Magliano, J. P., Loschky, L. C., Clinton, J., & Larson, A. M. (in press). Is reading the same as viewing? An exploration of the similarities and differences between processing text- and visually based narratives. In B. Miller, L. Cutting, and P. McCardle (Eds.), Unraveling the Behavioral, Neurobiological, & Genetic Components of Reading Comprehension, pp. 78-90. Baltimore, MD: Brookes Publishing Co.

CHAPTER 8

Is Reading the Same as Viewing?

An Exploration of the Similarities and Differences Between Processing Text- and Visually Based Narratives

Joseph P. Magliano, Lester C. Loschky, James A. Clinton, and Adam M. Larson

The title of this chapter explores to what extent there are shared cognitive systems that support the processing of narratives across print and visual media. An initially obvious answer to the question is no, given that viewing images and reading texts involve different cognitive and brain systems during encoding. In fact, we contend that there are aspects of how narratives are conveyed across media that may lead to differences in encoding that are less obvious. On the other hand, if we assume that mental models provide the basis of comprehension across modalities of experiences (McNamara & Magliano, 2009), surely there are shared cognitive systems that lead to the construction of these mental models (Gernsbacher, 1990; Gernsbacher, Varner, & Faust, 1990; Magliano, Radvansky, & Copeland, 2007).

Consider the shot sequence from the James Bond movie *Moonraker* (Broccoli & Gilbert, 1979), shown in Figure 8.1, which constitutes about 15 seconds of film. Just prior to this sequence, we find Bond on an airplane filled with enemy agents who want to kill him. Bond throws one of the bad guys (who has a parachute) out of the plane and then subsequently gets pushed out of the plane himself (without a parachute). Most viewers of this sequence infer that Bond has the goal to steal the parachute and will be successful in doing so at the start of Shot 6, when Bond starts his dive (Magliano, Dijkstra, & Zwaan, 1996). Upon review of the shot sequence in Figure 8.1, one can see that the sequence is carefully constructed to constrain the possible interpretation of the narrative events. More important, it illustrates some conventions in filmmaking that may not have correlates in text; but at the same time, the sequence also illustrates that there must be overlapping processes between film and text.

Consider Shots 1 and 2, which are extreme, high-angle, long shots of two figures in the air and heading toward the earth. These shots establish the nature of a complication or problem for Bond to solve, specifically that he is falling through the air without a parachute. Shot 2 also introduces a potential solution to that problem, namely, the parachute possessed by the bad guy. Although processing the events conveyed in these camera shots may be different in many ways from processing analogous sentences in a text-based version of the narrative, in doing so viewers need to engage in mapping processes across the images that are also central to text

BRP-MILLER-13-0103-008.indd 78 17/07/13 1:39 PM

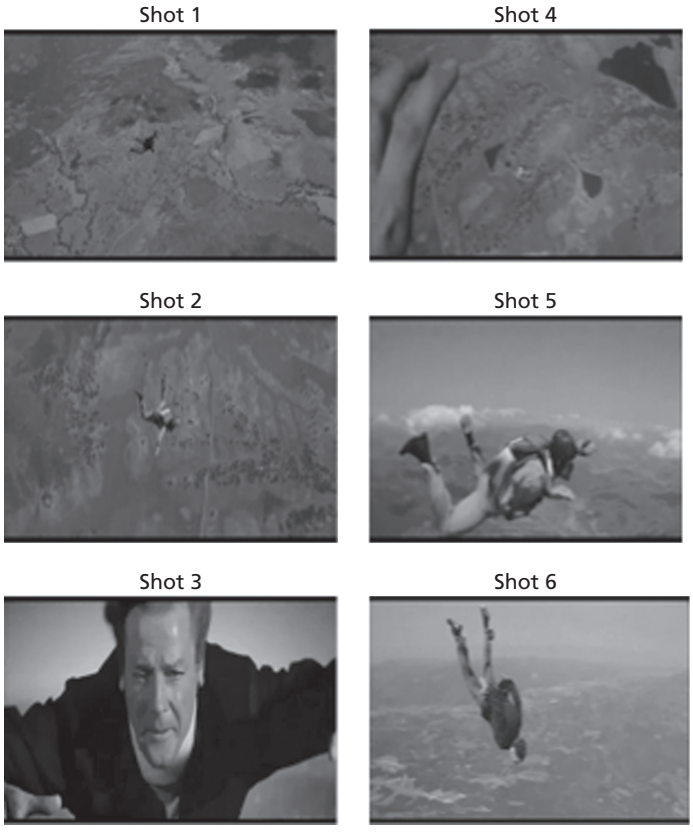

Figure 8.1. Shot summary of a scene from *Moonraker*.

comprehension (McNamara & Magliano, 2009). For example, viewers must establish that the figures in Shots 1 and 2 are the previously seen Bond and the bad guy, respectively, and reinstate knowledge of which of these entities has or does not have a parachute. These mapping processes are akin to the kinds of bridging inferences we know that readers generate (e.g., anaphoric reference; see Graesser, Singer, & Trabasso, 1994, for an extensive review).

Another issue to consider is that understanding the internal states of characters (goals, beliefs, emotions) is an important part of narrative comprehension (e.g., Graesser, Singer, & Trabasso, 1994). However, text and film may differ in the extent to which internal states can readily be conveyed. Authors can rely on a variety of narrative conventions to provide explicit information regarding the internal states of characters (e.g., free indirect speech, omniscient narration). However, filmmakers usually rely on dialogue to explicitly convey internal states of characters. Other techniques, such as voice-over, can be used to explicitly convey internal

BRP-MILLER-13-0103-008.indd 79 17/07/13 1:39 PM

states of characters, but these are rarely used and not well received by critics. That said, there are filmmaking conventions for conveying internal states that build upon editing and actors' facial expressions, body postures, and actions to help the viewer infer the internal states of a protagonist. For example, Shots 3 and 4 illustrate one important convention, namely, a point-of-view shot sequence. This is a shot sequence that shows a close-up of a character's face, followed by eye-line match shots (i.e., a camera shot from the implied location of the characters' eyes). This shot sequence is critical for conveying the internal states of characters (Bordwell & Thompson, 2003) but will always require the inference that the second shot is seen from the viewpoint of the character shown in the first shot. Based on that inference, in the case of the example in Figure 8.1, viewers must further infer that Bond is cognizant of the bad guy with the parachute, arguably a necessary condition for his generating the goal of stealing the parachute.

As this example illustrates, to be able to give a detailed answer to the question of whether reading and viewing narratives are the same requires a detailed systematic investigation. We contend that such an investigation is critical for directing empirical studies in order to answer the question meaningfully. The goal of this chapter is to provide the start of such an investigation. We structure our discussion around features of the media, the psychological processes involved in processing and comprehending narratives across media, and research comparing processing across media.

FEATURES OF THE MEDIA

Any discussion regarding the similarities and differences in the cognitive processes that support narrative comprehension across media should start with a discussion of the nature of the media. Such an exploration can yield hypotheses suggesting where and when there will be divergences and convergences between media, which can then lead to future empirical studies. To this end, we have identified four dimensions that warrant consideration across the media of texts, graphic narratives, and films, which are shown in Table 8.1. These dimensions are by no means exhaustive; rather, they reflect the dimensions that we feel warrant consideration, given the goals of this chapter.

One fundamental issue to consider is the *minimal production unit of meaning*, the smallest unit that needs to be processed to understand what is happening "now" in the narrative, namely, the salient events that make up the narrative plot. With narrative text, an obvious production unit of meaning is the sentence. It is easily recognized (in English) by starting with a capital letter and ending with specific punctuation (e.g., a period). Similarly, in such graphic narratives as comics, an obvious minimal production unit of meaning is the panel, which is also easily recognized by typically being bounded by a box. Finally, in film, a comparable minimal production unit of meaning is the shot, which is easily recognized by a change in camera

BRP-MILLER-13-0103-008.indd 80 17/07/13 1:39 PM

Table 8.1. Considerations regarding the media

Dimension	Text	Graphic narrative	Film
Unit of meaning	Sentence	Panel	Shot
Multimodality	Text	Text and pictures	Dialog, sounds, images
Presentation control	Self-paced	Self-paced	Generally not self-paced
Plot structure	Goal episodes	Goal episodes	Goal episodes

angle and/or an editing boundary (e.g., a fade-out or dissolve). Each of these minimal production units is widely recognized by producers and consumers of these respective media.

Nevertheless, there are certain problems with each of these minimal meaningful units from the standpoint of psychological processes that support comprehension. With text, a single sentence can include multiple clauses, each of which contains a subject and predicate (i.e., a verb phrase). Verbs in these clauses are particularly important in conveying the events that are being conveyed in a narrative (e.g., walking, stopping, and seeing; Zwaan, Langston, & Graesser, 1995). Much research has shown that as readers reach the end of each clause in a text, they update their representation to incorporate that information into the mental model that they are constructing (Gernsbacher, 1990). Thus, from a psycholinguistic standpoint, clauses seem more reasonable as a minimal unit of meaning than a sentence. Alternatively, one could also argue that the individual word is an even more minimal meaningful unit. Eye-movement studies have shown that readers fixate most words in a text sequentially (Rayner, 1998), and ease of word recognition is an important predictor of children's reading development (Rayner, Foorman, Perfetti, Pesetsky, & Seidenberg, 2001).

In the case of graphic narratives, the panel is also a somewhat problematic minimal meaningful unit. A given panel generally contains multiple objects (animate or inanimate) and one or more actions, either of which could be argued to be more appropriate minimal meaningful units. For example, in Figure 8.2 (*The Fantom of the Fair*, Gustavson, 1939), we see a man (the fantom) and a woman, each of which must be recognized, as well as their respective actions (flying, calling), in order to comprehend the event(s) in the panel. Such objects are rather comparable to concrete nouns in text processing, and their actions are comparable to verbs. Similarly, in film, a given shot usually contains multiple objects engaged in one or more actions. Thus, again, objects and their actions could be argued to be more appropriate minimal meaningful units. Furthermore, a given camera shot in film can convey several events, or several shots can be used to convey a single event.

Finally, both graphic narratives and films usually contain language. Consider a panel excerpt from Gustavson's 1939 graphic story *The Fantom of the Fair* (see Figure 8.2). The Fantom has just rescued a circus high-diver, Jane, and he is now flying away from her. Viewers have to integrate the

Figure 8.2. Example panel from *Fantom of the Fair*.

dialogue with the images to understand what is happening now in the narrative moment. The same is true in film, where integration of both visual and linguistic processing is necessary for comprehension. Thus, one can ask whether the sentence (or clause, or word) should also be a minimal meaningful unit in graphic narratives and film.

The next dimension that warrants discussion is the extent to which the narrative is a multimedia experience. Texts obviously are not, whereas graphic narratives and film are. Multimedia experiences can be complex to process because the comprehender has to integrate linguistic and perceptual information, which can place a higher cognitive load on the comprehender (Mayer, 2009). We suspect that graphic narratives are the most complex and resource demanding because they require text reading and pictorial comprehension. Consider again Figure 8.2 from *The Fantom of the Fair* (Gustavson, 1939). The dialog and image contain both convergent and divergent semantic content. An interesting empirical question is the extent to which the semantic overlap between these two sources of information affects comprehension.

It is also important to understand that there are differences in the extent to which comprehenders can control the pacing of processing the narrative. With text and graphic narratives, comprehenders can control the pace at which they read and view. This allows comprehenders to pause, look back, reevaluate, and repair any comprehension problems. On the other hand, when viewing a film there is generally no control over the pacing. If comprehension problems arise, the viewer very often cannot or does not view that part of the film again. The exception is if the viewer has the ability to rewind. Whereas this option is available to individuals watching a film on their own viewing device, films are often watched in theaters or other

BRP-MILLER-13-0103-008.indd 82 17/07/13 1:39 PM

social settings in which rewinding is either not an option or is not generally taken advantage of to repair comprehension problems.

Although the dimensions discussed thus far vary to one degree or another across the modalities, we contend that the final dimension, plot structure, does not. We argue that all narratives are structured around hierarchical goal episodes regardless of media (Trabasso, van den Broek, & Suh, 1989). Based on this assumption, we further assume that there will be commonalities in the mental models constructed by readers and viewers of comparable narratives. We will discuss empirical evidence supporting this claim in the context of the psychological processes involved in processing the different forms of media. One point that we do want to emphasize now, however, is that the different forms of media have different *affordances* in terms of which aspects of the goal episodes can be made explicit. As discussed above, films do not afford ways of explicitly conveying the internal states of characters (outside of dialogue or voice-over narration); thus viewers must instead infer them from the facial expressions, gestures, and actions of the actors. Conversely, this information can be stated explicitly in a text if an author chooses to do so. Given the importance of inferring and tracking goals of characters, one would think that this may make processing complex goals and internal states of characters more complex for narrative films more than for texts. However, basic emotions are readily inferred by facial expressions, and this inferencing skill is learned at an early age (McClure, 2000). Moreover, viewers appear to regularly monitor the goal states of characters when viewing a film (Magliano, Taylor, & Kim, 2005). The extent to which viewers can infer internal state of characters in film (i.e., the kinds of internal states and their complexity) and how they do so warrant further investigation.

FRONT-END AND BACK-END PROCESSES ACROSS MEDIA

We make a broad distinction between what we call front-end and back-end processing of narrative media (Loschky, Magliano, & Larson, 2012). Front-end processes are those involved in extracting information from the visual stream, which lead to the computation of the event model that represents the "now" moment in a narrative. Table 8.2 shows a list of processes that compose front-end processing across the three media considered here. These processes cluster into those processes involved in processing language and those involved in scene perception. It is beyond the scope of this chapter to present a detailed discussion of each of these processes, however. Given that most readers of this volume will be more familiar with the language-based processes than with those that are film based, we do not describe them here. Nonetheless, a brief description of the processes involved in scene perception is warranted.

When constructing a representation of the "now" moment in a scene, front-end processes begin by extracting visual information during each

Table 8.2. Considerations regarding front-end and back-end processes

	Text	Graphic narrative	Film
Front-end			
Processes			
Orthographic	X	X	
Phonological	X	X	X
Lexical–syntactic	X	X	X
Semantics	X	X	X
Gist processing		X	X
Object processing		X	X
Motion processing		(X)	X
Product			
Event model	X	X	X
Back-end			
Processes			
Event segmentation	X	X	X
Inferences	X	X	X
Structure building	X	X	X
Products			
Text base	X	X	?
Situation model	X	X	X

Note: (X) denotes that this process may be present.

eye fixation. In scene perception, eye fixations last, on average, a third of a second, after which the eyes move to fixate a new object or location (Rayner, 1998). During the first 150 ms of the first fixation, front-end processes extract global semantic meaning from the scene, called scene gist (Fei-Fei, Iyer, Koch, & Perona, 2007). For example, the viewer will rapidly understand whether a scene is "natural" or "man made," a "beach" or a "forest" (Loschky & Larson, 2010) and whether it has positive or negative emotional valence (Maljkovic & Martini, 2005). Likewise, the scene gist will also include global information about key objects in the scene, such as whether there are people or animals (Fei-Fei et al., 2007) and some idea of what action is being performed (Larson, Hendry, & Loschky, 2012). This scene gist representation will guide the eyes to fixate other objects to gather more detailed information (Eckstein, Drescher, & Shimozaki, 2006) and form the foundation of the event model. The event model will be further developed by the back-end processes in order to comprehend the dynamic structure of the visual narrative.

One obvious observation that can be derived from Table 8.2 is that there is some shared overlap in processes, but this is most evident for the two forms of visually based narratives. Obviously, reading and scene perception involve some different cognitive and brain systems associated with encoding linguistic versus scenic input. However, there is a growing body of evidence for the perspective that language understanding is "grounded," meaning that the brain systems that are involved in perceiving and moving through the world support the processing of language (e.g.,

BRP-MILLER-13-0103-008.indd 84 17/07/13 1:39 PM

Zwaan, 2004). It is beyond the scope of this chapter to discuss this literature; rather, we will focus on the shared overlap in language processing across the three media.

The primary difference between the linguistic demands of reading (whether it be in the context of texts or graphic narratives) and film is that graphemic processing is minimal in film because most of the language is presented aurally. As such, there are fewer reading demands on the cognitive system during front-end processing of film. Indeed, preliterate children begin to develop narrative comprehension skills by viewing purely visually based narratives (Trabasso & Nickels, 1992). However, a large body of evidence shows that oral language skills are important for narrative comprehension, as reflected in the Simple View of Reading model (Gough & Tunmer, 1986). This evidence raises an interesting empirical question as to whether these skills account for variance in individual differences in comprehending film.

Another interesting case is comprehending graphic narratives. As we have discussed, this multimedia experience involves both reading and visual processing skills. It makes sense that the visual images could serve as a scaffold for processing the language, but as mentioned above, this is completely contingent on the semantic overlap between the texts and the images. In the panel of examples in Figure 8.2, the text is in the form of dialogue between the characters and conveys additional information not contained in the image (e.g., the fantom has other pressing business), and thus elaborates upon the image and explains why the fantom is leaving Jane.

A critically important difference between graphic narratives and film is that film contains motion, but graphic narratives do not. This leads to greater *attentional synchrony*, in which different viewers look at the same things at the same time when viewing a film as opposed to when viewing a sequence of related still images (e.g., in a graphic narrative; Dorr, Martinetz, Gegenfurtner, & Barth, 2010).[1] Nevertheless, other research has shown that still images that imply motion, which is ubiquitous in comics, lead to greater activation of motion-processing brain regions (e.g., area MT) than do still images that do not imply motion (Osaka, Matsuyoshi, Ikeda, & Osaka, 2010). Thus, even graphic narratives with static images seem to involve some degree of motion processing. To our knowledge there is scant research comparing processing of graphic narratives versus film. However, there is a substantial literature on multimedia learning that could guide systematic study of these media.

We identified three back-end processes that lead to a mental model (i.e., the product of back-end processes). This mental model contains two products, the textbase, which corresponds to the explicitly experienced

[1] This difference in attentional synchrony is also likely related to the fact that in normal film viewing, the filmmaker sets the viewing pace, whereas with comics the viewer or reader controls the pace, likely producing less attentional synchrony.

BRP-MILLER-13-0103-008.indd 85 17/07/13 1:39 PM

content, and the situation model, which includes the inferences that are generated in order to coherently connect that content to world knowledge (e.g., Kintsch, 1998). The mental model is generally believed to be structured around a hierarchically organized sequence of goal episodes (e.g., Trabasso et al., 1989). There is good evidence that viewers of graphic narratives and film produce situation models, for they clearly draw inferences that require integrating information from what they saw and their prior knowledge (e.g., Magliano et al., 1996). However, we know of no clear evidence indicating whether film viewers form a language-like propositional representation of what they saw, which we could call a textbase.

The first of the back-end processes is *event segmentation*, which involves understanding the boundaries between story episodes (e.g., Kurby & Zacks, 2008). As will be discussed below, viewers and readers segment narratives based on changes in such story dimensions as space, time, causality, and the goal episode (e.g., Magliano, Kopp, McNerney, Radvansky, & Zacks, 2012). The next process involves *inferencing*, which primarily involves establishing relationships between explicitly experienced story elements and drawing upon relevant background knowledge (McNamara & Magliano, 2009). Finally, both event segmentation and inferencing are in the service of *structure building*, which constitutes several processes involved in building mental models (laying the foundation, mapping, shifting to new mental structures, inference generation; Gernsbacher, 1990).

PSYCHOLOGICAL RESEARCH COMPARING COMPREHENSION ACROSS MEDIA

A number of scholars who study discourse comprehension have argued that models of discourse comprehension can be viewed as general models that extend to other media and even naturalistic events (Gernsbacher, 1990; Kintsch, 1998; Magliano et al., 2007). However, there are very few empirical investigations of that claim, in particular when controlling for content (e.g., Baggett, 1979; Magliano et al., in press). There have been a number of studies that have assessed the extent to which comprehension skills and proficiencies share variance (e.g., Gernsbacher et al., 1990; Kendeou, Bohn-Gettler, White, & van den Broek, 2008; Pezdek, Lehrer, and Simon, 1984). In this section, we discuss both of these lines of research.

Baggett (1979) conducted a seminal study assessing whether mental models for text-based and visually based narratives were similar. Across several experiments, college student participants were asked to identify the goal episodes that composed the film *The Red Balloon* (Lamorisse, 1956). However, they either read a text-based version of the story, viewed the film-based version of it, or looked at still picture frames from it. Participants identified the same goal episodes regardless of the medium. In addition, in one experiment they were asked to recall the narratives; the content of the recall protocols was remarkably similar in content and structure. This

study provided fairly convincing evidence that back-end processes are surprisingly similar regardless of the modality of the experience.

A more recent study was conducted by Magliano, Kopp, McNerney, Radvansky, and Zacks (2012). The researchers had participants (college students or older adults) view a popular series of children's picture stories that contained only images and no language or read a version for which the researchers wrote text that captured the major events depicted. Participants identified when they perceived boundaries between the events that made up the stories. The researchers found remarkable convergences in the event boundary judgments within and across the two media, albeit convergence was greater within a medium than across the two.

Next consider research that has explored the extent to which comprehension is similar or different across media (i.e., text and film), which has produced conflicting evidence. For example, Pezdek, Lehrer, and Simon (1984) found that there was minimal shared variance in comprehension skill across media. They had children (8–11 years old) listen to (oral text), read (graphic narratives combining language and images), or view (TV) different versions of the same story. The researchers assessed comprehension in a variety of ways but found no significant overall difference between television comprehension and reading comprehension scores. However, they found that comprehension scores across the media were not significantly correlated, suggesting that comprehension proficiencies were not similar, at least for children at this age. On the other hand, Kendeou, Bohn-Gettler, White, and van den Broek (2008) had children (6–8 year olds) listen to, read, or view narratives and assessed comprehension proficiencies. A notable difference from the Pezdek et al. (1984) study was that the children's inferencing skills across the media were consistently, albeit moderately, correlated—demonstrating a stronger convergence across back-end processes than found in the earlier study. The divergence of these two studies clearly indicates that this issue warrants further investigation. Nevertheless, across the five studies considered here (Baggett, 1979; Gernsbacher et al., 1990; Kendeou et al., 2008; Magliano et al., in press; Pezdek et al., 1984), there is converging evidence for both shared processes and proficiencies in narrative comprehension across textual and visual media.

CONCLUDING REMARKS AND PROMISING FUTURE DIRECTIONS

This chapter was necessarily a general overview of an area of theoretical and empirical explorations that shows great promise. We believe that our distinction between front-end and back-end processes provides a heuristic for future research and deeper reflection that has validity in reading, vision, and comprehension sciences. We conclude this chapter by identifying what we believe to be some interesting future research directions. One

caveat is that each research suggestion warrants greater discussion than is afforded here.

One critical line of research that warrants further exploration is the extent to which comprehension processes are functionally similar across media. We provided a framework for this exploration with the introduction of the distinction between front- and back-end processes. Although we have argued that back-end processes likely share larger overlap across media than do front-end processes, it is important to recognize that such front-end processes as eye movements and fixations are heavily guided by back-end processes, such as what one is trying to comprehend at that moment. For example, consistent with results from reading (Anderson & Pichert, 1978), viewers of pictures of home interiors looked at different objects for different amounts of time depending on whether they took the perspective of a home buyer or a burglar, which in turn influenced their memory (Kaakinen, Hyönä, & Viljanen, 2011). Further research on the effects of back-end processes on front-end processes across media is greatly needed.

As a second issue, there is growing evidence that both linguistic and grounded systems (i.e., perceptual and embodied symbols) support comprehension processes (e.g., Fincher-Kiefer, 2001). However, still an open question is the extent to which these different types of representations are involved while perceiving and comprehending narratives presented in different media (e.g., text, graphic novel, or movie). Specifically, are the roles of linguistic and grounded symbols in the representation of linguistic and visually based narratives similar or different?

A final issue that warrants exploration pertains to the topic of this volume. Specifically, what factors affect individual differences in comprehension ability across different media? Do the same factors that predict individual differences in the comprehension of texts also account for individual differences in visually based narratives? The research discussed in this chapter indicates that the answer to this question is currently equivocal. Similarly, do individuals who have disorders known to negatively affect language comprehension, such as dyslexia or autism spectrum disorder, similarly experience troubles comprehending visually based narratives? If not, can visually based narratives be used in the context of interventions as a scaffold to promote language comprehension skills of such individuals?

Although the foregoing list of possibilities is by no means exhaustive, we believe it points to a rich future for research in comprehension across media. By conducting this research, we will gain better insights into the nature of comprehension and the cognitive and brain systems that support it.

REFERENCES

Anderson, R.C., & Pichert, J.W. (1978). Recall of previously unrecallable information following a shift in perspective. *Journal of Verbal Learning and Verbal Behavior, 17*(1), 1–12. doi:10.1016/s0022-5371(78)90485-1

Baggett, P. (1979). Structurally equivalent stories in movie and text and the effect of the medium on recall. *Journal of Verbal Learning & Verbal Behavior, 18*(3), 333–356.

Bordwell, D., & Thompson, K. (2003). *Film art: An introduction*. New York, NY: McGraw-Hill.

Broccoli, A.R.P., & Gilbert, L.D. (Writers). (1979). *Moonraker* [Film]. Available from CBS/Fox Video, Industrial Park Drive, Farmington Hills, MI 48024.

Dorr, M., Martinetz, T., Gegenfurtner, K.R., & Barth, E. (2010). Variability of eye movements when viewing dynamic natural scenes. *Journal of Vision, 10*(10), 1–17.

Eckstein, M.P., Drescher, B.A., & Shimozaki, S.S. (2006). Attentional cues in real scenes, saccadic targeting, and Bayesian priors. *Psychological Science, 17*(11), 973–980.

Fei-Fei, L., Iyer, A., Koch, C., & Perona, P. (2007). What do we perceive in a glance of a real-world scene? *Journal of Vision, 7*(1), 1–29.

Fincher-Kiefer, R. (2001). Perceptual components of situation models. *Memory & Cognition, 29*, 336–343.

Gernsbacher, M.A. (1990). *Language comprehension as structure building* (Vol. 11). Hillsdale, NJ: Erlbaum.

Gernsbacher, M.A., Varner, K.R., & Faust, M.E. (1990). Investigating differences in general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*(3), 430–445.

Gough, P., & Tunmer, W. (1986). Decoding, reading, and reading disability. *Remedial and Special Education, 7*, 6–10.

Graesser, A.C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review, 101*, 371–395.

Gustavson, P. [Artist]. (1939). *The fantom of the fair* [Public domain comic strip]. Available from http://goldenagecomics.co.uk

Kaakinen, J.K., Hyönä, J., & Viljanen, M. (2011). Influence of a psychological perspective on scene viewing and memory for scenes. *Quarterly Journal of Experimental Psychology, 64*(7), 1372–1387. doi:10.1080/17470218.2010.548872

Kendeou, P., Bohn-Gettler, C., White, M.J., & van den Broek, P. (2008). Children's inference generation across different media. *Journal of Research in Reading, 31*(3), 259–272. doi:10.1111/j.1467-9817.2008.00370.x

Kintsch, W. (1998). Comprehension: A paradigm for cognition. New York, NY: Cambridge University Press.

Kurby, C.A., & Zacks, J.M. (2008). Segmentation in the perception and memory of events. *Trends in Cognitive Sciences, 12*, 72–79.

Lamorisse, A. (Writer). (1956). *The red balloon* [Film]. Paris, France: Films Montsouris.

Larson, A.M., Hendry, J., & Loschky, L.C. (2012, May). *Scene gist meets event perception: The Time course of scene gist and event recognition*. Poster presented at the 12th annual meeting of the Vision Sciences Society, Naples, FL.

Loschky, L.C., & Larson, A.M. (2010). The natural/man-made distinction is made prior to basic-level distinctions in scene gist processing. *Visual Cognition, 18*(4), 513–536.

Loschky, L.C., Magliano, J.P., & Larson, A.M. (2012). The need for an integrated theory of film perception and comprehension. Paper presented at the Workshop on Intelligent Cinematography and Editing 2012 workshop, Raleigh, NC.

Magliano, J.P., Dijkstra, K., & Zwaan, R. (1996). Generating predictive inferences while viewing a movie. *Discourse Processes, 22*, 199–224.

Magliano, J.P., Kopp, K., McNerney, M.W., Radvansky, G.A., & Zacks, J.M. (2012). Aging and perceived event structure as a function of modality. *Aging, Neuropsychology, and Cognition, 19*, 264–282.

Magliano, J.P., Radvansky, G.A., & Copeland, D.E. (2007). Beyond language comprehension: Situation models as a form or autobiographical memory. In F. Schmalhofer & C. Perfetti (Eds.), *Higher level language processes in the brain: Inference and comprehension processes* (pp. 379–391). Mahwah, NJ: Erlbaum.

Magliano, J.P., Taylor, H.A., & Kim, H.J. (2005). When goals collide: Monitoring the goals of multiple characters. *Memory & Cognition, 33*, 1357–1367.

Maljkovic, V., & Martini, P. (2005). Short-term memory for scenes with affective content. *Journal of Vision, 5*(3), 215–229.

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

BRP-MILLER-13-0103-008.indd 89 17/07/13 1:39 PM

McClure, E.B. (2000). A meta-analytic review of sex differences in facial expression processing and their development in infants, children, and adolescents. *Psychological Bulletin*, *126*(3), 424–453.

McNamara, D.S., & Magliano, J.P. (2009). Towards a comprehensive model of comprehension. In B. Ross (Ed.), *The psychology of learning and motivation* (Vol. 51, pp. 297–384). New York, NY: Elsevier Science.

Osaka, N., Matsuyoshi, D., Ikeda, T., & Osaka, M. (2010). Implied motion because of instability in Hokusai Manga activates the human motion-sensitive extrastriate visual cortex: An fMRI study of the impact of visual art. *NeuroReport: For Rapid Communication of Neuroscience Research*, *21*(4), 264–267. doi:10.1097/WNR.0b013e328335b371

Pezdek, K., Lehrer, A., & Simon, S. (1984). The relationship between reading and cognitive processing of television and radio. *Child Development*, *55*, 2072–2082.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372–422.

Rayner, K., Foorman, B.R., Perfetti, C.A., Pesetsky, D., & Seidenberg, M.S. (2001). How psychological science informs the teaching of reading. *Psychological Science in the Public Interest*, *2*(2), 31–74.

Trabasso, T., & Nickels, M. (1992). The development of goal plans of action in the narration of picture stories. *Discourse Processes*, *15*, 249–275.

Trabasso, T., van den Broek, P., & Suh, S. (1989). Logical necessity and transitivity of causal relations in stories. *Discourse Processes*, *12*, 1–25.

Zwaan, R.A. (2004). The immersed experiencer: Toward an embodied theory of language comprehension. In B.H. Ross (Ed.), *The psychology of learning and motivation* (pp. 35–62). New York, NY: Academic Press.

Zwaan, R.A., Langston, M.C., & Graesser, A.C. (1995). The construction of situation models in narrative comprehension: An event-indexing model. *Psychological Science*, *6*, 292–297.

BRP-MILLER-13-0103-008.indd 90 17/07/13 1:39 PM