



Picture This: A Review of Research Relating to Narrative Processing by Moving Image Versus Language

Elspeth Jajdelska^{1*}, Miranda Anderson², Christopher Butler³, Nigel Fabb¹, Elizabeth Finnigan⁴, Ian Garwood⁵, Stephen Kelly⁶, Wendy Kirk⁷, Karin Kukkonen⁸, Sinead Mullally⁹ and Stephan Schwan¹⁰

¹English, University of Strathclyde, Glasgow, United Kingdom, ²Philosophy, University of Stirling, Stirling, United Kingdom, ³Department of Clinical Neuroscience, University of Oxford, Oxford, United Kingdom, ⁴English, Southern Regional College of Northern Ireland, Armagh, United Kingdom, ⁵Film and Television Studies, University of Glasgow, Glasgow, United Kingdom, ⁶Psychology, University of Strathclyde, Glasgow, United Kingdom, ⁷Glasgow Women's Library, Glasgow, United Kingdom, ⁸Comparative Literature, University of Oslo, Oslo, Norway, ⁹Neuropsychology, Newcastle University, Newcastle upon Tyne, United Kingdom, ¹⁰Psychology, Leibniz-Institut für Wissensmedien, Tübingen, Germany

OPEN ACCESS

Edited by:

Ramesh Kumar Mishra,
University of Hyderabad, India

Reviewed by:

Yang Jiang,
University of Kentucky College of
Medicine, United States
Angelique Christine Paulk,
Harvard Medical School,
United States

*Correspondence:

Elspeth Jajdelska
elspeth.jajdelska@strath.ac.uk

Specialty section:

This article was submitted to
Cognition,
a section of the journal
Frontiers in Psychology

Received: 02 January 2019

Accepted: 02 May 2019

Published: 26 June 2019

Citation:

Jajdelska E, Anderson M, Butler C,
Fabb N, Finnigan E, Garwood I,
Kelly S, Kirk W, Kukkonen K,
Mullally S and Schwan S (2019)
Picture This: A Review of Research
Relating to Narrative Processing by
Moving Image Versus Language.
Front. Psychol. 10:1161.
doi: 10.3389/fpsyg.2019.01161

Reading fiction for pleasure is robustly correlated with improved cognitive attainment and other benefits. It is also in decline among young people in developed nations, in part because of competition from moving image fiction. We review existing research on the differences between reading or hearing verbal fiction and watching moving image fiction, as well as looking more broadly at research on image or text interactions and visual versus verbal processing. We conclude that verbal narrative generates more diverse responses than moving image narrative. We note that reading and viewing narrative are different tasks, with different cognitive loads. Viewing moving image narrative mostly involves visual processing with some working memory engagement, whereas reading narrative involves verbal processing, visual imagery, and personal memory (Xu et al., 2005). Attempts to compare the two by creating equivalent stimuli and task demands face a number of challenges. We discuss the difficulties of such comparative approaches. We then investigate the possibility of identifying lower level processing mechanisms that might distinguish cognition of the two media and propose internal scene construction and working memory as foci for future research. Although many of the sources we draw on concentrate on English-speaking participants in European or North American settings, we also cover material relating to speakers of Dutch, German, Hebrew, and Japanese in their respective countries, and studies of a remote Turkish mountain community.

Keywords: narrative, media, reading, film, fiction, comprehension, literature, cognitive humanities

INTRODUCTION

Reading, and reading fiction in particular, for enjoyment has been positively correlated with young people's attainment in a wide range of studies across different countries, many, such as the OECD's PISA studies, involving large cohorts (Cunningham and Stanovich, 1998; PIRL, 2001, 2006; Organisation for Economic Co-operation and Development (OECD), 2002, 2010;

Verghese et al., 2003; Mar et al., 2009; Hughes et al., 2010; Sullivan and Brown, 2013; McGeown et al., 2015; Ritchie et al., 2015; Sikora et al., 2019). Moving image narrative is experienced as easier than written narrative both to process and access (Salomon, 1979, 1984; Salomon and Leigh, 1984; Beentjes and van der Voort, 1988, p. 393–4; Ennemoser and Schneider, 2007; Jensen et al., 2016). But there is no evidence that watching fiction films for enjoyment confers comparable benefits in attainment. Since the advent of television, and then laptops, tablets, and smart phones, reading for enjoyment has faced ever more competition from the moving image (Merga, 2016, 2017; Merga and Roni, 2017). This is causing concern among policy makers and educationalists, among others (Scottish Government, 2007; Scholastic, 2015).

It is not clear exactly why reading narrative for enjoyment is associated with attainment. An analysis of data on identical twins by Ritchie et al. suggests that an association between reading and either genetic traits or socio-economic background cannot explain all of the benefits of reading for pleasure (Ritchie et al., 2015; see also Organisation for Economic Co-operation and Development (OECD), 2002). Nor is it clear whether, for example, watching highly crafted moving image fiction for enjoyment could have comparable benefits. A clearer understanding of the underlying mechanism of the effect could help to make the case for encouraging young people to enjoy fiction and guide policy. In this review, we first examine evidence on the similarities and differences between narrative processing across these media. There are substantial structural and experiential overlaps between narratives in media. But our review suggests that moving image narrative produces superior performance in the recall of sensory detail and information; readers have more divergent experiences than viewers; and variation between individuals affects the experience of verbal narrative more than that of moving image narrative. However, findings are hard to compare. Researchers use varied methods in the attempt to produce comparable stimuli in the two media, and the body of research is often fragmented and inconclusive.

In the second part, we explore approaches for future research. We suggest a change of focus from attempting to create equivalent stimuli. Instead, it may be fruitful to compare how early visual and verbal processing interacts with global and temporally extended experiences of narrative. In particular, we consider internal scene construction, supported by the episodic memory system, and the potential for differential engagements of verbal and visual working memory in different media. These mechanisms may be implemented in the ways that both lower and higher level causal patterns in narrative are constructed, through flexibility of concept instantiation. Fiction reading may draw more extensively on personal memory than watching moving image narrative and may involve more complex causal inferences.

Throughout the review, we use the term “story” for a minimal narrative structure underlying a narrative experience. We use “narrative” for stories rendered in words and/or images; stimuli extended in time and depicting at least one agent-like entity initiating and/or experiencing change in their environment (Gennette, 1980; Fludernik, 2008). “Moving image” refers to all kinds of screen-based moving image stimuli, from those

seen in cinemas to those watched on phones. This term covers both animations and live action films. Most moving image narratives today include speech and other sounds. We follow many of the researchers discussed below and classify these as “moving image” but explore the question of hybridity in movies, the fact that they often combine speech, music, and either live action or animated images in a single narrative, where possible. “Verbal” refers to any language-based stimuli, spoken or written (or potentially signed, though we have no examples of signed narratives). “Text” refers to verbal stimuli organized at the level of discourse (rather than, for example, letter, word, or single clause), whether spoken or written. We use “transportation” to describe a mode of experiencing narration, more or less independent of medium. This is distinct from “immersion” or “presence,” terms used to describe the experience of a medium, more or less independent of narration (Busselle and Bilandzic, 2008).

SIMILARITIES IN NARRATIVE PROCESSING BETWEEN TEXT AND MOVING IMAGE

Our understanding of similarities is informed by a model of narrative processing developed in the 1980s and 1990s, initially for text. This model combines “situation modeling” of a narrative with “event segmentation.” Situation modeling emerged from a range of earlier work focused on narrative processing as an information-focused task relating to real-life schemas (Meyer, 1975; Rumelhart, 1975, 1977; Mandler and Johnson, 1977; Thorndyke, 1977; Poulsen et al., 1979; Stein and Glenn, 1979; Lichtenstein and Brewer, 1980). Event indexing also emerged from work on real-life processing (Newton, 1973; Zacks and Tversky, 2001). Using this model, a range of work by Zacks and colleagues suggests that in real life, verbal narrative and moving image narrative, we attend to the same dimensions (characters’ goals, their relationships with objects, and their location) using the same mechanisms of heightened attention to event boundaries at a range of timescales (Speer et al. 2007a; Zacks et al., 2009; Zacks, 2015, p. 29).

Situation modeling also provides evidence for distinctions made by narratologists, anthropologists, and folklorists between a minimal structure that can be reused (story) and the experience of a particular iteration of such a structure (narrative) (Aarne and Thompson, 1961; Propp, 1968; Genette, 1980; Shklovsky, 2011, p. 25, 31). The minimal structure consists of a sequence of events and relates to modeling and updating the narrative across dimensions, such as changes in characters’ goals (Zwaan and Radvansky, 1998). The model generates bounded episodes in a sequence through event perception which segments time into hierarchically organized events, by directing attention to points when several dimensions change at once.

There has been less empirical work on the experiential aspects of narrative, though here too there is evidence of a shared narrative processing across media. Investigating the potential for narrative to improve abilities to empathise, Kidd and Castano (2013), for example, found that literary fiction, as distinct from

popular genre fiction, produced improved performance in theory of mind tests, while Black and Barnes (2015) found the same for viewing award-winning, as opposed to less critically acclaimed, television drama (see also Mar et al., 2009; Koopman and Hakemulder, 2015). Hakemulder similarly finds that both films and texts which carefully craft form, the way the story is told, produce a richer experience of meaning (Hakemulder, 2007). Both media are capable of supporting transportation, a narrative state involving “imagery, emotional response and attentional focus” (Green and Brock, 2000; Green et al., 2008; Bal and Veltkamp, 2013).

It is not surprising then that theoretical and empirical work in both psychology and humanities reveals substantial overlaps in the structures and affordances of the two media, with both requiring readers, hearers, and viewers to learn or understand ways of representing material. Magliano et al. show how “Mapping processes” between shots, for example, are akin to the “bridging inferences we know that readers generate” in reading, as demonstrated by psycholinguistic research (Magliano et al., 2013, p. 79). Shot sequences of faces and postures can enable inferences about characters’ inner states, which authors supply through, for example, narrative voice or free indirect speech, and through direct information about inner states or indirect information about facial expressions and stance (Magliano et al., 2013, p. 80).

Aesthetic approaches to form also suggest continuity between narrative media. Hakemulder defines “literariness” as deviation from the norms of representation and finds that it can enhance both film and text narratives (Hakemulder, 2004, 2007). Simpson suggests cinematic analogues for stylistic textual features that create a sense of urgency (Simpson, 2014). Kraft compares manipulations of camera angle to variation in verbal narrators’ points of view (Kraft, 1987). Forceville makes a systematic comparison between Ian McEwan’s novel *The Comfort of Strangers* and Harold Pinter’s film adaptation to show how visual metaphors, cross cuts, and ambiguous “point of view shots” can replicate effects generated by McEwan using verbal style (McEwan, 1981; Pinter, 1990; Forceville, 2002). Lang et al. suggest that audiovisual narratives can share some of the structure of oral conversational narratives (Lang et al., 1995). Work in the humanities on “transmedia” narrative theory has also identified both affordances specific to particular media, and ways in which these can create formal equivalence (Ryan, 2004; Walsh, 2006; Kukkonen, 2011, 2013).

But both media seem to depend on similar kinds of medium-specific learning, rather than just carrying over generic knowledge into an understanding of what is being represented. For example, Schwan and Ildirar worked with adult first-time viewers of television to explore what formal processes they needed to comprehend moving image narrative. These viewers found some aspects of continuity across shots straightforward, but not all. For example, they interpreted changes in camera position across shots as changes in the position of the person filmed, and had to learn to see film as a crafted medium creating “a coherent whole” (Schwan and Ildirar, 2010; Ildirar and Schwan, 2015). Viewers of dynamic scenes similarly learn to combine different views which may maximize relevant information rather than being organized around real-life

experience (Friedman and Waller, 2008; Garsoffky et al., 2009; Huff and Schwan, 2012). Film viewers must also learn to identify event boundaries which are missing from shots or cross cuts (Schwan et al., 2000; Schwan and Garsoffky, 2004; Shimamura et al., 2014, 2015; Smith and Mital, 2015). In the same way, learning to understand spoken language and/or decode print is not sufficient to understand verbal narrative. Written discourse comprehension involves an ability to conceptualize a narrator who, unlike a face to face interlocutor, has only a notional location in time and space, and to correctly interpret deixis (terms relating to speaker location, such as “this” versus “that”) in relation to that notional location (Jajdelska, 2007). Without these skills, even children who are skilled at decoding written words and sentences (converting them into speech) cannot always draw appropriate inferences from narrative text (Yuill and Oakhill, 1991).

Both media then share structures of comprehension through situation modeling and the marking of event boundaries. Both can support transportation and enhanced empathy. Both require their audience to appreciate their status as artifacts, and both can vary the ways in which information is conveyed for comparable, if not equivalent, esthetic effects.

DIFFERENCES IN NARRATIVE PROCESSING BY TEXT AND MOVING IMAGE

Baggett compared film and text narrative processing in a series of experiments (Baggett, 1975, 1979, 1984; Baggett and Ehrenfeucht, 1982) and characterized by careful and painstaking efforts to create equivalent stimuli across the two media. These included using a popular children’s film with no dialogue (“The Red Balloon”) to generate a text narrative which subjects judged had exactly the same episodes (1979). These included taking a popular children’s film which had no dialogue (“The Red Balloon”), creating a text version of the narrative, and then inviting raters to confirm that the text and the film had the same number of episodes. The approach to generating materials drew on theories of narrative structure (Kintsch and Van Dyke, 1975) comparable to the later situation model and event boundary approach. Baggett and Ehrenfeucht later looked at content rather than episodic structure, using 180 stills and 20 free recall protocols from viewers to rewrite the written source for a 1953 animated film. They used this to create an audio text of exactly the same length as the film and using the same dialogue (Baggett and Ehrenfeucht, 1982).

Baggett’s findings on structure support the work on event indexing and situation modeling discussed above; both readers and viewers segmented the story in similar ways. In cued recall after 7 days, however, readers were more likely to draw on world knowledge than viewers, and their recall of correct information was worse, sometimes because they substituted real-world knowledge. They recalled fewer precise and vivid details. The authors also reported, though this was not measured systematically, that readers showed little emotion, whereas at

least two film viewers cried and some film groups applauded (1979). As the authors point out, this potential difference in emotion may have had an impact on memory. However, this need not indicate a general distinction between media: texts have the potential to generate emotion too. In this respect, the experiment illustrates rather how equivalence on one dimension (structure or content) inevitably interferes with other dimensions (doctoring the text to match structure or content can hamper the stylistic freedom that might make it more emotional). It is not clear then how far Baggett's findings relate to differences in psychological effect of media in general rather than differences between, for example, the comparative emotional power of these specific stimuli.

A second body of work contrasting moving image and verbal narrative arose from growing concerns in the 1970s about the rise of television and its power to reduce time spent by children reading (Neuman, 1995). Findings here are hard to compare because of varied uses of stimuli across moving images, audio, text, and still images, alone or in a range of combinations, and usually by young children, since researchers are usually interested in issues relating to children's competence in literacy (Singer and Singer, 1981, 2005; Beagles-Roos and Gat, 1983; Pezdek and Stevens, 1984; Gibbons et al., 1986; Greenfield et al., 1986; Pezdek et al., 1987; Beentjes and van der Voort, 1988, 1991). With these caveats in place, some findings are consistent with Baggett's. Beagles-Roos and Gat, for example, found that (child) readers are more likely to use real-world knowledge to form narrative inferences than viewers; similarly, Baggett found that (adult) readers were more likely to have real-world knowledge intrude into recall protocols (Beagles-Roos and Gat, 1983). Gibbons et al., as well as Beagles-Roos and Gat, found, like Baggett, that viewers had a better memory for content in general and detail in particular than readers (Gibbons et al., 1986). Levorato (1991), on the other hand, found that children who watched a video of a familiar line of action (getting ready for bed, laying the table) recalled minor actions in the sequence less well than those who read a verbal description and that the readers produced more linguistically complex reports than the video group.

Existing research also suggests that differences in processing verbal narrative may vary between individuals and groups, in ways that are likely to affect a comparison with a moving image narrative. Mental imagery abilities, for example, vary in vividness when measured by questionnaires such as the Vividness of Visual Imagery Questionnaire (Marks, 1973; Zwaan, 2009; for a critique of the questionnaire, see Troscianko, 2013). Brain imaging and clinical research also point to high levels of individual variation in imagery, with the size of area V1 both "predicting the sensory strength and precision of visual imagery" and "likely to vary enormously across individuals" (Pearson et al., 2015, p. 594). Denis (1982) found that high imagers have better recall for narrative and descriptive text than low imagers, an advantage which disappears for abstract non-imageable texts. Center et al. (1999) found that imagery training improved comprehension and recall for written narratives in children. Weibel et al. (2011) found that imagery abilities are not related to either the presence or enjoyment uniformly across the media

of text, film, and computer game. Individuals with high imagery abilities experienced more presence and enjoyment in text, while those with low imagery abilities reported "marginally higher enjoyment ratings" in film, but not higher presence. Individuals with high imagery ability, then, may find text easier and more rewarding than those with low imagery, a factor which will affect their differential response to medium. At least one set of findings, however, suggests limits on the benefits of high imagery for texts. Readers can adjust to some counterfactual norms, such as talking pigs, very rapidly (Foy and Gerrig, 2014), and "minimally counterintuitive concepts," where a natural law is broken (for example, a talking bush), are found routinely in fairy tales (Thompson, 1955–1958; Norenzayan et al., 2006). It is not clear whether the medium contributes to this easy adjustment, but Slone et al. (2007) find that while high imagery (for example, "apple" versus "justice") interacts with memory for both intuitive and maximally counterintuitive concepts, this was not the case for minimally counterintuitive concepts. "Unnatural narratologists," for example, argue that genre expectations create new norms, so that minimally counterintuitive concepts are not just accepted, but seen as natural for a reader/hearer of fairy tales (though see also Alber, 2016; Anderson and Iverson, 2018; Jajdelska, 2019).

As well as differences between individuals in imagery, there is evidence of differences between individuals in speeds of visual and verbal processing. These can be detected in infancy: verbal and spatial working memory capacities have high heritability. These individual differences interact with affect in ways that lead to "preferred brain pathways to process visual and cognitive information" (Parasuraman and Jiang, 2012, p. 77). Traits such as "high need for cognition" may also differentiate the text and movie experiences of individuals. Allbritton and Gerrig suggest that reading narratives will be more transportive than viewing them (Allbritton and Gerrig, 1991). But Green et al. (2008) found that while imagery was not related to transportation, high need for cognition individuals found text narrative more transportive, while the converse was true for low need for cognition individuals. The personality trait of openness has also been associated with a preference for print fiction, again suggesting a difference between the ease and pleasure with which individuals might process verbal and film narrative (Mar et al., 2009). Traits such as anxiety can affect readers' predictive inferences in relation to threat. More anxious readers are more likely than less anxious ones to make predictive inferences in response to threat-based content (Calvo and Castillo, 2001). It is not clear if these individual variations are equally relevant to the experience of moving image narrative. Finally, variation in verbal competence (written and spoken) plays an obvious role in differentiating between reader experiences compared to those of viewers (Van der Molen, 2000).

Age too can affect narrative processing by medium (Hayes and Birnbaum, 1980; Baggett and Ehrenfeucht, 1982; Bohn-Gettler et al., 2011). There is a marked difference between children of around 4 years old and those of 8 or 10 years, with younger children in particular performing better on movies than text on recall and comprehension (Reifel, 1984; Gibbons et al., 1986; Beentjes and van der Voort, 1991; Bordeaux and Lange, 1991;

Kendeou et al., 2008). Younger children (about 4 years old) tend to infer goals from explicit physical actions more than other sources of information. This may explain why they recall audiovisual narratives better than audio ones, whereas the older children and adults in the same studies, conducted over generations with different levels of exposure to visual media, tended to perform equally well in both media (Hayes and Birnbaum, 1980; Baggett and Ehrenfeucht, 1982; Reifel, 1984; Gibbons et al., 1986; Beentjes and van der Voort, 1991; Kendeou et al., 2008). As children get older, they get better at narrative processing in both media (Pezdek et al., 1987). But although, as we saw above, both verbal and moving image narrative comprehension require skill and practice, verbal narrative comprehension in particular seems to benefit from increased practice in childhood (Beentjes and van der Voort, 1988; McGeown et al., 2015; Richie et al., 2015; Jensen et al., 2016). Differences between children of the same age in processing narrative relative to medium may reflect different levels of practice as well as differences such as need for cognition or imagery ability.

As well as these differences between individual experiences of the two media, Magliano et al. (2013) propose structural differences. “Shots,” they argue, can be compared to “sentences” as “minimum units of production.” However, it can be argued that shots can just as plausibly be compared to words. And there are problems with either analogy. In McQueen and Walsh (2008), for example, a single 17-and-a-half-minute shot covers an unbroken dialogue between two characters (Maher, 2008). In ecological settings, film narrative often requires visual, verbal, and musical processing simultaneously. Silent readers control the pace of the narrative; solitary viewers can rewind or fast forward, but cannot otherwise control pace, and film watching is often social (Magliano et al., 2013, p. 80–82). The power of film to control visual attention through shot length, motion, luminescence, and other medium specific affordances has been described as “the tyranny of film” (Hasson et al., 2008; Cutting et al., 2011; Loschky et al., 2015). This characterization has some support from evidence of synchronization of eye blinks and saccades between viewers of film (Nakano et al., 2009; Shimamura et al., 2014; Loschky et al., 2015). However, this degree of synchronization between viewers may not apply to longer-term processes, such as identifying characters’ goals (Kauppi et al., 2010). There may also be a bias here to mainstream commercial film and a potentially reductive model of passive viewer and active filmmaker. A comparison between *auteur* film and *commercial* fiction books, or silent film and highly formulaic genres, might instead highlight text’s power to control attention through, for example, “rhetorical focusing,” and film makers’ power to make viewers work hard through choice of cuts or camera angle (Schwan and Garsoffky, 2004; Sanford and Emmott, 2012).

SUMMARY: SIMILARITIES AND DIFFERENCES

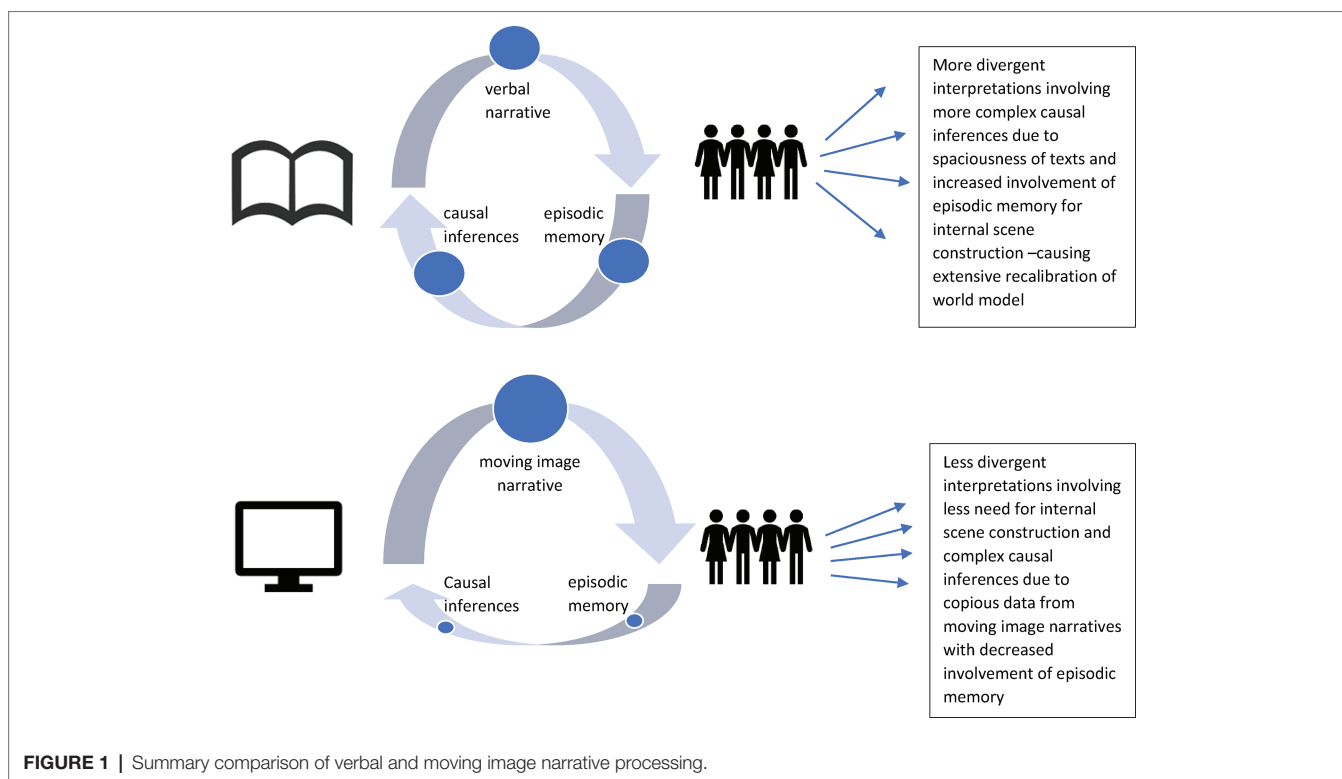
There are considerable overlaps in the affordances of different kinds of narrative media, and these are reflected in processing. Viewers and readers alike can be constrained by the attention

hierarchies of real life when they segment narratives into events using situation modeling. These sequences of events allow readers and viewers to retell the story as a new narrative in the original or a different medium, explaining, for example, why novels can be turned into recognizable films. Both media require viewers or readers to see the narrative as an artifact with the hand of a human maker behind it in order to comprehend it, and this can involve learning conventions of form. Each also has a body of formal techniques that can exploit their different affordances, varying *how* the story is told. At least some of these have comparators in the other medium, allowing for some equivalent or near equivalent esthetic effects. They can both produce empathy, transportation, vividness, and emotion.

Differences between the two processing experiences may be manifest in the greater variety among readers’ memories of the story, as they draw more heavily on their own experience of the world to form inferences (see **Figure 1**). In film groups, for example, the role of negotiating the self when reconstructing a story can be made open, allowing for high mutual influence in self-presentation to the group (Edwards and Middleton, 1986). In book groups, the process of jointly remembering and interpreting the text may involve both synchronization of mood and speech and divergence in interpretation and recall of the text (Steenberg et al., 2014). Readers’ memories may be not only less accurate than those of film viewers but also more creative in the sense of introducing new elements. Readers may be less likely than viewers to recall, or believe that they recall, vivid (in the sense of sensory) details. Their retellings may themselves have narrative form; recall of movie stories may be closer to reportage. Readers may also diverge more in retelling because of individual differences between them, including differences between infants, older children, and adults; traits like openness and need for cognition; and aptitudes in mental imagery. It should be noted, however, that all of these distinctions may be affected by both experimental and ecological task demands. A student preparing to answer questions on a narrative in a language exam may remember different details from someone reading, hearing, or viewing for leisure, for example. We also touch on the difficulties of creating equivalent task demands in relation to different media toward the end of this review.

NON-NARRATIVE VISUAL AND VERBAL PROCESSING

The body of work directly comparing film and text narratives is small, but work on other aspects of verbal and visual processing may also be relevant. Some aspects of visual processing may be “cost free” in comparison to their verbal equivalents (Lang et al., 1995 citing: Salomon, 1984; Salomon and Leigh, 1984; Graber, 1990; Grimes, 1991; Rolandelli et al., 1991; Basil, 1995). For example, visual perception of physical causation can be an automatic process, involving fine grained spatial and temporal congruity that verbal description cannot reproduce (Fugelsang et al., 2005). Similarly, visual processing of human action can engage automatic simulation at fine grained levels (Brass et al., 2000).



Moving images therefore might be expected to have an advantage over text in teaching new tasks. The evidence here is hard to interpret. Studies of learning to use a new computer application (Palmiter et al., 1999) or an asthma inhaler (Wilson et al., 2010) indicate an advantage for video in short-term recall (Wilson et al., 2010) and task performance (Palmiter et al., 1991). But this advantage may be limited to participants with lower literacy skills (Wilson et al., 2010), or reversed in longer-term task performance, where viewers were slower and less accurate in a related task (Palmiter et al.). Lower competence in literacy is also associated with the advantage of moving image news plus words over still image news in Van der Molen (2000), where children perform better with television and adults with print. Moving image information may involve greater engagement, sympathy, and perceived realism but not necessarily improved processing or recall (Yadav et al., 2011).

Another example of automaticity in visual processing relates to “theory of mind” or “mirroring” interpretations of agents’ goals (perception of actions *via* the mirror neuron system). This can be contrasted with “mentalizing” systems used for the same purpose (Overwalle and Baetens, 2009). The two have different implications for readers’/viewers’ narrative emotions and interpretation. Identification of outgroups, for example, may recruit mirror neuron system responses alone and even categorize these agents or characters as “infra-human” objects (Harris and Fiske, 2006). Verbal descriptions of actions can also engage the mirror neuron system, but much depends on exactly how the passage might be written (Gallagher et al., 2000; Hauk et al., 2004; Jajdelska et al., 2010). Again, careful attention to stimuli is needed.

A long line of research on the “picture superiority effect” (Paivio, 1971) investigates whether pictures of objects have greater effects than single-word depictions of those objects (any effects may extend to auditory representations of the objects; Crutcher and Beer, 2011). Insofar as there is a picture superiority effect, it need not be because of Paivio’s “dual coding” account of processing (Hockley, 2008). However, it is not clear that pictures really do have a general advantage over words in relation to either recall (Bartlett et al., 1980; Gati and Tversky, 1987; Koehler et al., 2005; Reinwein, 2012) or emotion (Otgaar et al., 2010; Schlochtermeyer et al., 2013).

The “verbal overshadowing effect” is associated with evidence that describing a face after seeing it can make it harder to recognize that face again (Klatzky et al., 1982; Schooler and Engstler-Schooler, 1990). One possible explanation for this is that face perception is holistic, whereas verbal descriptions tend to be feature by feature (Dodson et al., 1997). However, when verbal descriptions are holistic and less feature-based, they may sometimes improve recognition (Brown and Lloyd-Jones, 2006). In the case of dynamic scenes, rather than faces, Huff and Schwan (2008) find that verbalization (in this case in the form of reading a description rather than generating one) can involve both facilitation and overshadowing. Jajdelska et al. (2010) suggest that differences in the way a text describes a face may significantly improve memory for the described face.

An automaticity advantage of moving image may be plausible in some cases, such as direct perception of causation and action. In case of the picture superiority effects, and verbal overshadowing effects, much may depend on the precise words

and images used. Again, this suggests that great care is needed in choice of stimuli when attempting to compare narrative media.

INTERACTIONS BETWEEN TEXT AND IMAGE

This is an area that has been researched by educational psychologists. Mayer's "multimedia principle," for example, states that "People learn more deeply from words and pictures than from words alone" (Mayer, 2014, p. 43). Again, research on illustrated texts suggests a complex range of interacting effects between visual and verbal processing, often depending on the precise words and images used. In a review of research on how the sequencing of picture and texts relates to learning, for example, Eitel and Scheiter conclude that "the relative complexity of the information conveyed by the picture and by the text should determine which medium is better to be processed first" (Eitel and Scheiter, 2015). In some books, the illustrations and texts jointly supply the information needed, rather than one supporting the other (Corrigan and Surber, 2010). Elsewhere they aid comprehension in relation to some kinds of text, but not others (Moore and Skinner, 1985; Beveridge and Griffiths, 1987; Trabasso and Nickels, 1992). As with the picture superiority and verbal overshadowing effects, small variations in the fine detail of stimuli can have powerful impacts (Schlochtermeyer et al., 2013). Magliano et al. (2012), for example, speculate that an illustration of a leaping frog in a children's book will be more vivid than the accompanying text but do not explore the potential effects of a different text using, for example, kinesic imagery (Jajdelska et al., 2010).

Very few films in recent decades do not use language (through dialogue and sometimes onscreen text and voiceover) and music as well as images. Magliano et al. found that viewers of a popular action film relied on visual sources, music, and dialogue in that order to make predictive inferences (Magliano et al., 1996; Hoeckner et al., 2011; Strick et al., 2015). However, as the discussion of cinematic form earlier suggests, a different film genre might produce different interactions between these three elements. Work on graphic novels and comic books shows that formal features of image and text, such as viewpoint, can conflict as well as harmonize, supplying these media with a particularly rich and challenging range of affordances (Kukkonen, 2011, 2013; Sabeti, 2012). As with research on visual versus verbal processing, research on mixed media suggests that great care is needed if comparing narrative stimuli by medium.

ISSUES WITH METHODS USED IN EXISTING RESEARCH

Verbal processing, like mental imagery, is increasingly understood as engaging the resources of higher level visual processing (Barsalou et al., 2003; Hauk et al., 2004; Kosslyn et al., 2005; Reddy et al., 2010; Hauk and Tschentscher, 2013; Pulvermüller, 2013). Yet, as we have seen, establishing equivalent stimuli across words and images may be an impossible task. Episodic structure and duration, for example, cannot be made identical without

varying style, imagery, and form. Creating equivalent content again may be impossible; just one still from a movie can contain more visual information than the lengthiest description could fully capture. And although film and text share some formal affordances, these cannot often, if at all, be equated in fine detail. Again, it is problematic to vary dimensions one at a time, since each dimension potentially affects the others. The emotional effects of a story, for example, can potentially relate to interactions between all of these dimensions, and emotion cannot necessarily be untangled from perception and comprehension (Calvo and Castillo, 2001; Barrett and Bar, 2009).

Task demands in relation to the two media also present a problem. Verbal narrative is embedded in social life and performance, from everyday conversational narrative to high culture. Sociolinguists, for example, have shown how conversational narrative acts as a means to negotiate self-identity in relation to group identity, while anthropologists have discussed the role of performance in narrative production and audience reception (Labov and Waletzky, 1967; Bauman, 1986; Barber, 2007). On the other hand, few, if any, of us are "everyday storytellers" in film, despite the rapid increase of filmed material on smartphones. Baggett noted that her reading participants retold stories using "Once upon a time" and reproducing past tenses from the original texts, while the viewers reported what happens "when the movie begins" in the present tense (1979). Readers may experience "an illusion of truth" when reading fiction as they have a default position of trusting narrators, a trust harder to replicate in films even if they use a voiceover narrator (Gilbert, 1991; Prentice and Gerrig, 1999; Marsh et al., 2003), although some forms of fiction rely on readers' distrust of "unreliable narrators." The processing of verbal narrative, then, even in written form, is embedded in processes of social interaction, and readers/hearers also have a lifetime of experience as narrators of conversational narratives themselves. Recent directions in predictive language processing also emphasize the closeness of production and comprehension, as readers/hearers simulate the speech of interlocutors in order to time their own contributions to conversation (Pickering and Garrod, 2013; compare Overy and Molnar-Szakacs, 2009 on music). Visual narrative processing may involve the simulation of content but is less likely to involve simulation of production, if at all.

Equivalence of stimuli and tasks, then, may be an impossible goal. A possible alternative is to investigate graded effects on a continuum of stimuli. In an analysis of an oral storyteller's performance, Lwin (2010) shows how the combined effects of the performer's motion and prosody create a unified interpretation of the text encouraging "relatively uniform cognitive, emotive and evaluative responses" in the live audience. A continuum from silent reading, to audio only, to live performance with different degrees of motion could then be used to grade how far audience responses diverge from one another at each stage. Similarly, experiments on serially degraded images could be extended to moving images to identify the relationship between viewer experience and patterns of withheld information (Churchland, 2012, p. 66–67).

An additional approach is to pursue a clearer understanding of low-level processing mechanisms; "further research on the

effects of back-end processes on front-end processes across media is greatly needed" (Magliano et al., 2013, p. 88). In what follows, we identify internal scene construction and visual/verbal working memory as promising areas for future research.

FUTURE DIRECTIONS: INTERNAL SCENE CONSTRUCTION AND THE EPISODIC MEMORY SYSTEM

It is now understood that the episodic memory system is used in a range of tasks beyond event memory, including future planning and imagination (Hassabis and Maguire, 2007; Zeidman et al., 2015). There is also a relationship between episodic memory, the self, and narrative production. This can be seen in the co-emergence in childhood of narrative and episodic memories (Nelson and Fivush, 2004; Hoerl, 2007), the pressure on autobiographical memories to cohere with beliefs about the self (Conway, 2005), and the role of memory in developing a narrative of the self (Fitzgerald, 1988; Wang and Conway, 2006). This relationship may also be seen in findings on the relationships between personal experience and narrative processing (Chow et al., 2015), on false memories in pictures versus verbal narratives (Garry and Wade, 2005), and on a "self-reference" effect (Carson et al., 2016). The default mode network may moderate this relationship between episodic memory, narrative, and the self, through a role in generating narrative both during task-related activities and in "resting" or "screen-saver mode" (Gerrans, 2014, p. 5). Pearson et al. (2015) relate the physical closeness of "high level areas ... to memory-encoding structures" to the overlap in perception and imagery, and the claim that "mental imagery is presumably based on the recall and recombination of memories" (595). It is perhaps unsurprising, then, that readers can engage in episodic future thinking in which they project their selves on to characters (Buckner and Carroll, 2006; Finnigan, 2013, p. 151–160).

Experience of verbal narratives, rather than moving image ones, may supply models for the social presentation of the self through autobiographical narrative (Labov and Waletzky, 1967; Habermas and Paha, 2001; Rubin et al., 2011; Jobson et al., 2014). Fioretti and Smorti (2015) compared autobiographical reminiscences with narratives of those memories. The narratives included more emotion and emotional complexity than the reminiscences, especially in relation to surprise, recalling Bauman's account of narration of stories creating more suspense when performed orally to unfamiliar audiences (Bauman, 1986).

Neuropsychological evidence suggests that a crucial element in this set of relationships between narrative, episodic memory, and the self is the ability to imagine scenes (Hassabis et al., 2007). Damage to the medial temporal lobe (MTL) affects both episodic memories for one's own past and the ability to imagine the future (Tulving, 1985; Addis et al., 2007; Andelman et al., 2010). Race et al. (2011) asked patients with MTL lesions to: remember specific personal events from the past; imagine specific personal events in the future; and imagine that each of five detailed drawings of a scene, shown sequentially, was a scene from a movie and to tell a story about the scene. As

in previous studies, both episodic memories and future thoughts were significantly less detailed than those of controls. But this was not the case for the pictures, which acted as an external memory of the relevant scenes to afford narrative production (a role which illustrations may sometimes play in relation to text; Glenberg and Langston, 1992; Kruley et al., 1994). The capacity to generate internal scenes, then, is important in remembering, imagining, and producing narrative.

Verfaellie et al. (2014) worked with eight patients without the pictures of scenes used in Race et al. (2011). Researchers selected five fairy tales and four Bible stories. MTL patients were impaired in their ability to recount detailed semantic narratives, despite retaining narrative structure and recognizing story details. The loss appeared to be not of semantic knowledge but of the ability to recollect it in rich detail. This again suggests that the patients in Race et al. (2011) used the drawings of scenes for external scene construction in order to generate a richer narrative, and that the narrative obstacle experienced by the patients in Verfaellie et al. was the inability to imagine a scene in detail. As patients in Mullally et al. put it:

it's as if I have a lot of clothes to hang up in a wardrobe, but there's nothing to hang them on, so they all fall on the floor in a complete mess. (266)

I'm imagining different things happening, but there's no visual scene opening out in front of me. (266)

It's hard trying to get the space. It keeps getting squashed. (266) (Mullally et al., 2012)

Internal scene construction, then, appears to be an important element in processing verbal narrative. And the findings in Race et al. suggest that film viewers may be spared some or all of this cognitive effort. Readers and hearers with impaired scene construction abilities can, to varying degrees, recall story, the bare sequence of events associated with event perception and identified by anthropologists and narratologists as a vehicle for richer retellings. But they struggle to go beyond this.

The role of internal scene construction in verbal narrative may go beyond merely decorating underlying stories with visual detail. Ahmed et al. find that patients with Posterior Cortical Atrophy, who have damage to the visual cortex, perform poorly on tests of autobiographical recall. This is not because there is less overall information in their narratives but because their narratives lack visual and perceptual detail. This seems to be replaced by "semantic" detail, detail external to the event itself, which is included in an explanatory capacity (Ahmed et al., 2018). In the narratives of patients with mild Alzheimer's disease or amnesia compared with those of healthy participants, the loss extends to character motivations, the agents' goals which lie at the heart of situation modeling and earlier "story grammar" approaches to narrative structure (Addis et al., 2009; Schacter et al., 2013; Verfaellie, 2014). This leaves a temporal sequence of actions with under- or unspecified causal relations to one another, potentially leaving patients to rely on serial memory for recall (Gentner, 1976). This kind of skeletal sequence

of events, referred to in this review as “story,” with absent or ambiguous causation between episodes, has been identified by folklorists analyzing the underlying structure in variants of oral narratives such as folk and fairy tales (Aarne and Thompson, 1961; Bauman, 1986). It also recalls Bartlett’s findings, in which English participants retold a native American tale with new, more explicit character motivations, without realizing they had done so (Bartlett, 2014 [1932]). The different versions of a tale type can indeed vary dramatically from one another in causal links between episodes, especially when they cross cultures. A villain in one version of a tale can even become a victim in another version (Trompf et al., 1988). The transmission of folk tales, therefore, suggests that a story, or bare sequence of events, and the causal structure relating those events in a particular narrative telling, are separable. Narrative recall by patients supports this suggestion that episode sequence and causal relationships are separable. Moya et al., for example, found that right hemisphere damaged patients “were impaired on all aspects of visuospatial performance and verbal recall” of their narratives. These patients had difficulty in “interrelating components to one another, drawing inferences, and selecting an appropriate structure” (Moya et al., 1986, p. 387).

Internally constructed scenes, then, may be more than just a vehicle for vivid detail; they may build in the structures of causation missing from some patients’ narrative retellings. Research on boundary extension supports this claim. Boundary extension is an error in which scenes are remembered as extending further than they actually do (Intraub and Richardson, 1989):

When we initially encounter a scene, we are not limited to the information that is in front of our eyes, but have access to an automatically constructed and implicitly maintained internal representation of the scene...[which] extends well beyond the borders of the given scene and provides an overarching framework into which we rapidly embed what is currently in our field of view. This is a highly adaptive process that supports our experience of a continuous and coherent world, despite it being amassed from discontinuous sensory input. (Mullally et al., 2012, p. 261)

Patients with hippocampal lesions show attenuated boundary extension; the patients remember scene boundaries more accurately than controls *because* they appear to “have a fundamental problem generating internalized scene representations” (263). As with Verfaellie et al. (2014), patients did not have difficulties in perceiving the pictures of scenes. With scenes before them, they could also “anticipate what might be beyond the view in the scene” (263). However, when asked to imagine an extension of the scene, while patients could predict what might be there, their descriptions lacked spatial coherence. They could bring to mind contextual associations but not organize them in spatial relationships (266).

Spatial relationships are integral to causal relationships (Brass et al., 2000; Fugelsang et al., 2005). Even in a static scene, spatial information lets us calculate a range of potential causal interactions in the future. Simply knowing that a scene has a plant pot, a person, and a watering can in it, without seeing

that scene, already raises the possibility of the person watering the plants. But if we see that the plant and the can are both situated on a high shelf above the person’s sight line, then we can model more accurately the agent’s likely patterns of attention and behavior, the physical effort required to take down plant and can and so on, and our range of potential causal interactions is both expanded and refined. “Causality,” in this sense of forces in a space, need not be distinguished from the character goals, motivations, and interactions with objects identified in situation model theory (Zwaan and Radvansky, 1998). There is a certain probability that the person will reach for the watering can, another that the can will fall by accident on the person’s head alongside all the other potential options, including no interaction with can or plant at all. Both kinds of cause (pursuit of agents’ goals and the laws of physics) are built into the scene’s construction and potentially interact with one another.

Readers then, compared to viewers of film, may rely much more heavily on *internal* scene construction not just for ornamental vivid detail but for a unified sense of the set of possible interactions between all the elements of a scene, from agents and objects to surfaces and volumes, and therefore for local and global causal (and not just serial) narrative connections. Since the internal scenes generated by verbal narrative are likely to be more sparse than those generated by an ongoing film narrative, the causal relations between the elements of those sparser scenes are also likely to be more flexible and to have more variation between individuals. A film, for example, can capture fixed and precise relations of shadow and proximity, with causal implications for the way elements in the scene, including people, relate to one another. This suggestion is consistent with current views of concepts as “flexible, distributed representations comprised of modality specific conceptual features” (Kiefer and Pulvermüller, 2012, p. 805). Readers have more scope to exploit this flexibility than viewers as they instantiate concepts through scene construction, with ensuing flexibility in generating causal explanations at local and global levels (Jajdelska, 2016).

In this way, narrative causality may be more highly determined between individuals by film than by text. As Heider (1944) observed, “a change in the environment gains its meaning from the source to which it is attributed. This causal integration is of major importance in the organization of the social field” (372). Character traits and agency, key to narrative processing, may seem intuitively to be comprehensible without spatial representations. But Heider and Simmel (1944) show how pregnant with agent-directed change a spatial scene can be. Similarly, Finnigan observes that verbal descriptions often mimic the holistic early stages of visual scene perception (Finnigan, 2013), and Fugelsang observes that “extracting causal structure” is “an inherent property of the visual system” (Fugelsang et al., 2005, p. 45). Findings by Jahn (2004) also suggest that causal relevance is “a precondition for the spontaneous construction of spatial situation models” (see also Radvansky and Copeland, 2000). Even those narrative causal relations that seem to go beyond the direct phenomenal causality implied in a spatial representation, those which are often thought to require reasoning

(“It was probably the drink because he fell in love while drinking the cocktail”), can be related to a spatial framework (Oestermeier and Hesse, 2000). Kruley et al. (1994) suggest that spatial representations are used to represent non-spatial information. Negative and positive emotions, for example, can be mapped in contrasting areas of space (Myachykov et al., 2017). Internal scene construction, then, suggests how early processing can be related to higher level of interpretation over shorter or longer time scales, through support for actual and potential causal relations at different levels of abstraction. Indeed, visual scenes may just be a prominent example of a wider family of complex inter-relations between “objects” necessary in imagination. There is work, for example, on “auditory scenes,” which have a comparable relational structure, and the perception/imagery of which is also impaired in hippocampal patients (Teki et al., 2012). Social hierarchies also rely on comparable relational structures (Kumaran et al., 2012).

Their higher reliance on internal scene construction, then, suggests that readers will generate a more varied and flexible causal model to underpin the narrative than viewers. In addition, the sparser visual information of internally constructed scenes can allow for more flexible and individualized instantiation of concepts than that of externally visible moving images (Mahon and Caramazza, 2009; Kiefer and Pulvermüller, 2012). These differences may afford readers a closer relationship between the self and narrative processing than viewing does, as readers draw more heavily on memories of their own lives to construct internal scenes. Gordon et al. (2009), for example, found that memories for text narratives resembled memories of imagined events, whereas memories for film narratives resembled memories for real-world events. Garry and Wade (2005) found that readers of modified personal text narratives were more vulnerable to false memories than those who saw doctored photographs. Carson et al. (2016) find a “self-reference effect” for text narrative, where events related to the self are better remembered than those which are not.

Neuropsychological evidence, then, will likely prove important to the future of narrative research by medium, as will methods of measuring how far memory, imagination, and navigation systems are engaged, for example, through measures of glucose consumption in retrosplenial cortex (Vann et al., 2009).

FURTHER DIRECTIONS FOR FUTURE RESEARCH: WORKING MEMORY, PREDICTIVE PROCESSING, AND AI

We saw earlier that individual differences in imagery may interact with narrative medium. There is some evidence that high imagers also differ from low imagers in strategies for working memory tasks. Pearson et al. (2015), in a review of mental imaging research, find differences between higher and lower imagers in the strategies used for visual working memory tasks (594; see also Smith et al., 1996). Low imagers pick out details from a scene or array, encode them phonologically, and then compare to the subsequent stimulus. High imagers create a mental image and compare it directly with the stimulus

(Gur and Hilgard, 1975; Berger and Gaunitz, 1979; Harrison and Tong, 2009; Keogh and Pearson, 2014; Magliano et al., 2016). This suggests that the possibility of a set of interactions between the episodic buffer; the default mode network (associated with combining meanings to form narratives); episodic long-term memory; and the two “slave systems” of working memory (the phonological loop and the visuospatial sketchpad (Baddeley, 2008, 2018; Smallwood et al., 2016, p. 326, citing Olson et al., 2007; Domhoff, 2018, p. 154–161). This in turn suggests a potential pathway from local, short-term narrative inferencing mechanisms to larger narrative meanings and structures. Differential strategies at the working memory level between high and low imagers may offer a way to explore how that pathway is affected by narrative medium.

New theoretical frameworks for cognition as predictive, and involving high-dimensional vector spaces, may also suggest new methods to capture low-level narrative inferencing mechanisms. Spivey (2007), for example, has drawn attention to increasingly influential Bayesian models of the brain, continually generating and revising predictive hypotheses by drawing on all current input simultaneously, from bottom-up to top-down. Older, modular models of processing, in which one task is completed before the next can be begun, were suited to behavioral experiments in which task outcomes were measured once, at the conclusion of the task. Spivey argues that the Bayesian brain can be better understood by continuously measuring activity throughout the performance of the task. For example, tasks that require participants to choose one of two verbal alternatives to complete a phrase can be measured using the participants’ movements of a computer mouse, which tracks their movements to and from the potential targets leading up to their final decision. Adapting this approach to moving image processing might allow a comparison between low-level mechanisms for narrative processing that did not rely on attempts to produce equivalent stimuli. Emerging neural, or deep learning, AI networks use high dimensional vector spaces to learn processing skills. These are still some way from successfully processing narrative in any medium, but they can potentially be used for proof of concept. Inverting neural networks that give verbal labels to images has revealed that they create some surprising implicit causal relationships in the way they interpret their training data (Mordvintsev et al., 2015). Investigating this implicit causality might shed light on the process of generating causal links between mental images produced in response to verbal input. Here too there may be scope for exploring potential low-level processing differences by medium, which can illuminate internal scene construction by humans.

CONCLUSION

Comparing narrative processing by medium is difficult because it is not possible to have equivalent stimuli. To address this problem, we recommend a focus on neural mechanisms at comparatively low levels of processing and time scales, in combination with an awareness of the rich and holistic nature of narrative experience, which can encompass memory,

imagination, empathy, spatial resources, inference, emotion, and transportation as well as the potential to create meaning in relation to the self. The wide range of benefits associated with reading or hearing fiction, coupled with the richness of the experience, suggests that a simplistic explanation involving cognitive transfer may not be available (Melby-Lervåg et al., 2016). Existing work points to a greater diversity among readers/hearers of verbal stories than among viewers of moving image stories. We suggest that one explanation for this finding is that processing written fiction relies more heavily on the resources of the episodic memory system. This suggests a mechanism that could help explain the different effects of the two media, with written fiction using internal scene construction to make causal predictions, and in doing so, readers/hearers modify their existing model of the world more extensively than viewers and more distinctly from one another. Narrative processing, it seems, is more than information processing. Readers and viewers may extract an underlying plot structure

from films and books which share a narrative, but their experiences may nonetheless be very different.

AUTHOR CONTRIBUTIONS

EJ contributed to identifying articles and coordination of reviews, writing, and redrafting the manuscript. All authors have contributed to reading and reviewing the articles and commenting on drafts.

FUNDING

This work is supported by the Royal Society of Edinburgh, Arts and Humanities Research Network Award, 2017–2018, and by the Strategic Research Fund of the School of Humanities at the University of Strathclyde.

REFERENCES

- Aarne, A., and Thompson, S. (1961). *The types of the folktale: A classification and bibliography*. (Academia Scientiarum Fennica: Helsinki).
- Addis, R., Sacchetti, D., Brandon, A., and Schacter, D. (2009). Episodic simulation of future events is impaired in mild Alzheimer's disease. *Neuropsychologia* 47, 2660–2671. doi: 10.1016/j.neuropsychologia.2009.05.018
- Addis, D. R., Wong, A., and Schacter, D. (2007). Remembering the past and imagining the future: common and distinct neural substrates during event construction and elaboration. *Neuropsychologia* 45, 1363–1377. doi: 10.1016/j.neuropsychologia.2006.10.016
- Ahmed, S., Irish, M., Loane, C., Baker, I., Husain, M., Thompson, S., et al. (2018). Association between precuneus volume and autobiographical memory impairment in posterior cortical atrophy: beyond the visual syndrome. *Neuroimage Clin.* 18, 822–834. doi: 10.1016/j.nicl.2018.03.008
- Alber, J. (2016). *Unnatural narrative: Impossible worlds in fiction and drama*. (Lincoln: University of Nebraska Press).
- Allbritton, D., and Gerrig, R. (1991). Participatory responses in text understanding. *J. Mem. Lang.* 30, 603–626. doi: 10.1016/0749-596X(91)90028-I
- Andelman, F., Hoofen, D., Goldberg, I., Aizenstein, O., and Neufeld, M. (2010). Bilateral hippocampal lesion and a selective impairment of the ability for mental time travel. *Neurocase* 16, 426–435. doi: 10.1080/13554791003623318
- Anderson, M., and Iversen, S. (2018). Immersion and defamiliarization: experiencing literature and world. *Poetics Today* 39, 569–595. doi: 10.1215/03335372-7032760
- Baddeley, A. (2008). The episodic buffer: a new component of working memory? *Trends Cogn. Sci.* 4, 417–423.
- Baddeley, A. (2018). *Exploring working memory: Selected works of Alan Baddeley*. (Abingdon: Routledge).
- Baggett, P. (1975). Memory for explicit and implicit information in picture stories. *J. Verbal Learn. Verbal Behav.* 14, 538–548. doi: 10.1016/S0022-5371(75)80031-4
- Baggett, P. (1979). Structurally equivalent stories in movie and text and the effect of medium on recall. *J. Verbal Learn. Verbal Behav.* 18, 333–356. doi: 10.1016/S0022-5371(79)90191-9
- Baggett, P. (1984). Role of temporal overlap of visual and auditory material in forming dual media associations. *J. Educ. Psychol.* 76, 408–417. doi: 10.1037/0022-0663.76.3.408
- Baggett, P., and Ehrenfeucht, A. (1982). Information in content equivalent movie and text stories. *Discourse Process.* 5, 73–99.
- Bal, P., and Veltkamp, M. (2013). How does fiction reading influence empathy? An experimental investigation on the role of emotional transportation. *PLoS One* 8:E55341. doi: 10.1371/journal.pone.0055341
- Barber, K. (2007). *The anthropology of texts, persons and publics: Oral and written culture in Africa and beyond*. (Cambridge: Cambridge University Press).
- Barrett, L., and Bar, M. (2009). See it with feeling: affective predictions during object perception. *Philos. Trans. R. Soc. B* 364, 1325–1334. doi: 10.1098/rstb.2008.0312
- Barsalou, L., Simmons, W., Barbey, A., and Wilson, C. D. (2003). Grounding conceptual knowledge in modality-specific systems. *Trends Cogn. Sci.* 7, 84–91. doi: 10.1016/S1364-6613(02)00029-3
- Bartlett, F. (2014 [1932]). *Remembering: A study in experimental and social psychology*. (Cambridge: Cambridge University Press).
- Bartlett, J., Till, R., and Levy, J. (1980). Retrieval characteristics of complex pictures: effects of verbal encoding. *J. Verbal Learn. Verbal Behav.* 19, 430–449.
- Basil, M. (1995). "Secondary reaction tasks" in *Measuring psychological responses to media*. ed. A. Lang (Hillsdale, NJ: Erlbaum).
- Bauman, R. (1986). *Chapter five. Story, performance, and event: Contextual studies of oral narrative*. (Cambridge: Cambridge University Press), 78–111.
- Beagles-Roos, J., and Gat, I. (1983). Specific impact of radio and television on children's story comprehension. *J. Educ. Psychol.* 75, 128–137. doi: 10.1037/0022-0663.75.1.128
- Beentjes, J., and van der Voort, T. (1988). Television's impact on children's reading skills: a review of research. *Read. Res. Q.* 23, 389–413. doi: 10.2307/747640
- Beentjes, J., and van der Voort, T. (1991). Children's written accounts of televised and printed stories. *Educ. Technol. Res. Dev.* 39, 15–26. doi: 10.1007/BF02296435
- Berger, G., and Gaunitz, S. (1979). Self-rated imagery and encoding strategies in visual memory. *Br. J. Psychol.* 70, 21–24. doi: 10.1111/j.2044-8295.1979.tb02137.x
- Beveridge, M., and Griffiths, V. (1987). The effect of pictures on the reading processes of less able readers: a miscue analysis approach. 10, 29–42.
- Black, J., and Barnes, J. (2015). Fiction and social cognition: the effect of viewing award-winning television dramas on theory of mind. *Psychol. Aesthet. Creat. Arts* 9, 423–429. doi: 10.1037/aca0000031
- Bohn-Gettler, C., Rapp, D., van den Broek, P., Kendeou, P., and White, M. (2011). Adults' and children's monitoring of story events in the service of comprehension. *Mem. Cogn.* 39, 992–1011. doi: 10.3758/s13421-011-0085-0
- Bordeaux, B., and Lange, G. (1991). Children's reported investment of effort when viewing television. *Commun. Res.* 18, 617–635.
- Brass, M., Bekkering, H., Wohlschläger, A., and Prinz, W. (2000). Compatibility between observed and executed finger movements: comparing symbolic, spatial and imitative cues. *Brain Cogn.* 44, 124–143. doi: 10.1006/brcg.2000.1225
- Brown, C., and Lloyd-Jones, T. (2006). Beneficial effects of verbalization and visual distinctiveness on remembering and knowing faces. *Mem. Cogn.* 34, 277–286. doi: 10.3758/BF03193406
- Buckner, R., and Carroll, D. (2006). Self-projection and the brain. *Trends Cogn. Sci.* 11, 49–57. doi: 10.1016/j.tics.2006.11.004
- Busselle, R., and Bilandzic, H. (2008). Fictionality and perceived realism in experiencing stories: a model of narrative comprehension and engagement. *Commun. Theory* 18, 255–280. doi: 10.1111/j.1468-2885.2008.00322.x

- Calvo, M., and Castillo, M. (2001). Bias in predictive inferences during reading. *Discourse Process.* 32, 43–71. doi: 10.1207/S15326950DP3201_03
- Carson, N., Murphy, K., Moscovitch, M., and Rosenbaum, R. (2016). Older adults show a self-reference effect for narrative information. *Memory* 24, 1157–1172. doi: 10.1080/09658211.2015.1080277
- Center, Y., Freeman, L., Robertson, G., and Outhred, L. (1999). The effect of visual imagery training on the reading and listening comprehension of low listening comprehenders in year 2. *J. Res. Read.* 22, 241–256.
- Chow, H., Mar, R., Xu, Y., Liu, S., Wagage, S., and Braun, A. (2015). Personal experience with narrative events modulates functional connectivity within visual and motor systems during story comprehension. *Hum. Brain Mapp.* 36, 1494–1505. doi: 10.1002/hbm.22718
- Churchland, P. (2012). *Plato's camera: How the physical brain captures a landscape of abstract universals.* (Cambridge Mass: MIT Press).
- Conway, M. (2005). Memory and the self. *J. Mem. Lang.* 53, 594–628. doi: 10.1016/j.jml.2005.08.005
- Corrigan, R., and Surber, J. (2010). The reading level paradox: why children's picture books are less cohesive than adult books. *Discourse Processes* 47, 32–54.
- Crutcher, R., and Beer, J. (2011). An auditory analog of the picture superiority effect. *Mem. Cogn.* 39, 63–74. doi: 10.3758/s13421-010-0015-6
- Cunningham, A., and Stanovich, K. (1998). What reading does for the mind. *Am. Educ.* 22, 8–15.
- Cutting, J., Brunick, K., DeLong, J., Iricinschi, C., and Candan, A. (2011). Quicker, faster, darker: changes in Hollywood film over 75 years. *i-Perception* 2, 569–576. doi: 10.1068/i0441aap
- Denis, M. (1982). Imaging while reading text: a study of individual differences. *Mem. Cogn.* 10, 540–545. doi: 10.3758/BF03202436
- Dodson, C., Johnson, M., and Schooler, J. (1997). The verbal overshadowing effect: why descriptions impair face recognition. *Mem. Cogn.* 25, 129–139. doi: 10.3758/BF03201107
- Domhoff, G. W. (2018). *The emergence of dreaming: Mind-wandering, embodied simulation and the default network.* (Oxford: Oxford University Press).
- Edwards, D., and Middleton, D. (1986). Joint remembering: constructing an account of shared experience through conversational discourse. *Discourse Process.* 9, 423–459. doi: 10.1080/01638538609544651
- Eitel, A., and Scheiter, K. (2015). Picture or text first? Explaining sequence effects when learning with pictures or text. *Educ. Psychol. Rev.* 27, 153–180. doi: 10.1007/s10648-014-9264-4
- Ennemoser, M., and Schneider, W. (2007). Relations of television viewing and reading: findings from a longitudinal study. *J. Educ. Psychol.* 99, 349–368. doi: 10.1037/0022-0663.99.2.349
- Finnigan, E. (2013). A cognitive approach to spatial patterning in literary narrative. PhD thesis. (Glasgow: University of Strathclyde).
- Fioretti, C., and Smorti, A. (2015). How emotional content of memories changes in narrating. *Narrat. Inq.* 25, 37–56.
- Fitzgerald, J. M. (1988). Vivid memories and the reminiscence phenomenon: the role of a self narrative. *Hum. Dev.* 31, 261–273. doi: 10.1159/000275814
- Fludernik, M. (2008). "Narrative and drama" in *Theorizing narrativity.* eds. J. Pier, and J. Landa (Berlin: Walter de Gruyter), 355–384.
- Forceville, C. (2002). The conspiracy in *the comfort of strangers*: narration in the novel and the film. *Lang. Lit.* 11, 119–135.
- Foy, J., and Gerrig, R. (2014). Flying to neverland: how readers tacitly judge norms during comprehension. *Mem. Cogn.* 42, 1250–1259. doi: 10.3758/s13421-014-0436-8
- Friedman, A., and Waller, D. (2008). View combination in scene recognition. *Mem. Cogn.* 36, 467–478. doi: 10.3758/MC.36.3.467
- Fugelsang, J., Roser, M., Corballis, P., Gazzaniga, M., and Dunbar, K. (2005). Brain mechanisms underlying perceptual causality. *Cogn. Brain Res.* 24, 41–47. doi: 10.1016/j.cogbrainres.2004.12.001
- Gallagher, H., Happé, F., Brunswick, N., Fletcher, P., Frith, U., and Frith, C. (2000). Reading the mind in cartoons and stories: an fMRI study of 'theory of mind' in verbal and nonverbal tasks. *Neuropsychologia* 38, 11–21. doi: 10.1016/S0028-3932(99)00053-6
- Garry, M., and Wade, K. (2005). A picture is worth less than 45 words: narratives produce more false memories than photographs do. *Psychon. Bull. Rev.* 12, 359–366. doi: 10.3758/BF03196385
- Garsoffky, B., Schwan, S., and Huff, M. (2009). Canonical views of dynamic scenes. *J. Exp. Psychol. Hum. Percept. Perform.* 35, 17–27. doi: 10.1037/0096-1523.35.1.17
- Gati, I., and Tversky, A. (1987). Recall of common and distinctive features of verbal and pictorial stimuli. *Mem. Cogn.* 15, 97–100. doi: 10.3758/BF03197020
- Gennette, G. (1980 [1972]). in *Narrative discourse: an essay in method.* ed. J. I. Lewin (New York: Cornell University Press).
- Gentner, D. (1976). The structure and recall of narrative prose. *J. Verbal Learn. Verbal Behav.* 15, 411–418. doi: 10.1016/S0022-5371(76)90036-0
- Gerrans, P. (2014). Pathologies of hyperfamiliarity in dreams, delusions and déjà vu. *Front. Psychol.* 5:97. doi: 10.3389/fpsyg.2014.00097
- Gibbons, J., Anderson, D., Smith, R., Field, D., and Fischer, C. (1986). Young children's recall and reconstruction of audio and audiovisual narratives. *Child Dev.* 57, 1014–1023. doi: 10.2307/1130375
- Gilbert, D. (1991). How mental systems believe. *Amer. Psychol.* 46, 107–119.
- Glenberg, A., and Langston, W. (1992). Comprehension of illustrated text: pictures help to build mental models. *J. Mem. Lang.* 129–151.
- Gordon, R., Gerrig, R., and Franklin, N. (2009). Qualitative characteristics of memories for real, imagined and media-based events. *Discourse Process.* 46, 70–91. doi: 10.1080/01638530802629117
- Graber, D. (1990). Seeing is remembering: how visuals contribute to learning from television. *J. Commun.* 40, 134–156. doi: 10.1111/j.1460-2466.1990.tb02275.x
- Green, M., and Brock, T. C. (2000). The role of transportation in the persuasiveness of public narratives. *J. Pers. Soc. Psychol.* 79, 701–721. doi: 10.1037/0022-3514.79.5.701
- Green, M., Kass, S., Carrey, J., Herzig, B., Feeney, R., and Sabini, J. (2008). Transportation across media: repeated exposure to print and film. *Media Psychol.* 11, 512–539. doi: 10.1080/15213260802492000
- Greenfield, P., Farrar, D., and Beagles-Roos, J. (1986). Is the medium the message? An experimental comparison of the effects of radio and television on imagination. *J. Appl. Dev. Psychol.* 7, 201–218. doi: 10.1016/0193-3973(86)90029-8
- Grimes, T. (1991). Mild auditory-visual dissonance in television news may exceed viewer attention capacity. *Hum. Commun. Res.* 18, 268–298. doi: 10.1111/j.1468-2958.1991.tb00546.x
- Gur, R., and Hilgard, E. (1975). Visual imagery and the discrimination of differences between altered pictures simultaneously and successively presented. *Br. J. Psychol.* 66, 341–345. doi: 10.1111/j.2044-8295.1975.tb01470.x
- Habermas, T., and Paha, C. (2001). The development of coherence in adolescents' life narratives. *Narrat. Inq.* 11, 35–54. doi: 10.1075/ni.11.1.02hab
- Hakemulder, J. (2004). Foregrounding and its effect on readers. *Discourse Process.* 38, 193–218.
- Hakemulder, J. (2007). Tracing foregrounding in responses to film. *Lang. Lit.* 16, 125–139.
- Harris, S., and Fiske, T. (2006). Dehumanizing the lowest of the low: neuroimaging responses to extreme out-groups. *Psychol. Sci.* 17, 847–853. doi: 10.1111/j.1467-9280.2006.01793.x
- Harrison, S., and Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. *Nature* 458, 632–635. doi: 10.1038/nature07832
- Hassabis, D., Kumaran, D., Vann, S., and Maguire, E. (2007). Patients with hippocampal amnesia cannot imagine new experiences. *Proc. Natl. Acad. Sci. USA* 104, 1726–1731. doi: 10.1073/pnas.0610561104
- Hassabis, D., and Maguire, E. (2007). Deconstructing episodic memory with construction. *Trends Cogn. Sci.* 11, 299–306. doi: 10.1016/j.tics.2007.05.001
- Hasson, U., Landesman, O., Knappmeyer, B., Vallines, I., Rubin, N., and Heeger, D. (2008). Neurocinematics: the neuroscience of film. *Projections* 2, 1–26. doi: 10.3167/proj.2008.020102
- Hauk, O., Johnsrude, I., and Pulvermüller, F. (2004). Somatotopic representation of action words in the motor and premotor cortex. *Neuron* 41, 301–307. doi: 10.1016/S0896-6273(03)00838-9
- Hauk, O., and Tschentscher, N. (2013). The body of evidence: what can neuroscience tell us about embodied semantics? *Front. Psychol.* 4, 1–14. doi: 10.3389/fpsyg.2013.00050
- Hayes, D., and Birnbaum, D. (1980). Preschoolers' retention of televised events: is a picture worth a thousand words? *Dev. Psychol.* 16, 410–416. doi: 10.1037/0012-1649.16.5.410
- Heider, F. (1944). Social perception and phenomenal causality. *Psychol. Rev.* 51, 358–374. doi: 10.1037/h0055425

- Heider, F., and Simmel, M. (1944). An experimental study of apparent behaviour. *Am. J. Psychol.* 57, 243–259. doi: 10.2307/1416950
- Hockley, W. (2008). The picture superiority effect in associative recognition. *Mem. Cogn.* 36, 1351–1359. doi: 10.3758/MC.36.7.1351
- Hoekner, B., Wyatt, E., Decety, J., and Nusbaum, H. (2011). Film music influences how viewers relate to movie characters. *Psychol. Aesthet. Creat. Arts* 5, 146–153. doi: 10.1037/a0021544
- Hoerl, C. (2007). Episodic memory, autobiographical memory, narrative: on three key notions in current approaches to memory development. *Philos. Psychol.* 20, 621–640. doi: 10.1080/09515080701537988
- Huff, M., and Schwan, S. (2008). Verbalizing events: overshadowing or facilitation? *Mem. Cogn.* 36, 392–402. doi: 10.3758/MC.36.2.392
- Huff, M., and Schwan, S. (2012). Do not cross the line: heuristic spatial updating in dynamic scenes. *Psychon. Bull. Rev.* 19, 1065–1072. doi: 10.3758/s13423-012-0293-z
- Hughes, T., Chang, C., Vanderbilt, J., and Ganguli, M. (2010). Engagement in reading and hobbies and incident dementia in the community: the MoVIEW Project. *Am. J. Geriatr. Psychiatr.* 18, S100–S100.
- Ildirar, S., and Schwan, S. (2015). First-time viewers' comprehension of films: bridging shot transitions. *Br. J. Psychol.* 106, 133–151. doi: 10.1111/bjop.12069
- Intraub, H., and Richardson, M. (1989). Wide-angle memories of close-up scenes. *J. Exp. Psychol. Learn. Mem. Cogn.* 15, 179–187.
- Jahn, G. (2004). Three turtles in danger: spontaneous construction of causally relevant spatial situation models. *J. Exp. Psychol. Learn. Mem. Cogn.* 30, 969–987. doi: 10.1037/0278-7393.30.5.969
- Jajdelska, E. (2007). *Silent reading and the birth of the narrator*. (Toronto: University of Toronto Press).
- Jajdelska, E. (2016). Being there yet not there: why don't embodied responses to literary texts jar with one another? *J. Lit. Semant.* 45, 1–20. doi: 10.1515/jls-2016-0002
- Jajdelska, E. (2019). The flow of narrative in the mind unmoored: an account of narrative processing. *Philos. Psychol.* 32, 560–583. doi: 10.1080/09515089.2019.1585796
- Jajdelska, E., Butler, C., Kelly, S., McNeill, C., and Overy, K. (2010). Crying, moving and keeping it whole: what makes literary description vivid? *Poetics Today* 31, 433–463.
- Jensen, J., Martins, N., Weaver, J., and Ratcliff, C. (2016). Educational TV consumption and children's interest in leisure reading and writing: a test of the validated curriculum hypothesis. *J. Broadcast. Electron. Media* 60, 213–230. doi: 10.1080/08838151.2016.1164161
- Jobson, L., Moradi, A., Rahimi-Movaghar, V., Conway, M., and Dalgleish, T. (2014). Culture and the remaining of trauma. *Clin. Psychol. Sci.* 2, 696–713. doi: 10.1177/2167702614529763
- Kauppi, J., Jääskeläinen, I., Sams, M., and Tohka, J. (2010). Inter-subject correlation of brain hemodynamic responses during watching a movie: localization in space and frequency. *Front. Neuroinform.* 4, 1–10. doi: 10.3389/fninf.2010.00005
- Kendeou, P., Bohn-Gettler, C., White, M., and van den Broek, P. (2008). Children's inference generation across different media. *J. Res. Read.* 31, 259–272. doi: 10.1111/j.1467-9817.2008.00370.x
- Keogh, R., and Pearson, J. (2014). The sensory strength of voluntary visual imagery predicts visual working memory capacity. *J. Vis.* 14, 1–13. doi: 10.1167/14.12.7
- Kidd, D., and Castano, E. (2013). Reading literary fiction improves theory of mind. *Science* 342, 377–380. doi: 10.1126/science.1239918
- Kiefer, M., and Pulvermüller, F. (2012). Conceptual representations in mind and brain: theoretical developments, current evidence and future directions. *Cortex* 48, 805–825. doi: 10.1016/j.cortex.2011.04.006
- Kintsch, W., and Van Dijk, T. A. (1975). Comment on se rappelle et on resume des histoires [How stories are recalled and summarized]. *Langages* 40, 98–128.
- Klatzky, R., Martin, G., and Kane, R. (1982). Semantic interpretation effects on memory for faces. *Mem. Cogn.* 10, 195–206. doi: 10.3758/BF03197630
- Koehler, M., Yadav, A., Phillips, M., and Cavazos-Kottke, S. (2005). What is video good for? Examining how media and story genre interact. *J. Educ. Multimed. Hypermedia* 14, 249–272.
- Koopman, E., and Hakemulder, F. (2015). Effects of fiction on empathy and self-reflection: a theoretical-empirical framework. *J. Lit. Theory* 9, 79–111.
- Kosslyn, S., Thompson, W. L., Sukel, K. E., and Alpert, N. M. (2005). Two types of image generation: evidence from PET. *Cogn. Affect. Behav. Neurosci.* 5, 41–53. doi: 10.3758/CABN.5.1.41
- Kraft, R. N. (1987). The influence of camera angle on comprehension and retention of pictorial events. *Mem. Cogn.* 15, 291–307.
- Krusey, P., Sciamia, S. C., and Glenberg, A. (1994). On-line processing of textual illustrations in the visuospatial sketchpad: evidence from dual-task studies. *Mem. Cogn.* 22, 261–272. doi: 10.3758/BF03200853
- Kukkonen, K. (2011). Comics as a test case for transmedial narratology. *SubStance* 40, 34–52.
- Kukkonen, K. (2013). *Studying comics and graphic novels*. (Oxford: Wiley-Blackwell).
- Kumaran, D., Melo, H., and Duzel, E. (2012). The emergence and representation of knowledge about social and non-social hierarchies. *Neuron* 76, 653–666. doi: 10.1016/j.neuron.2012.09.035
- Labov, W., and Waletzky, J. (1967). Narrative analysis. *J. Narrat. Life Hist.* 7, 3–38.
- Lang, A., Sias, P., Chantrill, P., and Burek, J. (1995). Tell me a story: narrative elaboration and memory for television. *Commun. Rep.* 8, 102–110. doi: 10.1080/08934219509367616
- Levorato, M. (1991). Children's memory for goal-directed events. *Discourse Process.* 14, 443–467. doi: 10.1080/01638539109544796
- Lichtenstein, E., and Brewer, W. (1980). Memory for goal-directed events. *Cogn. Psychol.* 12, 412–445. doi: 10.1016/0010-0285(80)90015-8
- Loschky, L., Larson, A., Magliano, J., and Smith, T. (2015). What would Jaws do? The Tyranny of film and the relationship between gaze and higher-level narrative film comprehension. *PLoS One* 10:E0142474. doi: 10.1371/journal.pone.0142474
- Lwin, S. (2010). Capturing the dynamics of narrative development in an oral storytelling performance: a multimodal perspective. *Lang. Lit.* 19, 357–377. doi: 10.1177/0963947010373029
- Magliano, J., Dijkstra, K., and Zwaan, R. (1996). Generating predictive inferences while viewing a movie. *Discourse Process.* 22, 199–224.
- Magliano, J., Kopp, K., McNeerney, M., Radvansky, G., and Zacks, J. (2012). Aging and perceived event structure as a function of modality. *Neuropsychol. Dev. Cogn. B Aging Neuropsychol. Cogn.* 19, 264–282.
- Magliano, J., Larson, A., Higgs, K., and Loschky, L. (2016). The relative roles of visuospatial and linguistic working memory systems in generating inferences during visual narrative comprehension. *Mem. Cogn.* 44, 207–219. doi: 10.3758/s13421-015-0558-7
- Magliano, J., Loschky, L., Clinton, J., and Larson, A. (2013). "Is reading the same as viewing? An exploration of the similarities and differences between processing text- and visually-based narratives" in *Unravelling the Behavioral, Neurobiological & Genetic Components of Reading Comprehension*. eds. B. Miller, L. Cutting, and P. McCordle (Baltimore, MD: Brookes Publishing), 78–90.
- Maher, K. (2008). *Steven McQueen's Hunger: Featuring one of cinema's greatest ever scenes*. (London: The Times).
- Mahon, B., and Caramazza, A. (2009). Concepts and categories: a cognitive neuropsychological perspective. *Annu. Rev. Psychol.* 60, 27–51. doi: 10.1146/annurev.psych.60.110707.163532
- Mandler, J. M., and Johnson, N. S. (1977). Remembrance of things parsed: story structure and recall. *Cogn. Psychol.* 9, 111–151. doi: 10.1016/0010-0285(77)90006-8
- Mar, R., Oatley, K., and Peterson, J. (2009). Exploring the link between reading fiction and empathy: ruling out individual differences and examining outcomes. *Communications* 34, 407–428.
- Marks, D. (1973). Visual imagery differences in the recall of pictures. *Br. J. Psychol.* 64, 17–24. doi: 10.1111/j.2044-8295.1973.tb01322.x
- Marsh, E., Meade, M., and Roediger, H. (2003). Learning facts from fiction. *J. Mem. Lang.* 49, 519–536. doi: 10.1016/S0749-596X(03)00092-5
- Mayer, R. E. (2014). "Cognitive theory of multimedia learning" in *The Cambridge handbook of multimedia learning*. 2nd ed (Cambridge: Cambridge University Press), 43–71.
- McEwan, I. (1981). *The comfort of strangers*. (London: Jonathan Cape).
- McGeown, S., Johnston, R., Walker, J., Howatson, K., Stockburn, A., and Dufton, P. (2015). The relationship between young children's enjoyment of learning to read, reading attitudes, confidence and attainment. *Educ. Res.* 57, 389–402. doi: 10.1080/00131881.2015.1091234
- McQueen, S. (director and writer), and Walsh, E. (wrtier). (2008). *Hunger*. Film4, Channel Four Film, Northern Ireland Screen: Ireland and UK.

- Melby-Lervåg, M., Redick, T., and Hulme, C. (2016). Working memory training does not improve performance on measures of intelligence or other measures of “far transfer”: evidence from a meta-analytic review. *Perspect. Psychol. Sci.* 11, 512–534. doi: 10.1177/1745691616635612
- Merga, M. (2016). What would make them read more? Insights from Western Australian adolescents. *Asia Pac. J. Educ.* 36, 409–424.
- Merga, M. (2017). What would make children read for pleasure more frequently? *Research Journal of the National Association for the Teaching of English* 51, 207–223.
- Merga, M., and Roni, S. (2017). The influence of access to eReaders, computers and mobile phones on children's book reading frequency. *Comput. Educ.* 109, 187–196. doi: 10.1016/j.compedu.2017.02.016
- Meyer, B. J. (1975). *The organization of prose and its effects on memory*. (New York: North Holland Publishing Company).
- Moore, P., and Skinner, M. (1985). The effects of illustrations on children's comprehension of abstract and concrete passages. *J. Res. Read.* 8, 45–56.
- Mordvintsev, A., Olah, C., and Tyka, M. (2015). Inceptionism: going deeper into neural networks. Available at: <https://research.googleblog.com/2015/06/inceptionism-going-deeper-into-neural.html> (Accessed May 29, 2019).
- Moya, K., et al. (1986). Covariant defects in visuospatial abilities and recall of verbal narrative after right hemisphere stroke. *Cortex* 22, 381–397. doi: 10.1016/S0010-9452(86)80003-X
- Mullally, S., Intraug, H., and Maguire, E. (2012). Attenuated boundary extension produces a paradoxical memory advantage in amnesic patients. *Curr. Biol.* 22, 261–268. doi: 10.1016/j.cub.2012.01.001
- Myachikov, A., Chapman, A., and Fischer, M. (2017). Cross-representational interactions: interface and overlap mechanisms. *Front. Psychol.* 7:2028. doi: 10.3389/fpsyg.2016.02028
- Nakano, T., Yamamoto, Y., Kitajo, K., Takahashi, T., and Kitazawa, S. (2009). Synchronisation of spontaneous eye blinks while viewing video stories. *Proc. Royal Soc. B* 1673, 3635–3644. doi: 10.1098/rspb.2009.0828
- Nelson, K., and Fivush, R. (2004). The emergence of autobiographical memory: a social cultural developmental model. *Psychol. Rev.* 111, 486–511. doi: 10.1037/0033-295X.111.2.486
- Neuman, S. (1995). *Literacy in the television age: the myth of the TV effect*. (Norwood, New Jersey: Ablex).
- Newton, D. (1973). Attribution and the unit of perception of ongoing behavior. *J. Person. Soc. Psychol.* 28, 28–38.
- Norenzayan, A., Atran, S., Faulkner, J., and Schaller, M. (2006). Memory and mystery: the cultural selection of minimally. *Cogn. Sci.* 30, 531–553. doi: 10.1207/s15516709cog0000_68
- Oestermeier, U., and Hesse, F. (2000). Verbal and visual causal arguments. *Cognition* 75, 65–104. doi: 10.1016/S0010-0277(00)00060-3
- Olson, I. R., Plotzker, A., and Ezzyat, Y. (2007). The enigmatic temporal pole: a review of findings on social and emotional processing. *Brain* 130, 1718–1731. doi: 10.1093/brain/awm052
- Organisation for Economic Co-operation and Development (OECD) (2002). *PISA in focus*. (Washington DC: Author).
- Organisation for Economic Co-operation and Development (OECD) (2010). *PISA 2009 results: Executive summary*. (Washington DC: Author).
- Otgaar, H., Smeets, T., and Bergen, S. (2010). Picturing survival memories: enhanced memory after fitness-relevant processing occurs for verbal and visual stimuli. *Mem. Cogn.* 38, 23–28. doi: 10.3758/MC.38.1.23
- Overwalle, F., and Baetens, K. (2009). Understanding others' actions and goals by mirror and mentalizing systems: a meta-analysis. *NeuroImage* 48, 564–584. doi: 10.1016/j.neuroimage.2009.06.009
- Overy, K., and Molar-Szakacs, I. (2009). Being together in time: musical experience and the mirror neuron system. *Mus. Percept.* 26, 489–504. doi: 10.1525/mp.2009.26.5.489
- Paivio, A. (1971). *Imagery and verbal processes*. (New York: Holt, Rinehart & Winston).
- Palmiter, S., Elkerton, J., and Baggett, P. (1991). Animated demonstrations vs written instructions for learning procedural tasks: a preliminary investigation. *Int. J. Man Mach. Stud.* 34, 687–701. doi: 10.1016/0020-7373(91)90019-4
- Parasuraman, R., and Jiang, Y. (2012). Individual differences in cognition, affect and performance: behavioral, neuroimaging, and molecular genetic approaches. *NeuroImage* 59, 70–82. doi: 10.1016/j.neuroimage.2011.04.040
- Pearson, J., Naselaris, T., Holmes, E., and Kosslyn, S. (2015). Mental imagery: functional mechanisms and clinical applications. *Trends Cogn. Sci.* 19, 590–602. doi: 10.1016/j.tics.2015.08.003
- Pezdek, L., and Stevens, E. (1984). Children's memory for auditory and visual information on television. *Dev. Psychol.* 20, 212–218. doi: 10.1037/0012-1649.20.2.212
- Pezdek, S., Stoeckert, J., and Kiely, J. (1987). Individual differences in television comprehension. *Mem. Cogn.* 15, 428–435. doi: 10.3758/BF03197732
- Pickering, M., and Garrod, S. (2013). An integrated theory of language production and comprehension. *Behav. Brain Sci.* 36, 329–392. doi: 10.1017/S0140525X12001495
- Pinter, H. (1990). *The comfort of strangers and other screenplays*. (London: Faber & Faber).
- PIRL (Progress in International Reading Literacy Study International Report). (2001). Available at: https://timssandpirls.bc.edu/pirls2001i/PIRLS2001_Pubs_IR.html
- PIRL (Progress in International Reading Literacy Study International Report). (2006). Available at: https://timssandpirls.bc.edu/pirls2006/intl_rpt.html
- Poulsen, D., Kintsch, E., Kintsch, K., and Premak, D. (1979). Children's comprehension and memory for stories. *J. Exp. Child Psychol.* 28, 379–404. doi: 10.1016/0022-0965(79)90070-5
- Prentice, D., and Gerrig, R. (1999). “Exploring the boundary between fiction and reality” in *Dual-process Theory in Social Psychology*. eds. S. Chaiken and Y. Trope, 529–549.
- Propp, V. (1968). in *Morphology of the folk tale*. ed. L. Scott (Austin, TX: University of Texas Press).
- Pulvermüller, F. (2013). How neurons make meaning: brain mechanisms for embodied and abstract-symbolic semantics. *Trends Cogn. Sci.* 17, 458–470. doi: 10.1016/j.tics.2013.06.004
- Race, E., Keane, M., and Verfaellie, M. (2011). Medial temporal lobe damage causes deficits in episodic memory and episodic future thinking not attributable to deficits in narrative construction. *J. Neurosci.* 3, 10262–10269.
- Radvansky, G., and Copeland, D. (2000). Functionality and spatial relations in memory and language. *Mem. Cogn.* 28, 987–992. doi: 10.3758/BF03209346
- Reddy, L., Tsuchiya, N., and Serre, T. (2010). Reading the mind's eye: decoding category information during mental imagery. *NeuroImage* 50, 818–825. doi: 10.1016/j.neuroimage.2009.11.084
- Reifel, S. (1984). Symbolic representation at two ages: block buildings of a story. *Discourse Process.* 7, 11–20.
- Reinwein, J. (2012). Does the modality effect exist? And if so, which modality effect? *J. Psycholinguist. Res.* 41, 1–32.
- Ritchie, S., Bates, T. C., and Plomin, R. (2015). Does learning to read improve intelligence? A longitudinal multivariate analysis in identical twins age 7 to 16. *Child Dev.* 86, 23–36. doi: 10.1111/cdev.12272
- Rolandelli, D., Wright, J., Huston, A., and Eakins, D. (1991). Children's auditory and visual processing of narrated and non-narrated television programming. *J. Exp. Child Psychol.* 51, 90–122. doi: 10.1016/0022-0965(91)90078-7
- Rubin, D., Dennis, M., and Beckham, J. (2011). Autobiographical memory for stressful events: the role of autobiographical memory in posttraumatic stress disorder. *Conscious. Cogn.* 20, 840–856. doi: 10.1016/j.concog.2011.03.015
- Rumelhart, D. E. (1975). “Notes on a schema for stories” in *Representation and understanding: Studies in cognitive science*. eds. D. G. Bobrow, and A. Collins (New York: Academic Press), 211–236.
- Rumelhart, D. E. (1977). “Understanding and summarising brief stories” in *Basic processes in reading: Perception and comprehension*. eds. D. LaBerge, and S. J. Samuels (Hillsdale, New York: Lawrence Erlbaum).
- Ryan, M. (2004). “Introduction” in *Narrative across media: The languages of storytelling*. ed. M. Ryan (Lincoln: University of Nebraska Press), 1–40.
- Sabeti, S. (2012). Reading graphic novels in schools: texts, contexts and the interpretative work of critical reading. *Pedagog. Cult. Soc.* 20, 191–210. doi: 10.1080/14681366.2012.672336
- Salomon, G. (1979). *Interaction of media, cognition and learning*. (San Francisco: Jossey-Bass).
- Salomon, G. (1984). Television is ‘easy’ and print is ‘tough’: the differential investment of mental effort in learning as a function of perceptions and attributions. *J. Educ. Psychol.* 76, 647–658. doi: 10.1037/0022-0663.76.4.647
- Salomon, G., and Leigh, T. (1984). Predispositions about learning from print and television. *J. Commun.* 34, 119–135. doi: 10.1111/j.1460-2466.1984.tb02164.x
- Sanford, A., and Emmott, C. (2012). *Mind, brain and narrative*. (Cambridge: Cambridge University Press).

- Schacter, D., Gaesser, B., and Addis, D. (2013). Remembering the past and imagining the future in the elderly. *Gerontology* 59, 143–151. doi: 10.1159/000342198
- Schlottermeier, L., Kuchinke, L., Pehrs, C., Urton, K., Kappelhoff, H., and Jacobs, A. (2013). Emotional picture and word processing: n fMRI study on effects of stimulus complexity. *PLoS One* 8. doi: 10.1371/journal.pone.0055619
- Scholastic. (2015). *Kids and family reading report*. 5th Edn. Available at: https://www.scholastic.com/content/dam/KFRR/PastReports/KFRR2015_5th.pdf (Accessed May 29, 2019).
- Schooler, J., and Engstler-Schooler, T. (1990). Verbal overshadowing of visual memories: some things are better left unsaid. *Cogn. Psychol.* 22, 36–71. doi: 10.1016/0010-0285(90)90003-M
- Schwan, S., and Garsoffky, B. (2004). The cognitive representation of filmic event summaries. *Appl. Cogn. Psychol.* 18, 37–55. doi: 10.1002/acp.940
- Schwan, S., Garsoffky, B., and Hesse, F. (2000). Do film cuts facilitate the perceptual and cognitive organization of activity sequences? *Mem. Cogn.* 28, 214–223. doi: 10.3758/BF03213801
- Schwan, S., and Ildirar, S. (2010). Watching film for the first time: adult viewers interpret perceptual discontinuities in film. *Psychol. Sci.* 2, 970–976. doi: 10.1177/0956797610372632
- Scottish Government (2007). Progress in International Reading Literacy Survey: highlights from Scotland's results. Available at: <http://www.gov.scot/Publications/2007/11/23111011/1> (Accessed May 29, 2019).
- Shimamura, A., Cohn-Sheehy, B., Pogue, B., and Shimamura, T. (2015). How attention is driven by film edits: a multimodal experience. *Psychol. Aesthet. Creat. Arts* 9, 417–422. doi: 10.1037/aca0000025
- Shimamura, A., Cohn-Sheehy, B., and Shimamura, T. (2014). Perceiving movement across film edits: a psychocinematic analysis. *Psychol. Aesthet. Creat. Arts* 8, 77–80. doi: 10.1037/a0034595
- Shklovsky, V. (2011 [1970]) in *The bowstring: On the dissimilarity of the similar*. ed. S. Avagyan (Champaign, IL: Dalkey Archive Press).
- Sikora, J., Evans, M. D. R., and Kelley, J. (2019). Scholarly culture: how books in adolescence enhance adult literacy, numeracy and technology skills in 31 societies. *Soc. Sci. Res.* 77, 1–15. doi: 10.1016/j.ssresearch.2018.10.003
- Simpson, P. (2014). Just what is narrative urgency? *Lang. Lit.* 23, 3–32. doi: 10.1177/0963947013510650
- Singer, D., and Singer, J. (1981). Television and the developing imagination of the child. *J. Broadcast.* 25, 373–387.
- Singer, J., and Singer, D. (2005). Children's narrative play as precursor of narrative consciousness. *Imagin. Cogn. Pers.* 25, 97–117.
- Slone, J., Gonce, L., Upal, A., Edwards, K., and Tweney, R. (2007). Imagery effects on recall of minimally counterintuitive concepts. *J. Cogn. Cult.* 7, 355–367. doi: 10.1163/156853707X208558
- Smallwood, J., Karapanagiotidis, T., Ruby, F., Medea, B., de Caso, I., Konishi, M., et al. (2016). Representing representation: integration between the temporal lobe and the posterior cingulate influences the content and form of spontaneous thought. *PLoS One* 11:e0152272. doi: 10.1371/journal.pone.0152272
- Smith, E., Jonides, J., and Koeppe, R. (1996). Dissociating verbal and spatial working memory using PET. *Cereb. Cortex* 6, 11–20. doi: 10.1093/cercor/6.1.11
- Smith, T., and Mital, P. (2015). Attentional synchrony and the influence of viewing task on gaze behaviour in static and dynamic scenes. *J. Vis.* 13:16.
- Speer, N. J., Reynolds, K., and Zacks, J. (2007a). Human brain activity time-locked to narrative event boundaries. *Psychol. Sci.* 18, 449–455. doi: 10.1111/j.1467-9280.2007.01920.x
- Spivey, M. (2007). *The continuity of mind*. (Oxford: Oxford University Press).
- Steenberg, M., Bräuner, P., and Wallot, S. (2014). Text technology: building subjective and shared experience in reading. *J. Cogn. Cult.* 14, 357–372. doi: 10.1163/15685373-12342131
- Stein, N. L., and Glenn, C. G. (1979). "An analysis of story comprehension in elementary school children" in *New Directions in Discourse Processing*. ed. R. O. Freedle (Norwood, NJ: Ablex), 53–120.
- Strick, M., De Bruin, H., De Ruiter, L., and Jonkers, W. (2015). Striking the right chord: moving music increases psychological transportation and behavioural intentions. *J. Exp. Psychol. Appl.* 21, 57–72. doi: 10.1037/xap0000034
- Sullivan, A., and Brown, M. (2013). *Social inequalities in cognitive scores at age 16: The role of reading*. (London: Centre for Longitudinal Studies).
- Teki, S., Kumar, S., von Kriegstein, K., Stewart, L., Lyness, C. R., Moore, B., et al. (2012). Navigating the auditory scene: an expert role for the hippocampus. *J. Neurosci.* 32, 12251–12257. doi: 10.1523/JNEUROSCI.0082-12.2012
- Thompson, S. (1955-1958). *Motif-Index of Folk Literature*, vol. 6. (Bloomington: Indiana University Press).
- Thorndyke, P. W. (1977). Cognitive structures in comprehension and memory of narrative discourse. *Cogn. Psychol.* 9, 77–110. doi: 10.1016/0010-0285(77)90005-6
- Trabasso, T., and Nickels, M. (1992). The development of goal plans of action in the narration of a picture story. *Discourse Processes* 15, 249–275.
- Trompf, G., Garry, J., and Otto, E. (1988). Western folktales in changing Melanesia. 99, 204–220.
- Troscianko, E. (2013). Reading imaginatively: the imagination in cognitive science and cognitive literary studies. *J. Lit. Semant.* 42, 181–198.
- Tulving, E. (1985). Memory and consciousness. *Can. Psychol.* 26, 1–12.
- Van der Molen, J., and Van der Voort, T. (2000). Children's and adults' recall of television and print news in children's and adult news formats. *Commun. Res.* 27:1320160.
- Vann, S., Aggleton, J., and Maguire, E. (2009). What does the retrosplenial cortex do? *Nat. Rev. Neurosci.* 10, 792–802. doi: 10.1038/nrn2733
- Verfaellie, M., Bousquet, K., and Keane, M. (2014). Medial temporal and neocortical contributions to remote memory for semantic narratives: evidence from amnesia. *Neuropsychologia* 61, 105–112. doi: 10.1016/j.neuropsychologia.2014.06.018
- Verghese, J., Lipton, R., Katz, M., Hall, C., Derby, C., Kuslansky, G., et al. (2003). Leisure activities and the risk of dementia in the elderly. *N. Engl. J. Med.* 348, 2508–2516. doi: 10.1056/NEJMoa022252
- Walsh, R. (2006). Narrative imagination across media. *Mod. Fiction Stud.* 52, 855–868. doi: 10.1353/mfs.2007.0012
- Wang, Q., and Conway, M. (2006). The stories we keep: autobiographical memory in American and Chinese middle-aged adults. *J. Pers.* 72, 911–938. doi: 10.1111/j.0022-3506.2004.00285.x
- Weibel, D., Wissmath, B., and Mast, F. (2011). Influence of spatial imagery on spatial presence and enjoyment assessed in different types of media. *Cyberpsychol. Behav. Soc. Netw.* 14, 607–612. doi: 10.1089/cyber.2010.0287
- Wilson, E., Park, D., Curtis, L., Cameron, K., Clayman, M., Makoul, G., et al. (2010). Media and Memory: the efficacy of video and print materials for promoting patient education about asthma. *Patient Educ. Couns.* 80, 393–398.
- Xu, J., Kemeny, S., Park, G., Frattali, C., and Braun, A. (2005). Language in context: emergent features of word, sentence, and narrative comprehension. *NeuroImage* 25, 1002–1015. doi: 10.1016/j.neuroimage.2004.12.013
- Yadav, A., Phillilden, K., and Dirkin, K. (2011). If a picture is worth a thousand words is video worth a million? Differences in affective and cognitive processing of video and text cases. *J. Comput. High. Educ.* 23, 15–37. doi: 10.1007/s12528-011-9042-y
- Yuill, N., and Oakhill, J. (1991). *Children's problems in text comprehension: An experimental investigation*. (Cambridge: Cambridge University Press).
- Zacks, J. M. (2015). *Flicker: Your brain on movies*. (Oxford: Oxford University Press).
- Zacks, J., Speer, N., and Reynolds, J. (2009). Segmentation in reading and film comprehension. *J. Exp. Psychol.* 307–327. doi: 10.1037/a0015305
- Zacks, J. M., and Tversky, B. (2001). Event structure in perception and conception. *Psychol. Bull.* 127, 3–21. doi: 10.1037/0033-2909.127.1.3
- Zeidman, P., Mullally, S., and Maguire, E. A. (2015). Constructing, perceiving and maintaining scenes: hippocampal activity and connectivity. *Cereb. Cortex* 25, 3836–3855. doi: 10.1093/cercor/bhu266
- Zwaan, R. A. (2009). Embodied cognition, perceptual symbols, and situation models. *Discourse Process.* 28, 81–88.
- Zwaan, R. A., and Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychol. Bull.* 123, 162–185. doi: 10.1037/0033-2909.123.2.162

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Jajdelska, Anderson, Butler, Fabb, Finnigan, Garwood, Kelly, Kirk, Kukkonen, Mullally and Schwan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.