the visual path they take while viewing it.

The main measurement in eye-tracking research is a "fixation," a brief moment when the eye is paused—e.g., on or around a letter, word, or word group. By tracking fixations, a number of useful metrics can be derived, including: (a) *fixation duration*, the cumulative duration and average spatial location of a series of consecutive fixations within an *area of interest* (AOI), the area of a display or visual environment that is of interest to the research team; (b) *mean fixation duration*, an indicator of information complexity and task difficulty (Rayner, 1998); and (c) *number of fixations*. Such metrics have been "highly reliable and useful in inferring the moment-to-moment processing of individual words and larger segments of text" (Starr & Rayner, 2001, p. 156).

Researchers have argued that fixation duration and cognitive effort are related, with greater fixation durations linked to greater cognitive effort (Goldberg & Kotval, 1999). Work examining the processing of text and images together reveals that people favor text over pictures when reading material that includes both (e.g., Beymer, Orton, & Russell, 2007; Hegarty & Just, 1993; Rayner, Rotello, Steward, Keir, & Duffy, 2001; Schmidt-Weigand, Kohert, & Glowalla, 2010). In examining how people look at cartoons consisting of a single picture with a relevant caption, Carroll, Young, and Guertin (1992) found that processing of the two seemed to be relatively isolated events; the picture frequently was not given full inspection until the text had been read. Rayner et al. (2001) found similar results with text and pictures presented in print ads.

Eye-tracking has been also used to examine how people read non-traditional texts. Omori, Igaki, Ishii, Kurata, & Masuda (2004) showed that both page layout and balloon placement influence how easily comics are read. Studies using eye-tracking reveal that inexperienced comic readers tend to focus more on text than on visuals and tend to wander around the page (Nakazawa, 2004; Allen & Ingulsrud, 2007).

The dynamic nature of video game play complicates the analysis of eye-tracking data as compared with non-dynamic interfaces such as still images (Jennett, Cox, Cairns, Dhoparee, Epps, Tijs, & Walton, 2008). This challenge has not stopped researchers from investigating usability issues (Mat Zain, Abdul Razak, Jaafar, & Zulkipli, 2011), learning processes (Knoepfle, Wang, & Camerer, 2009), and learning in video games (Kiili & Ketamo, 2010; Law, Mattheiss, Kickmeier-Rust, & Albert, 2010). Kiili & Ketamo (2010) found that players' perception patterns varied greatly when playing an educational game, and that extraneous elements on the interface caused confusion. Similarly, Law et al. (2010) found the existence of textual feedback did not have a significant effect on attention, reporting a negative correlation between number of fixations and the understandability of the game.

Thus far, the case has been made that text and graphics play a major role in society today through "new" media such as comic books and video games, and that these media are potential vehicles for literacy exploration and instruction. The present study aims to contribute to the dialogue about the importance of understanding how text and images affect comprehension by presenting students with a single narrative in two formats and using a combination of paper-based comprehension measures and eye-tracking metrics.

PARTICIPANTS, INSTRUMENTS, PROCEDURES

Twenty (10 female and 10 male) sixth-grade public school students in New York City, all native English speakers with normal vision, participated in this study. Participants were randomly assigned to either a Game Group (GG, n = 10) or Comic-book Group (CG, n = 10).

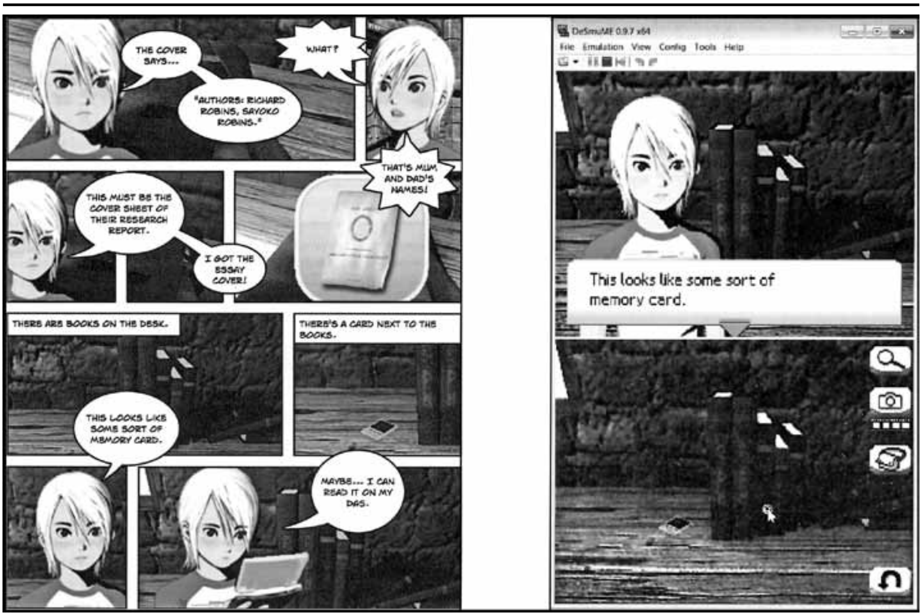
The Narrative

The narrative used in this study was from *Trace Memory* (released outside of North America as *Another Code: Two Memories*), an adventure video game developed by CiNG (2005) and published by Nintendo for the Nintendo DS. In adventure games, players assume the role of a protagonist in an interactive story driven by exploration and puzzle-solving instead of physical challenge (Adams, 2006).

Trace Memory was chosen for its story-driven game play, and the unlikeliness of sixth-grade students in 2011 having played a game released in 2005. It was described as a "touchable mystery novel" with "lengthy conversations" (Staff Reviewer, 2005; Hruschak, 2006) and reviewers claimed it was "well written" with a "good amount of suspense" (Harris, 2005; Parish, 2005). Also, the protagonist, Ashley Mizuki Robbins, who is searching for her father on Blood Edward Island, is a 13-year-old girl—a similar age to the target audience of the study.

A 10-minute, logically complete section of the game was selected for use in the study based on its representativeness of the game's mechanics and story. Using the selected portion, a 20-page comic book was created (see Figure 1), using procedures described previously by Kinzer et al. (2011, p. 268). All images and text came directly from the game; participants in all groups thus saw and

Figure 1. Sample Comic Book Page (Left) and Nintendo DS Emulator "Pages" (Right)



read the same content. Figure 1 shows a sample comic book page and "pages" from *Trace Memory* running on the Nintendo DS game emulator.

Because the portion of the story used occurred after the beginning of the story, to ensure consistency in the "lead up" to the section to be played/read, a 2.5-minute video was created to familiarize participants with *Trace Memory*'s characters and plot to that point in the story.

Participants played *Trace Memory* on a computer using DeSmuME (DeSmuME, 2009), an open source, Nintendo DS emulator. Using the emulator ensured that eye-tracking could be done in a consistent manner, that is, both the game and the comic could be read on screen, eliminating confounds in the presentation of the narrative. This also improved accuracy in the analysis, which is most successful when eye-tracking on a fixed screen as opposed to a mobile, handheld device.

Eye-Tracking Apparatus

A Tobii X60 eye-tracking system and Tobii Studio software were used to collect real-time eye movement data, using a computer to present the materials to be read/played. The Tobii X60 eye-tracking cameras track viewers' eye-movements, capturing fixations, fixation durations, and saccades as they view information on the computer screen. The eye-tracker uses a small camera bar that sits under the screen, and the reader simply sits in front of it and views what is presented, while remaining able to use the computer's mouse and keyboard.

Three different measures were derived from the eye-tracking recordings: task duration, fixation duration, and fixation count. Fixation count is a tally of the number of fixations in a given AOI. It is generally accepted that higher numbers on these measures indicate greater difficulty or more cognitive resources being used to process information in the AOI. However, while eye-tracking shows how long one attends to a given area, it does not tell us about the success of comprehending what is being looked at (De Koning, Tabbers, Rikers, & Paas, 2010), and thus must be complemented with other performance measures.

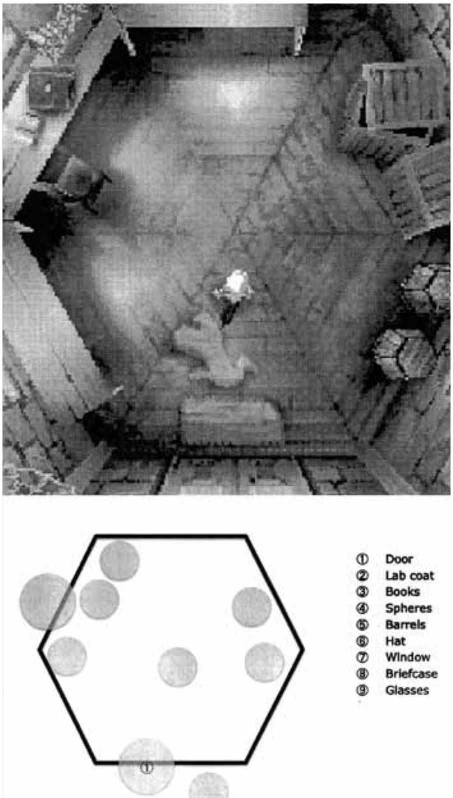
Comprehension Measures

A comprehension test was designed by the authors to measure both literal (15 multiple-choice, three open-ended questions) and inferential comprehension of the story (seven short-answer questions). The questions were developed by experts in reading/literacy education, and vetted by three graduate assistants with knowledge about literal and inferential comprehension, including two who previously taught reading/language arts in public schools. Among the literal questions, nine could be answered only by reading the text, while four could be answered only by using information in the images. The remaining questions could be answered using information that occurred in both the text and the images. All inferential questions asked participants to explain their answers. Participants were encouraged to answer as many questions as possible and were told that they could guess if they were not sure. There was no time limit; participants completed all questions.

Visual Recall Task

In one part of the narrative, the main character searches a small hexagonal cottage filled with items that were clues about her father's whereabouts. To capture information about participants' retention of the items and information found, after reading/playing the narrative, they were asked

Figure 2. Room Image from the Game and Comic (Top) and Corresponding Mapping Task Question



to indicate the items' position in the cottage on a hexagonal drawing/representation of the room (see Figure 2).

Procedure

Data collection took place in the school's computer lab. On arrival, participants were placed in front of a computer and positioned at an appropriate distance from the screen and eye-tracking camera. A research assistant introduced the eye-tracking camera and the participants practiced reading with a similar, sample narrative (either a comic book or a game). After the practice activity, participants put on headphones and watched the introductory video about the narrative they were about to read/play, and were able to ask questions about the information presented. Then participants either read the comic book or played the video game. There was no time limit. After completing the activity, they were given the comprehension measures and the visual recall task described above.

RESULTS

Measures of Pre-Task Group Equivalence

After participants were randomly assigned to one of the two groups (Game Group, GG; Comic Group, CG), their scores from the *Iowa Tests of Basic Skills* (ITBS; Hoover, Dunbar, & Frisbie, 2005) and eye-tracking data sampling rates were used to check for equivalence between groups. An independent samples *t*-test found the two groups to be equivalent on ITBS Reading Total and Language Total scores, with no statistically significant difference: Reading Total score, CG $M = 61.40$, $SD = 23.39$; GG $M = 60.30$, $SD = 30.21$; Language Total score, CG $M = 64.20$, $SD = 26.1$; GG $M = 62.90$, $SD = 30.97$. The two groups were also considered to be equivalent on eye-tracking sampling rate. Sampling rates above 80% are acceptable in terms of accuracy. An independent samples *t*-test found the two groups did not differ significantly on sampling rate: CG $M = 91.10$, GG $M = 91.10$.

To determine if knowledge related to prior use of the Nintendo DS or familiarity with comic books differed across groups and might therefore affect results, participants were asked about Nintendo DS ownership and use, and comic book reading. There was no statistical significance on number of participants who owned and used the Nintendo DS across groups (CG = 9, GG = 8) or who reported reading comics (CG = 9, GG = 9).

Comprehension Findings

The first research question investigated the impact of the presentation format of a narrative on middle school students' comprehension of the story. This was explored by using comprehension tests presented to participants after they had read the comic or played the game. Two types of comprehension questions were analyzed: literal and higher-order. Means and standard deviations for all measures discussed below appear in Table 1.

Table 1. Means and Standard Deviations for the Measures Used in Each Condition

MEASURES	CONDITION			
	Comic Group (CG) $n = 10$		Game Group (GG) $n = 10$	
COMPREHENSION MEASURES	Mean	*SD*	Mean	*SD*
Overall Literal Comprehension (n=18 items)	12.60	3.02	15.10	1.10
Text Only Literal Comprehension (n = 9 items)	7.30	1.49	8.50	0.53
Visual Only Literal Comprehension (n = 4 items)	2.50	1.08	3.40	0.52
Text + Visual Literal Questions (n = 5 items)	2.80	1.03	3.20	0.63
Higher-Order Comprehension (n = 7 items)				
accuracy	4.90	0.74	5.10	1.45
number of details	4.80	2.52	7.30	3.30
word count	86.10	25.55	97.30	26.16
Visual Recall Test	3.40	1.89	6.60	2.27
EYE-TRACKING MEASURES (seconds)	Mean	*SD*	Mean	*SD*
Fixation Duration Sum (Text Only Literal)	82.63	35.74	112.23	25.86
Fixation Duration Mean (Text Only Literal)	0.59	0.10	0.45	0.11
Fixation Count Mean (Text Only Literal)	15.90	5.64	28.57	3.04
Fixation Duration Sum (Image Only Literal)	5.43	1.87	22.94	9.52
Fixation Duration Mean (Image Only Literal)	0.33	0.06	0.35	0.04
Fixation Count Mean (Image Only Literal)	4.67	1.30	15.92	6.08

Literal Questions

An independent-samples t-test was conducted to examine the possible differences between the CG and the GG on the number of correct responses given to the literal questions. There is a statistically significant difference between the GG's and CG's overall literal question accuracy scores on the 18 questions ($t = 2.455$, $p < 0.05$), with the GG answering more questions accurately than the CG.

Higher-Order Questions

Higher-order questions were analyzed based on accuracy, level of detail, and word count (see Kinzer et al., 2011 for detailed procedures on this scoring system). A correct answer received one point. An independent samples *t*-test indicated no statistically significant difference between the CG and the GG. Level of detail was scored by giving one point for an accurate and relevant descriptor, or an appropriate reason. An independent samples *t*-test indicated no significant difference between the CG and the GG on level of detail scores, or for overall word count, which was examined within the inferential comprehension analysis because the provision of more details would be confirmed by a higher number of words used.

The second research question explored whether or not the groups differed in their comprehension of information that appeared only in the narrative's text and only in the narrative's images. Independent sample *t*-tests found a statistically significant difference between the number of accurate responses to literal questions that could be answered only by reading the text ($t = 2.395$, $p < 0.05$) and questions that could be answered only by attending to the images ($t = 2.377$, $p < 0.05$). For the questions that could be answered using information in either the text or image, there was no statistically significant difference between groups.

Visual Recall Task

To complement the multiple-choice and short-answer comprehension questions, we investigated differences between the two groups on visual recall of information. This required remembering story-relevant items appearing in a room within the narrative, and their locations. The room contained nine objects, all clues to solving the mystery in the narrative. The GG remembered significantly more object locations than did the CG ($t = -3.420$, $p < 0.05$).

Eye-Tracking Findings

The third research question explored the participants' attention to text and images in the narrative. Eye-tracking captured fixation points and duration to address this question and allows discussion of the possible effect of participants' attention on their comprehension of the narrative. Using such data is consistent with the majority of multimedia learning studies (e.g., see van Gog & Scheiter, 2010). Using total fixation time as a measure of time on task, it was found that the GG spent more time in the game ($M = 665.0$ seconds, $SD = 130.30$) than the CC group spent reading the comic ($M = 437.4$ seconds, $SD = 51.95$).

Fixation Duration on Text

Using total fixation duration as a measure of how long participants looked at the text and images reveals that the GG spent more time looking at the text in the narrative ($M = 271.52$, $SD = 78.32$) than did the CG ($M = 225.38$, $SD = 112.17$). This difference was not statistically different ($t = 1.067$; $p > 0.05$). However, as the total time within the game and comic differed, a percentage of overall time is a more appropriate unit of analysis. The CG fixated on the text 51.50% ($SD = 24.46$) of their total time, while the GG fixated on text 41.16% ($SD = 6.97$) of the time; there was no statistically significant difference between groups ($t = 1.286$; $p > 0.05$).

Text-Only Questions

There were nine questions that could be answered only by using information contained in the text. Results presented earlier showed that there was a statistically significant difference between the CG and GG in the number of these questions answered correctly, favoring the GG.

An analysis of fixation duration on the text containing information needed to answer those questions showed that there was a statistically significant difference between the GG and CG on total fixation durations ($t = 2.122$; $p < 0.05$), mean fixation durations ($t = 2.844$; $p < 0.05$), and mean fixation count ($t = 6.27$; $p < 0.05$). Thus, the participants in the GG looked at the text containing information allowing them to answer these questions longer than did the CG.

Visual-Only Questions

There were four questions that could be answered only by using information contained in the images. Results presented earlier showed that there was a statistically significant difference between the CG and GG in the number of these questions answered correctly, in favor of the GG. An analysis of fixation duration on the images containing information needed to answer those questions shows that although mean fixation duration did not differ significantly, there was a statistically significant difference in total fixation duration between the GG and CG ($t = 5.708$; $p < 0.001$). The GG also had statistically significantly more attention points (fixation count) on those images than did the CG ($t = 5.721$; $p < 0.001$).

Comprehension and Fixations

Next, we wanted to investigate the differences in eye-tracking metrics for participants who answered questions correctly versus incorrectly. To do so, we divided students into two groups, high accuracy (HA, above the mean score) and low accuracy (LA, below the mean score), for both text-only and visual-only literal questions. For questions that could be answered only by reading the text, we found no statistically significant difference on fixation durations (mean and sum) and fixation count between the HA ($n = 10$) and the LA ($n = 10$) groups. However, when the two groups were created based on the number of visual-only questions they answered correctly, five people were in the LA group and 15 people in the HA group. Because of the small visual-only LA group size, inferential statistics are problematic, though we report them here for informational purposes: Although there was no statistically significant difference on fixation duration means, there were statistically significant differences on the total fixation duration ($t = 3.48$; $p < 0.05$) and mean fixation count ($t = 3.46$; $p < 0.05$). A focus on perhaps more relevant, descriptive statistics illustrates that students with high accuracy scores spent nearly three times as long looking at the relevant area of interest (see Table 2). They also tended to fixate more often, as measured by fixation count, on the relevant part of the image.

Table 2. Descriptive Statistics on Eye-Tracking Measures for High Accuracy and Low Accuracy Groups

	Condition							
	Text-Only Questions				**Image-Only Questions**			
	High Accuracy Scores (*n* = 10)		Low Accuracy Scores (*n* = 10)		High Accuracy Scores (*n* = 15)		Low Accuracy Scores (*n* = 5)	
Measures	M	*SD*	M	*SD*	M	*SD*	M	*SD*
Fixation Duration Total (seconds)	104.13	37.37	90.73	30.61	16.94	11.67	5.92	2.16
Mean Fixation Duration (seconds)	0.49	0.11	0.55	0.14	0.35	0.05	0.32	0.07
Mean Fixation Count	24.72	8.37	19.74	6.80	12.06	7.49	5.02	1.43

DISCUSSION

This study examined readers' attention to and comprehension of a narrative presented in a comic book and an adventure video game. When understanding multimodal texts such as games or comics, information is available from various combinations of sources including printed text only, images only, or both text and images. Thus, when a reader does not understand information that is presented, it is often difficult to know: (a) whether or not the reader has looked at the text or image containing the information, or (b) has looked at it but did not process it in a way that leads to understanding. This study addressed the former possibility, using eye-tracking measures to determine whether or not comprehension differences might be due to the sixth-grade participants not looking at the area containing the information tested. Of course, if participants looked at the appropriate locus of information and did not provide correct answers, there are clear implications for future research that examines the processing of the information.

Additionally, different modes of texts, in this case a comic and a video game, have distinct requirements for interaction and use, and present information in different ways. We noted earlier that comics have been advocated in reading instruction for some time, and that more recently video games' use in instruction has been suggested. Yet, before implementing video games in instructional settings, we must know more about how students interact with them, whether or not they provide a means for understanding narrative elements, and in what way(s). Questions that must be considered include whether or not players attend to text in video games or ignore text in favor of the images in the game. In summary, this aspect of the study explored the potential relationship between attention to information-bearing units of text and graphics, as measured by eye fixation duration, in a narrative presented in a comic book and a video game.

What do the results reveal about whether or not sixth-grade readers are looking at the text and images within narratives? To begin with, the cumulative eye fixation totals show that the GG spent more time playing the game than did the CG in reading the comic largely because of the game's interactive features and ability to explore regions of the game world. For example, when the GG's avatars were in the hexagonal room (see Figure 2) they were able to move around, touch objects that were also described in text boxes, return to objects previously observed, and so on, while the CG saw images of the room and the objects and read about them, but did not directly interact with them. The time spent in the video game as opposed to the comic was approximately 11.1, as opposed to 7.3 minutes. Clearly, a difference in time on task was found, but how did this impact narrative comprehension? Did the increased time on task for the GG group arise from a tension between narrative and interactivity? Did the game play distract players' focus, resulting in decreased understanding of the narrative?

Our results indicate this was not the case, as the GG answered significantly more literal post-test comprehension questions correctly than did participants in the CG. The GG also answered more post-test questions that targeted information located only in the game's text. Furthermore, the GG answered more post-test questions correctly that targeted information located only in the video game's images. Overall, for the literal portion of the comprehension results, the GG performed better on the questions regardless of how the salient information was presented. Regarding the higher-order comprehension questions, the results reveal a similar trend but the difference was not statistically significant. Overall, the post-test comprehension measures indicate that the video game format did not detract from understanding the narrative.

Given these results, an obvious follow-up question is why students' comprehension scores were different between the two groups. Were the text and graphics attended to differently at the visual level? The eye-tracking results help explore this. The results clearly show that the GG spent more time "playing" than the comic group spent "reading." But when the fixation duration on images is considered as a percentage of total time, there is no statistical difference between the two groups. Despite the lack of statistical significance, the results do show that the GG spent a lower percentage (41.5%) of their time fixated on text than the CG group (51.5%). To be clear, the students in the comic book group spent a greater amount of their total fixation time on text related to the text-only questions, but scored significantly lower on the comprehension questions related to this text. When it came to fixating on images in the narrative, the students in the GG looked at the images for a significantly longer time and subsequently answered more image-related questions correctly.

These results suggest that by almost all measures, the GG performed differently than the CG in their respective narrative format. The GG also outscored the CG in the comprehension measure. Are these results due to some inherent difference in the narrative formats? Recall that the two groups had no statistical difference in their reading ability as measured by standardized test scores. What properties of the two narrative formats might lead to these differences?

Possibly, the more exploratory, open mind-set promoted by adventure games helps students comprehend narrative elements. Players can roam around the environment exploring and investigating the world and the objects in it. This openness is limited only by "boundary components" (Björk & Holopainen, 2005, p. 14), such as rules and goals, which limit what players are allowed to do in the game. It does not take long for an adventure game player to realize that trying to race through the game may be counter-productive, and that a more investigative, slower play style may prove more fruitful. From this perspective, the comic book and the adventure game begin to look different, despite being deliberately designed to be as close as possible in terms of text and graphics. Comic book readers are more limited in action possibilities: They can move from panel to panel, turn the page back and forth, look at the pictures repeatedly, or reread. Beyond this, there is not much "to do" in the narrative space of the comic.

A second possibility relates to the difference in the point of view used in comic books and adventure games. In both formats, the reader/player follows the main character, Ashley, as the story unfolds through text and images. However, in the game the reader/player often must complete or initiate an action that results in the appearance of an image or text, whereas the comic book reader encounters panels as they appear, without overt action. Thus, even though the actual image and text read in each format is the same, interactivity and "distance" from the narrative may vary perceptually and cognitively. For example, throughout the game there are periods when players watch Ashley do things such as talk to other characters. At other times, however, players take on the role of Ashley more directly, and must decide what they, operating as Ashley, do by making choices about what actions to take.

Examples of these choices include whether to try to unlock a door, pick up a stone, or examine an old book. In the comic, readers see the same images and read the same text, but without accompanying physical interactivity. The difference might result in a kind of "narrative distance" (see Booth, 1983) between the reader/player and the events of the story. Said another way, in the comic book format the reader is "once removed" from the action, whereas in the adventure game, the player may need to decide and take action before a particular event occurs. Could these processes help players better comprehend the narrative in the game? Previous work in social psychology shows that actors who perform behaviors, compared to observers who watch, often come to different causal analyses of the same situations (Gerrig & Jacovina, 2009, p. 229).

Results from the visual mapping task, where participants were compared on their ability to recall the location of specific objects found inside a room, are informative when considering possible answers to the above question. The CG read about the objects and saw their location in repeated images when reading the comic. The GG was able to move Ashley around the room, have her touch the objects, walk back and revisit objects, and generally consider the room as an active space within which they could control Ashley's exploration. On the post-test that asked participants to place objects in a representation of the room (see Figure 2), the GG remembered significantly more objects' locations. Taken together with the results discussed above, the impact of decision processes and aspects that relate to theories of embodied cognition (see Glenberg & Kaschak, 2002) may be fruitful areas for further research. We note also that eye-tracking, while labor intensive, provides a method of answering important questions related to the above discussion, especially in multimodal, digital environments such as encountered in video games.

CONCLUSION

Today's teachers must be flexible and adapt their lessons and instructional materials to capitalize on students' needs and interests as a means of inspiring them towards literacy. Popular culture materials, such as comic books and video games, may be underrepresented vehicles for doing so. Materials popular with children in free-reading time have been used for many years as motivating tools to encourage comprehension, critical thinking, characterization, plot, vocabulary and other aspects of literacy (e.g., see Fader & McNeil, 1968, for a discussion on the positive effects of using high-interest books and comics to enhance literacy instruction). Similarly, Krashen (2005) notes the potential beneficial effects of comic books used in reading instruction, especially for children who do not read well, and notes, "There is evidence suggesting that comic book reading can be a conduit to 'heavier' reading."

We feel that appropriate, adventure video games may also provide a means of using popular culture materials in reading instruction, although additional research in this area is needed to provide evidence about literacy practices within games. The work reported here shows that comprehension of narrative is part of adventure video game play and may be greater in game players than in comic book readers, and provides evidence that game play may have beneficial educational uses in literacy development.

ENDNOTE

This work was funded in part by Microsoft Research through the Games for Learning Institute. The content and opinions herein are the authors' and may not reflect the views of Microsoft Research, nor does mention of trade names, products, or organizations imply endorsement.

REFERENCES

Adams, E. (2006). *Fundamentals of game design* (2nd ed.). Indianapolis, IN: New Riders Publishing.

Allen, K., & Ingulsrud, J. E. (2007). *Strategies used by children when reading manga.* Paper presented at the Annual Congress of the Applied Linguistics Association of Australia, New South Wales, Australia.

Anglin, G. J., Vaez, H., & Cunningham, K. L. (2004). Visual representation and learning: The roles of static and animated graphics. In D. Jonassen (Ed.), *Handbook of research on educational communications and technology,* (2nd ed., pp. 865-916). Mahwah, NJ: Lawrence Erlbaum Associates.

Beymer, D., Orton, P. Z., & Russell, D. M. (2007). An eye tracking study of how pictures influence online reading. In C. Baranauskas, P. Palanque, J. Abascal, & S. Barbosa (Eds.), *Human-computer interaction – INTERACT 2007* (pp. 456-460). Heidelberg, Germany: Springer.

Blair, H. A., & Sanford, K. (2004). Morphing literacy: Boys reshaping their school-based literacy practices. *Language Arts, 81*(6), 452-460.

Björk, S., & Holopainen, J. (2005). *Patterns in game design*. Hingham, MA: Charles River Media, Inc.

Booth, W. C. (1983). *The rhetoric of fiction* (2nd ed.). Chicago, IL: University of Chicago Press.

Carney, R. N., & Levin, J. R. (2002). Pictorial illustrations still improve students' learning from text. *Educational Psychology Review, 14*(1), 5–26.

Carroll, P. J., Young, J. R., & Guertin, M. S. (1992). Visual analysis of cartoons: A view from the far side. In K. Rayner (Ed.), *Eye movements and visual cognition: Scene perception and reading* (pp. 444-461). New York, NY: Springer-Verlag.

Chun, C. W. (2009). Critical literacies and graphic novels for English-language learners: Teaching Maus. *Journal of Adolescent & Adult Literacy, 53*(2), 144–153.

CiNG. (2005). *Trace Memory* [Nintendo DS game]. Fukoko, Japan.

Coan, J. A., & Allen, J. J. B. (2007). Organizing the tools and methods of affective science. In J. A. Coan & J. J. B. Allen (Eds.), *Handbook of emotion elicitation and assessment* (pp. 3-6). New York, NY: Oxford University Press.

Commeyras, M. (2009). Drax's reading in Neverwinter Nights: With a tutor as henchman. *eLearning, 6*(1), 43-53.

DeSmuME (2009). DeSmuME Team (Version 0.9.8) [Software]. Available from http://desmume.org/download/

De Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2010). Attention guidance in learning from a complex animation: Seeing is understanding? *Learning and Instruction, 20*(2), 111-122.

Eikmeier, G. M. (2008). D'oh! Using the Simpsons to improve student response to literature. *English Journal, 97*(4), 77-80.

Emmison, M., & Smith, P. (2000). *Researching the visual: Images, objects, contexts and interactions in social and cultural inquiry*. New York, NY: SAGE.

Fader, D. N., & McNeil, E. B. (1968). *Hooked on books: Program and proof.* New York, NY: Berkeley Publishing Corp.

Fang, Z. (1996). Illustrations, text, and the child reader: What are pictures in children's storybooks for? *Reading Horizons, 37*(2), 137–142.

Gambrell, L. B., & Jawitz, P. B. (1993). Mental imagery, text illustrations, and children's story comprehension and recall. *Reading Research Quarterly, 28*(3), 265-273.

Gee, J. P. (2003). *What video games have to teach us about learning and literacy.* New York, NY: Palgrave Macmillan.

Gentile, D. A., Lynch, P. J., Linder, J. R., & Walsh, D. A. (2004). The effects of violent video game habits on adolescent hostility, aggressive behaviors, and school performance. *Journal of Adolescence, 27*(1), 5-22.

Gerrig, R. J., & Jacovina, M. E. (2009). Reader participation in the experience of narrative. In B. H. Ross (Ed.), *The psychology of learning and motivation, 51*, (pp. 223-254). New York, NY: Academic Press.

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review, 9*(3), 558–565.

Goldberg, H. J., & Kotval, X. P. (1999). Computer interface evaluation using eye movements: Methods and constructs. *International Journal of Industrial Ergonomics, 24*(6), 631-645.

Harris, C. (2005). Trace Memory: The DS interactive mystery finally hits the US. Was it worth the wait? *IGN.* Retrieved from http://ds.ign.com/articles/652/652873p1.html

Hegarty, M., & Just, M. A. (1993). Constructing mental models of machines from text and diagrams. *Journal of Memory and Language*, *32*(6), 717-742.

Hoover, H. D., Dunbar, S. B., & Frisbie, D. A. (2005). *Iowa tests of basic skills*. Chicago, IL: Riverside Publishing Company.

Hruschak, P. (2006). Player 1: A DS double whammy. *CiN Weekly*. Retrieved from http://web.archive.org/web/20060603141054/http://www.cinweekly.com/apps/pbcs.dll/article?AID=/20060111/ENT09/601110350/1063

Jennett, C., Cox, A. L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., & Walton, A. (2008). Measuring and defining the experience of immersion in games. *International Journal of Human Computer Studies*, *66*(9), 641-661.

Jewitt, C., & Oyama, R. (2001). Visual meaning: A social semiotic approach. In T. van Leeuwen & C. Jewitt (Eds.), *Handbook of visual analysis* (pp. 134-156). London, England: Sage.

Kiili, K., & Ketamo, H. (2010, October 22). Eye-tracking in educational game design. In Meyer, B. (Ed.). *Proceedings of the 4th European conference on game based learning* (pp. 160-167). Copenhagen, Denmark: The Danish School of Education, Aarhus University.

Kinzer, C. K., Hoffman, D. L., Turkay, S., Gunbas, N., & Chantes, P. (2011). Exploring motivation and comprehension of a narrative in a video game, book, and comic book format. In P. J. Dunston, L. B. Gambrell, K. Headley, S. King Fullerton, P. M. Stecker, V. Gillis, & C. C. Bates (Eds.), *60th Yearbook of the Literacy Research Association* (pp. 263-278). Oak Creek, WI: Literacy Research Association.

Krashen, S. (2004). *The power of reading: Insights from the research* (2nd ed.). Portsmouth, NH: Libraries Unlimited.

Krashen, S. (2005). The "decline" of reading in America, poverty and access to books, and the use of comics in encouraging reading. *Teachers College Record*. Retrieved from http://www.sdkrashen.com/articles/decline_of_reading/all.html

Kress, G., & van Leeuwen, L. (2006). *Reading images: The grammar of visual design.* New York, NY: Routledge.

Knoepfle, D. T., Wang, J. T., & Camerer, C. F. (2009). Studying learning in games using eye-tracking. *Journal of the European Economic Association, 7*(2-3), 388-398.

Kuechenmeister, B. (2009). Reading comics rhetorically: Orality, literacy, and hybridity in comic narratives, *Scan*, *6*(1). Retrieved from http://scan.net.au/scan/journal/display.php?journal_id=132

Law, E. L. C., Mattheiss, E., Kickmeier-Rust, M. D., & Albert, D. (2010). Vicarious learning with a digital educational game: Eye-tracking and survey-based evaluation approaches. In G. Leitner, M. Hitz, & A. Holzinger (Eds.), *Lecture notes in computer science - HCI in work and learning, life and leisure*, *6389* (pp. 471-488). Berlin: Springer.

Levin, J. R., Anglin, G. J., & Carney, R. N. (1987). On empirically validating functions of pictures in prose. In D. M. Willows & H. A. Houghton (Eds.), *The Psychology of Illustration: Basic Research* (pp. 51–85). New York, NY: Springer-Verlag.

Liu, J. (2004). Effects of comic strips on L2 learners' reading comprehension. *TESOL Quarterly*, *38*(2), 225-243.

Martin, C., & Steinkuehler, C. (2011, June). *Information literacy and online reading comprehension: Two interconnected practices.* Paper presented at the Games+Learning+Society Conference, Madison, WI.

Mat Zain, N., Abdul Razak, F., Jaafar, A., & Zulkipli, M. (2011). Eye-tracking in educational games environment: Evaluating user interface design through eye tracking patterns. In H. Zaman, P. Robinson, M. Petrou, P. Olivier, T. Shih, S. Velastin & I. Nyström (Eds.) *Proceedings of the Second International Conference on Visual Informatics: Sustaining research and innovations - Volume Part II* (pp.64-73). Berlin/ Heidelberg: Springer.

Mayer, R. E. (2009). *Multimedia learning*. (2nd ed.). New York, NY: Cambridge University Press.

Mayer, R. E. (2011). Multimedia learning and games. In S. Tobias & J. D. Fletcher (Eds.), *Computer games and instruction* (p. 281-305). Charlotte, NC: Information Age Publishing.

McCloud, S. (1993). *Understanding comics: The invisible art*. Northampton, MA: Kitchen Sink Press.

Miller, W. A. (1938). Reading with and without pictures. *The Elementary School Journal, 38*(9), 676-682.

Moje, E. B., Overby, M., Tysvaer, N., & Morris, K. (2008). The complex world of adolescent literacy: Myths, motivations, and mysteries. *Harvard Educational Researcher, 71*(1), 107-154.

Nakazawa, J. (2004). Manga (comic) literacy skills as determinant factors of manga story comprehension. *Manga Studies, 5,* 7-25.

Norton, B. (2003). The motivating power of comic books: Insights from Archie comic readers. *The Reading Teacher, 57*(2), 140-147.

Omori, T., Igaki, T., Ishii, T., Kurata, K., & Masuda, N. (2004). *Eye catchers in comics: Controlling eye movements in reading pictorial and textual media.* Paper presented at the 28th International Congress of Psychology. Beijing, China.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology, 45*(3), 255-287.

Parish, J. (2005). *Nintendo presents: My first adventure game.* Retrieved from http://www.1up.com/do/reviewPage?cId=3144099&did=1

Peeck, J. (1993). Increasing picture effects in learning from illustrated text. *Learning and Instruction, 3*(3), 227–238.

Ranker, J. (2007). Using comic books as read-alouds: Insights on reading instruction from an English as a second language classroom. *The Reading Teacher, 61*(4), 296-305.

Rayner, K. (1998). Eye movements and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372-343.

Rayner, K., Rotello, C. M., Stewart, A. J., Keir, J., & Duffy, S. A. (2001). Integrating text and pictorial information: Eye movements when looking at print advertisement. *Journal of Experimental Psychology: Applied, 7*(3), 219-226.

Rideout, V. J., Foehr, U. G., & Roberts, D. F. (2010). *GENERATION M2: Media in the lives of 8- to 18-year-olds.* Menlo Park, CA: Henry J. Kaiser Family Foundation. Retrieved from http://www.kff.org/entmedia/upload/8010.pdf

Rose, T. L. (1986). Effects of illustrations on reading comprehension of learning disabled students. *Journal of Learning Disabilities, 19*(9), 542–544.

Samuels, S. J. (1967). Attentional process in reading: The effect of pictures on the acquisition of reading responses. *Journal of Educational Psychology, 58*(6), 337-342.

Schmidt-Weigand, F., Kohert, A., & Glowalla, U. (2010). A closer look at split visual attention in system- and self-paced instruction in multimedia learning. *Learning and Instruction*, *20*(2), 100-110.

Sherman, R., & Wright, G. (1996). Orchestra. *Reading Improvement, 33*(2), 124-128.

Sipe, L. R. (1998). How picture books work: A semiotically framed theory of text-picture relationships. *Children's Literacy Education, 29*(2), 97-108.

Staff Reviewer. (2005). GamesMaster deciphers the latest fiendish puzzler for DS. *Gamesradar.com*. Retrieved from http://tinyurl.com/tracememory

Starr, M. S., & Rayner, K. (2001). Eye movements during reading: Some current controversies. *Trends in Cognitive Sciences, 4*(4), 156-163.

Steinkuehler, C. (2007). Massively multiplayer online gaming as a constellation of literacy practices. *eLearning, 4*(3), 297-318.

Steinkuehler, C., Compton-Lilly, C., & King, E. (2009, April). *Literacy practice & reading performance in the context of MMO games.* Presented at the Annual Meeting of the American Educational Research Association, Denver, CO.

Steinkuehler, C. (2011). *The mismeasure of boys: Reading and online videogames* (WCER Working Paper No. 2011-3). Retrieved from http://www.wcer.wisc.edu/publications/workingPapers/papers.php

Tiemensma, L. (2009). *Visual literacy: To comics or not to comics? Promoting literacy using comics.* Paper presented at the World Library and Information Congress: 75th IFLA General Conference and Council, Milan, Italy.

Turkay, S., Kinzer, C., Hoffman, D., Gunbas, N., & Nagle, C. (2010). A snapshot on youths' activities on online gaming forums: Internet and informal learning. In *Proceedings of the World Conference on Educational Multimedia, Hypermedia and Telecommunications 2010* (pp. 3987-3992). Chesapeake, VA: AACE.

Ujiie, J., & Krashen, S. D. (1996). Comic book reading, reading enjoyment, and pleasure reading among middle class and chapter 1 middle school students. *Reading Improvement*, *33*(1), 51-54.

van Gog, T., & Scheiter, K. (2010). Eye tracking as a tool to study and enhance multimedia learning. *Learning and Instruction, 20*(2), 95-99.

Vernon, M. D. (1953). The value of pictorial illustrations. *British Journal of Educational Psychology*, *23*(3), 180-187.

Wang, A. Y., & Newlin, M. H. (2000). Characteristics of students who enroll and succeed in psychology web-based classes. *Journal of Educational Psychology, 92*(1), 137–143.

Watkins, J. K., Miller, E., & Brubaker, D. (2004). The role of visual image: What are students really learning from pictorial representations? *Journal of Visual Literacy, 24*(1), 23–40.

Watson, D. J. (1989) Defining and describing whole language. *The Elementary School Journal*, *90*(2), 129-141.

Weintraub, S. A. (1966). Illustrations for the beginning reader. *Reading Teacher, 20*(1), 61-67.

Woodard, E. H. & Gridina, N. (2000). *Media in the Home 2000: The Fifth Annual Survey of Parents and Children.* The Annenberg Public Policy Center of the University of Pennsylvania.

Wright, G., & Sherman, R. (1994). What is black and white and read all over? The funnies! *Reading Improvement*, *31*(1), 37-48.

Wright, G., & Sherman, R. (1999). Let's create a comic strip. *Reading Improvement, 36*(2), 66-71.