

Reading Beyond the Lines: A Critical Review of Cognitive Approaches to Literary Interpretation
and Comprehension

Kathryn S. McCarthy

University of Illinois at Chicago, Department of Psychology

Accepted for Publication in *Scientific Study of Literature*, June 30, 2015

Please address correspondence to:

Kathryn S. McCarthy
Department of Psychology
1007 W. Harrison St. M/C 285
Chicago, IL 60607

E-mail: kmccart5@uic.edu

Acknowledgement

This manuscript was originally written as fulfillment of the preliminary examination for Ph.D. candidacy. The author wishes to thank her committee: Susan R. Goldman, Joseph P. Magliano, and Gary E. Raney as well as the reviewers for their suggestions and comments that have greatly strengthened the manuscript. This work was partially funded by Project READI (Reading, Evidence, Argumentation in Disciplinary Instruction), a multi-institution collaboration to improve complex comprehension of multiple forms of text in literature, history and science. It is supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305F100007 to University of Illinois at Chicago. The opinions expressed are those of the author and do not represent views of the Institute or the U.S. Department of Education.

Abstract

Reading literature requires not only understanding the literal meaning of the text, but also constructing a nonliteral interpretation of the text's deeper meaning (Goldman, 2004; Langer, 2010; Lee, 2007; 2011; Schraw, 1997) yet little is known about the psychological processes involved when interpretations are constructed. The current paper presents a review of the extant work from literary theory, empirical studies of literature, and research from more general cognitive text comprehension to explore the conditions under which literary interpretations are made and what this discipline-specific reading behavior can tell us about more general text comprehension.

Keywords: Cognition; Discourse Comprehension; Literary Interpretation; Cognitive Psychology; Inferences; Reader Goals

Reading Beyond the Lines: A Cognitive Approach to Literary Interpretation and Comprehension

Literacy in the 21st century requires not only the ability to read proficiently, but to do so in specialized content domains (Goldman, 2012; Goldman & Bisanz, 2002; Snow & Biancarosa, 2003). In literature, authors often write stories to have a message that speaks to a universal of the human condition (Goldman, 2004; Langer, 2010; Lee, 2007, 2010; Levine & Horton, 2013). Readers must have an understanding of the characters, setting, and events found in the text (a literal representation), but the successful reader must also construct a representation of the deeper meaning of the work (nonliteral representation). Literary theorists have written extensively on interpretation, yet there has been far fewer empirical studies conducted to understand the cognitive processes that underlie how real readers with different backgrounds, skills, interests, and abilities interpret literary works. Consequently, further discussion must occur across fields regarding theories of literary comprehension and about potential directions in empirical literary research. This paper, heavily grounded in cognitive psychology, presents a comprehensive review of literary theory, empirical studies of literature, and more general text comprehension as a means of fueling this discussion..

The Role of Interpretation in Literary Works

The majority of text comprehension research has attended to the generation of inferences in expository texts (McNamara & Magliano, 2009). One reason researchers have shied away from literary text and interpretation is because these stimuli and behaviors are “inconsiderate” (Zwaan, 1993) in the sense that they are difficult to manipulate and to assess in a controlled, empirical way. This ambiguity has resulted in a relatively unsystematic investigation of literature and literary reasoning. Before discussing the research, I will first discuss the constructs of

interpretation and authentic literary text to (a) establish the parameters for this review and (b) suggest definitions that may help to reconcile terms in the extant work.

Studies exploring interpretations have used a multitude of terms, including *point* (Dorfman & Brewer, 1994), *authorial intention* (Gibbs, 2001), *theme* (Kurtz & Schober, 2001; Graesser, Pomeroy, & Craig, 2002; Magliano & Graesser, 1991; Zhang & Hoosain, 2005), *interpretive inference* (McCarthy & Goldman, in press), *thematic generalization* (Peskin, 2007), *subtext* (Schraw, 1997), *significance* (Peskin, 2007), and *signification* (Rabinowitz, 1987). While these terms have nuanced differences, they all share the idea that they reflect some kind of “deeper meaning.” I refer to such interpretations as *nonliteral interpretations*: a representation of the text or part of the text “without reference to specific story elements” (Kurtz & Schoeber, 2001, p. 141). A nonliteral interpretation is made when the reader generates inferences that connect information in the text with ideas about the world beyond the story, such as identifying an object in the text to have a symbolic meaning, or the text as a whole, such as an understanding of how the story speaks to the state of the real world in which we live. A central focus of text comprehension has been the generation of inferences – connecting or elaborating on the text (Graesser, Singer, & Trabasso, 1994; Zwaan & Radvansky, 1998). Nonliteral interpretations can be thought of as higher-order inferences, as they require the bringing together of multiple pieces of information from throughout the text, as well as information from prior knowledge (Magliano, Baggett, & Graesser, 1996; Magliano & Graesser, 1991). Thus, a focus of this work is in understanding when and how interpretive inferences are generated.

Work in discourse comprehension and the generation of inferences has been extensive, but, as previously mentioned, much of the work in discourse comprehension research has avoided the use of literary works. Instead, most research in discourse comprehension has been

conducted using lab-designed expository and simple narrative texts (McNamara & Magliano, 2009). While these “textoids” afford excellent control and internal validity, one consistent criticism is that they do not reflect the complex nature and purpose of authentic literature (Claassen, 2012; Graesser, Millis, & Zwaan, 1997; Kreuz & Roberts, 1993; Magliano, Baggett, & Graesser, 1996; Miall & Kuiken, 1994; Peskin, 1998; Rapp, Komeda, & Hinze, 2011; Zwaan, 1993). Consequently, it may be inappropriate to generalize the findings of these studies and to assume that the models of comprehension that have emerged from them are able to capture the kinds of processes and behaviors involved in the reading of authentic literary works.

But what makes authentic literature different from these lab-designed texts or even distinct from other kinds of texts? Identifying authentic texts is relatively straightforward. Determining what qualifies as text as literary or nonliterary has proven to be a complicated categorization. There is general agreement that literature comprises its own genre, yet there is little agreement on exactly what qualifies a text¹ as literary or nonliterary (Dixon & Bortolussi, 2011). Studies suggests that both computational models and real readers can distinguish literary texts from nonliterary ones (McCarthy, Myers, Briner, Graesser, & McNamara, 2009), that readers read literary texts differently than nonliterary texts (Hanauer, 1998; Zeitz, 1994), and that literary texts have differential effects on other cognitive processes (Comer Kidd & Castano, 2013; Dixon, Bortolussi, Twilley, & Leung, 1993; Djikic, Oatley & Moldoveanu, 2013), and yet, exactly what makes these literary texts different from nonliterary texts has not been well-defined.

The studies cited in this review have used “authentic” works: published stories, novels, or

¹ It should be noted that “literature” can also be used to describe other mediums, such as film, theatre, or television (Levin & Horton, 2013), but for the purposes of this paper, the focus will be on text-based literature, including folk tales, fables, novels, short stories, and poetry.

poems not written with the purpose of being experimental stimuli. By using these authentic works, these studies have ecological validity and may shed light upon reading behaviors that are not typically demonstrated when using more simplistic texts. It is likely that interpretive inferences are one of these behaviors. In most studies using literature (indeed, the majority of the works cited in the remainder of the paper), texts are chosen because they have won literary prizes or because they are typical in English/Language Arts classrooms. The way in which researchers typically study literary reading is by taking texts deemed literary by experts and submitting them to various analyses of complexity (e.g. Louwerse, Benesh, & Zhang, 2008; McCarthy, Myers, Briner, Graesser, & McNamara, 2009) or comparing the comprehension and reasoning of these texts with texts that are not considered literary, such as expository texts (Hanauer, 1998; Zeitz, 1994), or popular fiction that has not been given literary awards (Comer Kidd & Castano, 2013; Dixon et al., 1993).

Looking across literary studies, the implicit agreed upon definition of “literature” is a text that literary experts agree possesses “literary merit” (Dixon, Bortolussi, Twilley, & Leung, 1993). However, the criterion for “literary merit” is unclear. The most articulate definition in the research comes from Schraw (1997). He defines a literary texts as narratives that are “richly symbolic and include both an interpretable surface meaning and one or more coherent subtexts (i.e., implicit thematic interpretations that run parallel to the explicit surface-level meaning of the text)” (p. 436). While it is questionable whether or not interpretability is the definitive feature that makes a work literary, this definition is implicitly supported by the texts that have been selected for research, suggesting that interpretability might be one quality of literary merit and is an important aspect of literary reading.

But what quality makes a text “richly symbolic” or “interpretable”? In literary theory,

some have argued that this interpretability is a function of formal features inherent in the text (Miall & Kuiken, 1994), whereas others have argued that interpretability is solely within the reader (Culler, 1994; Fish, 1980). In prior work, we have proposed that a more fruitful way of thinking about literature is considering “literary” as a means of describing a way of reading a text, rather than a particular quality of the text in itself (Goldman, McCarthy, & Burkett, 2015; Vipond & Hunt, 1984). That is to say that there may be aspects of the text, reader, and situation that determine whether a text is interpreted or not. Thus, in order to better understand literary interpretation, it is of value to explore these factors and their individual and interactive effects.

Factors That Influence Literary Text Comprehension

Both aspects of the text and aspects of the reader can be implicated in the generation of nonliteral interpretations. These factors, undoubtedly, work together in the construction of meaning, but for purposes of structure and clarity, I will discuss them independently before addressing interactive effects.

Aspects of the Text

The formalist approach to literary comprehension suggests that inherent features of a text activate interpretive responses (Miall & Kuiken, 1998). While the evidence below indicates that features of the text facilitate the construction of interpretations, it also suggests that the use of textual features is an effortful and strategic process.

Genre. Genre can be defined as groups of texts that share "prototypic language patterns, a standard way of reading and particular discursive functions" (Hanauer, 1998, p. 63). Research reveals that the genre of a text can be identified by textual features. The genre of a text can be rapidly and accurately identified by computational models that use frequency of specific words and sentence structures (Louwerse, Benesh, & Zhang, 2008; McCarthy, Myers, Briner, Graesser,

& McNamara, 2009), as well as actual readers working without explicit inclusion criteria (McCarthy, Myers, Briner, Graesser, & McNamara, 2009; Zwaan, 1993). This indicates that readers have some sense of what constitutes a particular genre, and, consequently, the way the text should be read. Hanauer (1998) proposed a genre-specificity hypothesis of reading that suggests different genres of text are read differently. This would predict that literary texts are read differently than those texts that are of a different genre. Indeed, Hanauer showed that readers have better recall, slower reading times, and lower judgments of comprehension for poetry than encyclopedia entries. Additional work shows that literary experts produce interpretations when reading literary texts, but not non-literary expository texts (Zeitz, 1994). This suggests that inherent textual features do, at least partially, dictate how a text is processed. Specifically, it appears that the genre of literature demands attention to the surface language of the text.

Dixon, Bortolussi, Twilley, and Leung (1993) were able to show not only that literary texts were read differently than nonliterary texts, but also that the text features affected appreciation for a literary work. They proposed that one unique text effect (a particular mental state that emerges from reading) of literature is an increased appreciation for the work over time and upon reflection. Thus, literary texts should show greater enjoyment after a second reading, whereas non-literary texts should not. To test this, participants read the literary detective story, *Emma Zunz*, or a true crime narrative that possessed similar plot points. They rated their enjoyment after an initial reading and a re-reading. As predicted, the literary text yielded greater *depth of appreciation* (an improvement in enjoyment rating from initial reading to re-reading) than the non-literary text. In a second experiment, the researchers manipulated the presence of the text feature *narrator ambiguity* (many questions and words indicating uncertainty). These

ambiguous phrases were replaced with more certain statements. Without the narrator ambiguity in the text, participants no longer rated the re-reading as more enjoyable than the initial reading. This research suggests that textual features can create effects that are unique to the genre of literature. However, it is not evident if greater depth of appreciation reflects any differences in processing or the likelihood of constructing nonliteral interpretations.

Foregrounding. Harker (1996) argued that in order for a reader to gain perspective on the world (as is the purpose of reading literature), the text must force the reader to experience everyday events or objects in a new way. Literary theorists call this *defamiliarization*, and suggest that it is “central to the literary experience” and the “hallmark of literariness” (Miall & Kuiken, 1994a, p. 337). They propose that defamiliarization is achieved through literary *foregrounding*— use of variations and devices in language to make some parts of the text more important or striking (in the foreground) than other parts (the background) (Miall & Kuiken, 1994a, 1994b).

To test this notion, Hoffstaedter (1987) asked both poetry experts² (literature students) and non-experts (engineering students) to go through 24 poems and underline the parts of the texts that made them poetic. Both groups underlined sentences that were highly metaphoric and stylistically irregular from normal language. While the non-experts underlined fewer sentences and were less confident in their ratings, the sentences they selected as poetic were highly correlated with those selected by the experts. Similarly, Miall and Kuiken (1994b) had participants rate the “strikingness” of sentences in passages of literature. They found sentences that possessed more foregrounded features (based on phonetic, grammatical, and semantic variations) were rated as more striking than background sentences, suggesting that there are

² In studies of literary reading, the distinction between experts and novices refers to amount of formal literary training, not general reading ability.

concrete aspects of text (stylistic variations) that make a text more or less likely to be defamiliarized. However, these studies do not speak to if these features are responsible for the generation of interpretive inferences.

If defamiliarization is, indeed, a critical feature of literature, then it is likely involved in the production of nonliteral interpretations. Indeed, Miall and Kuiken (1994a, 1994b, 1998), suggest that defamiliarization forces an affective response, and the act of refamiliarizing the content requires the activation of similarly emotionally-valenced prior knowledge, which helps the reader to find an appropriate interpretation. It is implied that this affective information is activated automatically. Miall and Kuiken's data do show evidence of a relationship between foregrounding and affect in literary reading, but their research failed to test their claims about affect yielding interpretation.

Recent work by Levine and Horton (2013) more directly investigated affect and nonliteral interpretation. The study compared a classroom intervention focused on affective evaluation to a control condition that used a traditional unit on identifying symbolism in poetry and short stories that did not mention affect. The intervention condition also taught symbolism, but encouraged students to practice evaluating how a part of the text made them feel, and how to identify what aspects of the text had lead them to that affective response. The students composed essays about the same poem at pre-test and post-test. Statements in the essays were coded as *literal descriptive* (paraphrase of the text), *literal interpretive* (connecting parts of the text via inferences), and, of particular interest, *thematic* (discusses meaning that goes beyond the story world). The classes showed no differences during pre-test, and both showed interpretive gains at post-test. However, the intervention condition showed a greater gain, such that students in the intervention produced more thematic statements than the control. Additionally, only the

intervention group showed a transfer benefit to a second poem. The findings suggest a connection between the specific features of a text, an affective response, and the construction of nonliteral interpretations. However, it must also be said that this process appears to be less automatic than proposed by Miall and Kuiken. The text may automatically activate affective response, but there is little empirical evidence for this and more evidence supporting the idea that readers must be explicitly taught to assess these emotions and relate them to interpretation. This emphasizes the importance of factors outside of the text itself.

Aspects of the Reader

The *conventionalist* perspective of literary criticism suggests that interpretation is not a function of the text, but, instead, a construction of the reader (Culler, 1994; Fish, 1980). While the data from the previous section suggest that a staunch conventionalist perspective (which would deny any effect of text) is implausible, the following research suggests that the reader also plays an integral part in how literary texts are read. Aspects of the reader can further be broken into those that are relatively static across contexts (i.e., prior knowledge, epistemic beliefs, interest in the domain or topic) and those are more context-dependent (i.e., task instructions, situational interest in a particular text).

Reader expectations. As mentioned earlier, Hanauer (1998) provided evidence for the genre-specific reading hypothesis – that different genres afford different reading processes. The use of authentic poetry and encyclopedic excerpts provided ecological validity, and the use of multiple texts ensured that the effect was not due to the content in a specific passage. However, it could be that the differences in reading times and comprehension ratings were not due to the texts, but rather the reader's expectations for how different texts should be read. In a study on expectations, the same text was given to all participants, but participants received different

information about the genre of the text. The participants who believed they were reading a literary short story performed better on surface level questions, whereas participants who thought they were reading a newspaper article had better situation model representations (Zwaan, 1993, 1994). These data are consistent with Hanauer's work in showing that literary readers have better memory for surface features of the text, but indicate an expectation effect, such that the way a text is processed is a function, not only of the text itself, but of the reader's expectations for the text.

Researchers propose that these effects occur because readers' expectations for the kind of text they are reading enact different schemas for meaning-making. These schemas help to determine what information in the text is important and to activate relevant knowledge. Zwaan (1993) referred to these schemas as *cognitive control systems*, whereas Vipond and Hunt (1984) discussed literary texts as *point-driven*, as opposed to expository texts that are *information-driven* and basic narratives that are *story-driven*. The research indicates that a literary schema focuses the reader on attending to surface code information, as well as the holistic meaning of the text, rather than textbase or plot-level situation model information. This is likely because readers are using the text as clues to construct the underlying meaning of the work.

Genre expectation manipulations have also been used to explore the generation of nonliteral interpretations in different kinds of literary texts. A think-aloud study by Peskin (2007) found that a sentence presented in a traditional poem format (visually similar to a haiku) yielded more references to literary conventions and interpretive operations (such as the recognition of multiple meanings, attention to specific language, and the construction of abstract meaning) than the same sentence presented in prose form. Peskin also replicated the genre expectation effect with two short, authentic poems. It is important to note, however, that these students were high-

level students enrolled in an English course, suggesting that, while certainly not experts, they possess domain-specific knowledge, and that may have been particularly salient in the context under which they read these texts. This sheds light upon the potential role of training within the domain of literature as a factor in literary interpretation.

Expertise. Much of what is known about literary comprehension has come from studies that compared the performance of novices and experts. In these comparison studies, the non-experts are mostly middle and high school students or early undergraduates with very basic declarative knowledge about the domain, whereas the experts range from graduate students through senior English faculty who possess extensive knowledge and experience with conventions in the domain (Warren, 2011). It is worth noting that, as is true in studies of expertise in other domains, expertise is often confounded with other factors, such as age and dispositional factors that may lead to a selection bias in the sample. Several investigations have attempted to isolate training from these other factors and have shown that formal training affords different kinds of processing and reasoning about literary works than those with similar levels of education (Zeitz, 1994) or formal training in a different domain or genre (Bortolussi & Dixon, 1996). These studies indicate that readers develop domain-specific schemas that help them to construct interpretations of the text.

Graves and Frederiksen (1991) asked two English professors and six undergraduates from the same university to read an excerpt from *The Color Purple*. The resulting think-aloud data were subjected to a discursive pattern analysis that categorized statements as *linguistic*, *propositional*, or *conceptual*. The conceptual statements were further characterized as *text-based* or *derived* descriptions, where the latter reflected “high-level” inferences drawn from the text, as well as bringing in prior knowledge about the world at large. This distinction from text-based

descriptions suggests that derived representations can be thought of as nonliteral interpretations. The analysis revealed that the novices produced mostly paraphrases of the events of the story (propositional statements), whereas the experts' think-alouds yielded mostly linguistic and conceptual statements. Looking specifically at derived descriptions, it was found that the novices produced significantly more text-based descriptions than the experts, and the experts produced significantly more derived descriptions than the novices. An additional study by Peskin (1998) indicates a similar expertise effect with authentic poetry.

Zeitz (1994) included two novice groups in her think-aloud study: high school students and engineering experts. These engineering experts possessed the same level of education as the literary experts (graduate training), but different domain specialization. All three groups (high school novices, engineering experts, and literary experts) read a poem, short story, and expository science text. Replicating expertise research in other domains (Chase & Simon, 1973; Chi, Feltovich, & Glaser, 1981; Means & Voss, 1985), experts showed a recall advantage in their domain. For the poem and short story, literary experts yielded more interpretation than the other groups. Interestingly, there were no differences in the amount of interpretations made across the three groups of readers for the science text; the engineering experts did not show an interpretive advantage, nor did the literary experts show a domain-general tendency to interpret more than the others.

The construction of interpretive inferences seems to be scant amongst non-experts. Vipond and Hunt (1984) asked more than 100 undergraduates to answer questions about a short story. Some of these questions were intentionally leading to encourage readers to think about a possible meaning behind the text. Despite this, less than 5% of the students considered the possibility of an interpretive message. Similarly, Claassen (2012) found that novice readers

activated author-relevant information during reading, but in a think-aloud protocol, less than 5% of their statements were related to author intent, suggesting that even though the information is activated, it is not being strategically selected for interpretation.

Across these studies, the literature experts not only produced more interpretations, but also drew from their extensive knowledge of conventions and common themes to contextualize and connect the current text with other works and ideas. Such information manifested itself in more time spent with the text during reading (Peskin, 1998) and in more sophisticated arguments in essays (Zeitz, 1994). Given that authors often rely upon familiar devices, tropes, and conventions, expert readers are likely to have developed relevant schemas that help them recognize and integrate information into a complex representation of the text.

For example, one of the experts in Graves and Frederiksen (1991) recognized the irregular syntax of a passage as the “kind of social dialect we associate with Southern black perhaps” (p. 19), and quickly pointed out evidence of white-and-black conflict that is prevalent in literary works of this nature. The novices, on the other hand, were unfamiliar with such written representations of dialect. Thus, they merely noted that the text was irregular and were unable to use the information to help them contextualize the story. This indicates that the experts are able to construct more inferences about literary texts because they have more relevant prior knowledge. This is consistent with more general text comprehension that suggests inference generation is affected by “reading skill, understanding of the demands of the reading task, working memory capacity, and background knowledge relevant to the text topics” (van den Broek, Lorch, Linderholm, & Gustafson, 2001, p. 1081).

Literary epistemology. It may also be that experts possess not only more content knowledge, but that they also have different beliefs about the nature of literature that assist them

in constructing interpretations. Both general and topic-specific epistemologies have been shown to affect comprehension in various domains (e.g., Bråten & Strømsø, 2006; Schommer, 1990; Stathopoulou & Vosniadou, 2007). For literature-specific epistemology, the Reader Belief Inventory (RBI) assessed *transmission* and *transaction* beliefs (Schraw, 2000; Schraw & Bruning, 1996). A transmission belief means that the reader believes that he or she must reconstruct *the* author's message (implying that there is one correct interpretation of the text). A transactional belief is constructive; the reader believes it is his or her job to interact with the author to create a meaning that might be idiosyncratic to that particular reader in that moment. Schraw and Bruning found that transmission beliefs were unrelated to interpretations, but an endorsement of transaction beliefs was highly related to the production of thematic, critical, and personal connection statements when reading and reasoning about literature and attention to multiple or ambiguous meanings, acknowledgement of author as an entity, and attention to and evaluation of the author's linguistic choices.

Possibility of multiple meanings. One difficulty in literary interpretation is that it is inappropriate to say that one interpretation is correct, or that one interpretation is better than another (Hillocks & Ludlow, 1984; Kurtz & Schoeber, 2001). This makes the assessment of literary reading tasks difficult for cognitive researchers. To attenuate this problem, researchers have employed "simple" literary texts with canonical interpretations, but even these texts appear to not have a decisive right or wrong interpretation. One way in which researchers have addressed this is by using "canonical" interpretation or interpretations that have been provided by a majority of readers. Abrahamson and Sprouse (1995) asked young students to listen to fables, select a moral from a list, and then to explain why they selected that moral. Though a throwaway point in their scoring section, Abrahamson and Sprouse reveal that two texts from

their stimuli set were excluded from analysis because consulting literary experts did not select the same interpretation of the story. This does not mean that one expert was right and one expert was wrong, rather a literary work could have a variety of possible meanings. Indeed, it is suggested that “exploring the possibilities” of meaning is a central activity in successful literary reading (Langer, 2010). This means not only that interpretations may differ across readers, but also that the same reader may entertain multiple interpretations of a single literary work. The willingness to accept multiple possible meanings suggests that the reader takes a transaction belief. Indeed, evidence showed that experts searched for and explored multiple interpretive meanings of the same passage more frequently than novice readers (Peskin, 1998). This indicates that experts understand that reading literature invites the reader to bring in his or her own knowledge and experience into the meaning-making process, whereas novices are strictly seeking to find the “right” answer, which limits the kinds of information to which they attend (Graves & Frederiksen, 1991; Levine & Horton, 2013). Exploring different interpretations also shows that the readers are holding off on making decisions about the text’s overall message, suggesting that readers worry less about maintaining local coherence in favor of establishing global coherence (Vipond & Hunt, 1984).

Metalinguistic awareness. Experts seem to possess and utilize metalinguistic awareness when reading literary texts. In most kinds of reading, the surface code of the text plays a minor role compared to the textbase and situation model representations (McNamara & Magliano, 2009). Given the emphasis on point-over-plot, one might think that the surface code of the text is even *less* important in literary reading. However, as has been mentioned, literary readers seem to devote more attention to the surface code of the text than they do with non-literary texts. While this seems contradictory, it may be that readers rely on the surface code as a means of assessing

the author's aesthetic choices. Because they have a rich knowledge-base of literary devices, conventions, and common themes, experts produce a potential interpretation early in their thinking about the text and then shift their attention to understanding how the author used the language to produce such a meaning (Graves & Frederiksen, 1991; Peskin, 1998; Rabinowitz, 1987). This indicates that within their community of practice, they must not only produce a nonliteral interpretation, but also support that interpretation with evidence that comes from the foregrounded textual features. A taxonomy tested by Hillocks and Ludlow (1984) suggests that using the structure or linguistic aspects of a text, as evidence to justify a particular interpretation of the text, is the most complex and sophisticated reasoning performed when reading fiction (the authors do not make distinction between general fiction and literature, but their examples are all drawn from what are taught as literary works in English/Language Arts classrooms). Given its taxonomic nature, readers need to have a sense of a possible interpretation before they can make sense of metalinguistic features and these textual aspects are used as justification and reasoning. As novice readers show little evidence of interpretation, it is unsurprising that they show little metalinguistic awareness.

Reader goals. Goals guide the manner in which the reader makes meaning from the text (Graesser, Singer, & Trabasso, 1994; van den Broek, Lorch, Linderholm, & Gustafson, 2001). In the text comprehension field at large, *goal* often implies a particular task instruction: however, recent research proposes that goals should be considered an interaction between both external given intentions and internal personal intentions (Kendeou, Bohn-Gettler, & Fulton, 2011; McCrudden, Magliano, & Schraw, 2010).

Internal intentions. One commonly studied internal intention is interest. Interest can be categorized as personal or situational. Personal interest (or individual interest) is relatively static,

such as preference for one genre more than another, or books by a certain author or on a certain topic. Personal interest has been shown to relate to prior knowledge (Tobias, 1994). Thus, it is unsurprising that literary experts have more interest, and, thus, more internal motivation to engage in literary reading tasks. Situational interest, as its name suggests, varies from situation to situation and text to text, and contributes unique variance to comprehension above and beyond personal interest. Schraw (1997) asked undergraduate students to read the story *The Book of Sands* and rate their interest via two questionnaires: one regarding personal interest, and one regarding situational interest. The participants then performed a recognition task of 20 factual and inferential multiple-choice questions, and wrote an essay about *what the story meant and what kind of personal thoughts and feelings it evoked*. These essays were coded in three ways. First, statements were coded as *text retellings, elaborations, interpretations of symbols, or thematic inferences*. Then, the number of personal responses were tallied and categorized into *affective responses, thought provoking aspects of the text, empathy with the text events and characters, and relating experiences in the text to one's own life*. Finally, the essays were scored on a scale from 1-6, reflecting the sophistication of the reader's interpretation of the text as a whole.

The recognition test data revealed that ratings of personal interest had no effect on understanding the literal meaning of the text; all participants performed well, regardless of condition. In contrast, situational interest predicted a participant's holistic interpretation score above and beyond personal interest, and the contribution of specific textual features, indicating that situational interest leads to richer interpretive representations of the text. Breaking the participants into high, average, and low situational interest also revealed that participants with low interest produced significantly fewer personal connections than those with average or high

situational interest. Additionally, those with high situational interest produced significantly more thematic inferences than those in the average or low interest groups. These data suggest that interest in a given text affects the kinds of inferences that are constructed, and the kinds of prior/personal knowledge brought into a reader's representation.

Elfenbein (2006) analyzed four different expert criticisms of the same poem by Robert Browning. He explored how the critics' reading goals influenced the way he or she understood the poem. The critics had self-motivated reasons for approaching the texts as they did. The examination showed that some critics attended mostly to trying to understand what the poem locally, whereas others attempted to understand the poem more holistically. Those critics who focused upon making sense of the poem at a semantic level seemed to express confusion and distaste for the poem, whereas those who attended to the feelings of the text as a whole, seemed to better appreciate its aesthetic merit and discussed the poem's overall meaning. This archival analysis not only provides support for the notion that taking a global, point-driven stance supports the construction of nonliteral interpretations, but also indicates that a reader's self-defined goal or intention affects the way a text is processed and, subsequently, understood.

External intentions. As the expert-novice research revealed, novice readers do not spontaneously engage in nonliteral interpretation (Graves & Frederiksen, 1991; Peskin, 1998; Zeitz, 1994) and, yet, when prompted to do so, they have little difficulty in *identifying* the moral of a story (Abrahamsen & Sprouse, 1995; Johnson & Goldman, 1987; Narvaez, Bentley, & Samuels, 1998). Thus, it is not that novices are incapable of interpretation, but, rather, they must be explicitly instructed to discuss it. This suggests that the particular task presented to the reader affects the engagement in interpretation. These findings are consistent with research in other domains (e.g., Kotovsky, Hayes, & Simon, 1985; Wiley & Voss, 1999). In a study that directly

tested the effect of task instruction on the construction of nonliteral interpretations in literature, undergraduates were asked to read a story (*Harrison Bergeron* by Kurt Vonnegut or *The Elephant* by Slowomir Mrozek) and to respond to an essay prompt. The prompt *What happened in this story?* yielded essays containing close-to-text *paraphrases*, similar to what was found in think-aloud protocols in Graves and Frederiksen (1991) and Zeitz (1994). However, when the prompt directed the readers towards an interpretation, the essays showed a significantly higher proportion of *interpretive inferences* that reflected either discussion of symbolic meanings or the moral or point of the story (McCarthy & Goldman, in press). This more specific instruction led the readers to construct more interpretive inferences. Thus, the reader's goal affected how the work was comprehended. One possible explanation for why interpretive statements only emerge in the essays under explicit instruction is that interpretations are generated automatically, but novice readers ignore them because they deem that information irrelevant to the essay-writing task. However, it seems if this were the case, such interpretations would have emerged in novice reader think-aloud data. This suggests that interpretations are not generated automatically, but as a strategic process in response to a particular reading goal.

McCrudden, Magliano, and Schraw (2011) combined qualitative and quantitative methods to explore the effects of both internal and external intentions on reading. In the study, all participants were given the same four expository texts about different countries. Two groups were asked to read the texts and to take a particular perspective on the pros and cons of living in one of the countries (Honduras or Pitcairn). The control group was asked simply to read for understanding. Their recall data replicated previous perspective-taking findings (Anderson & Pichert, 1978) in that students recalled more information about the country they were asked to focus on. Interestingly, post-test interviews found that some of the readers took a *narrowing*

approach, where they attended to only the information for the country they were asked to focus on, and skipped information about the other countries, while other students took a *broadening* approach – attending to information about all of the countries in order to compare and contrast. These interviews suggest that the readers consciously attended to their given external goal, but that readers enacted different internal strategies to realize this goal. One might predict that a similar interaction of internal and external intentions might affect how literary texts are read, and if nonliteral interpretations will be constructed. However, such a hypothesis has not been tested empirically.

The adoption of a *point-driven* stance (Vipond & Hunt, 1984) can be thought of as having the reading goal of constructing the text's underlying meaning. While Vipond and Hunt (1984) describe the kinds of behaviors each stance might include, they do not specify the mechanism through which a particular reading goal is enacted. Van den Broek and colleagues (2001) have shown that reading goals control the kinds of inferences that are generated during reading narrative texts. Their work shows that those reading to study made more inferences than those who were reading for leisure (van den Broek, Lorch, Linderholm, & Gustafson, 2001). They believe readers in the study condition had a more rigorous criterion for which to evaluate their understanding of the text and constructed a more thorough representation to meet this criterion. From these findings, they propose that a reader's goals result in activation of different *standards of coherence*. These standards establish what qualifies as successful comprehension, and these standards guide what inferences need to be constructed to make sense of the text. Transferring these findings to literature, readers, presumably, produce elaborate inferences as a means of meeting the goal of understanding the author's message. These inferences result in a richer representation upon which they can realize an interpretation.

Text and Reader Interactivity

This comprehensive review has identified several factors related to literary reading and the production of nonliteral interpretations: textual features of stylistic variation and rules of notice, prior knowledge from expertise and expectations, and reader goals. It has suggested that both text and reader play roles in successful literary reading. Additionally, findings suggest that manipulations of reader aspects are affected by the nature of the text, and manipulations of texts are affected by the kind of reader in question. The following studies provide evidence that these two general aspects interact to determine the kind of representation that is constructed.

In his work on situational interest, Schraw (1997) explored how text features were related to literary interpretation. He showed that ratings of suspense, text coherence, and thematic complexity were all correlated with quality of holistic interpretation. A hierarchical regression indicated that both thematic complexity ($R^2 = .082$) and text coherence ($R^2 \text{ change} = .022$) predicted the quality of the reader's holistic interpretation. However, when situational interest was entered in the first step of the model, only thematic complexity yielded a significant contribution to the model, and its effect was much attenuated ($R^2 \text{ change} = .020$) (Schraw, 1997). These data suggest the relationship between text characteristics and quality of holistic interpretations is partially mediated by situational interest. As this analysis was conducted on a single text, it is unclear whether these particular text features are the only ones that contribute to situational interest, or if this result is text-specific. Replication of such effects with other literary texts would be of value. Yet, what is apparent is that the relationship between text and reader is complex. In another study, it was shown that the conventionality of a text also influences a reader's interest, which then affects the production of an interpretation. Researchers manipulated fables to follow conventional storylines (fable schemas) or end in a schema-inconsistent way.

The versions of the stories that more closely followed traditional schemas were rated as more typical and more interesting. These interest ratings were correlated with the readers' ability to select an appropriate interpretation (Dorfman & Brewer, 1994).

It has also been shown that experience and expertise affects whether or not the stylistic variations that make up the textual features of literature are noticed. Dixon, Bortolussi, Twilley, and Leung (1993) found that the relationship between the literary features and literary effects was moderated by familiarity with literature. Specifically, the removal of narrator ambiguity from a text only affected the depth of appreciation in readers who more frequently read literary works. The presence or absence of literary features did not alter the less-familiar readers' depth of appreciation scores. In other words, these readers did not possess enough prior knowledge to recognize literary qualities, and the lack of awareness of these features resulted in the lack of a literary effect. By extension, it seems that these less familiar readers would be less likely to produce nonliteral interpretations if they are not attuned to these literary features, though this has yet to be tested.

Furthering the importance of prior knowledge on the comprehension of literary textual features, evidence suggests experts and novices differentially understand foregrounded text. When faced with complex foregrounded passages, experts recognize the purpose, and search for interpretive meanings. In contrast, novices tend to disengage from the task. Graves and Frederiksen (1991) noted, "When confronted with ambiguity or lack of clarity, students reacted to it as a reflection of their own inadequacy. It would seem that not only do students work more from the surface information, but that they are so dependent on it that when it doesn't make sense, they have no recourse but to flounder" (p. 21). These observations are echoed by the observations of non-experts in Peskin (1998) and Vipond and Hunt (1984). As Miall and Kuiken

(1994a, 1994b) note, authentic literary work is heavily influenced by, if not defined by, its use of foregrounding and defamiliarization. Syntactic complexity is beneficial for an experienced literary reader because it not only clues the reader into activating the appropriate reading goal, but the more complex the text is, the more linguistic features the reader can use to help construct the interpretation of the text. However, the aspects that make a text good literary reading are what novices struggle with most. It may be that the variance in epistemic beliefs between experts and novices might account for these differences. However, there has been no research on the change of epistemic beliefs across the trajectory of literary expertise, so how readers learn to overcome this obstacle is unclear.

The complex interactions of text, tasks, and experience are highlighted in a study by Mason, Scirica, and Salvi (2006). In this study, 52 students responded to the Reader Belief Inventory (Schraw, 2000; Schraw & Bruning, 1996), and were categorized into one of four epistemological groups: 1) low transmission, low transaction (LL), 2) high transmission, low transaction (HL), 3) low transmission, high transaction (LH), and 4) high transmission, high transaction (HH). They predicted a positive relationship between transaction beliefs and interpretive behavior. These same students read an excerpt from *The Little Prince*, answered text-based multiple-choice comprehension questions, and were given either (a) generic instructions that asked them to write their comments or (b) specific instructions that asked them to write their comments. Replicating Schraw (2000) and Schraw and Bruning (1996), they found no differences across the groups in their performance on the multiple choice comprehension questions, but a significant difference in the amount of interpretations, such that those in the HL group produced more personal responses than those in the LH group. However, the analyses revealed no main effect of task instruction, nor a task instruction by beliefs interaction. These

same patterns emerged when the essays were coded with a holistic 6-point sophistication of interpretation scale. Looking at the mean differences qualitatively suggests trends towards a significant interaction, but the lack of statistical effects is likely driven by the small sample size and uneven cells in the belief groups. To remedy this, Mason and colleagues used a larger sample, a different short story, and split the middle and high school students to test for grade-level differences in a second experiment. This complex design revealed main effects for beliefs, task instructions, and grade level in the number of interpretive responses and the holistic interpretation score. Only the task instruction by grade level interaction reached significance.

In the older students, the inclusion of specific instructions provided only a modest benefit in interpretive responses. In the younger students, the specific instructions yielded a large benefit. While there were no direct cross-experiment comparisons, the mixed results may be a function of the texts' complexity or affordances. The authors noted that the second text, *The Garden of the Obstinate Cats*, seemed to be a more difficult text for the students. Indeed, Experiment 2 yielded differences in multiple-choice comprehension question performance that were not seen in Experiment 1. This study suggests an extremely complex relationship between text characteristics, task instructions and reading goals, epistemological beliefs, and experience.

Summary

Before looking specifically at text comprehension models, perhaps it is best to summarize and highlight the themes that have emerged from the previous sections. The data indicate that the construction of a nonliteral interpretation requires a shift from a literal, story-driven stance to an interpretive, point-driven stance. This interpretive stance puts more emphasis on global coherence, which requires the integration of multiple parts of the text, as well the activation of prior knowledge into the reader's representation.

When real people read real texts, the adoption of a point-driven, interpretive stance, and the subsequent generation of appropriate interpretive inferences, appears to be a strategic and effortful process. Readers must move beyond a literal stance to an interpretive stance. This leap from one to the next is guided by textual and extratextual features that may be more or less relevant, given a particular reader who is reading a particular text for a particular purpose. Because of their knowledge of, and familiarity with the domain, experts more easily adopt the appropriate reading strategy. They are sensitive to rhetorical devices, rules of notice, and foregrounding in the text. These textual features act as signals to switch to a point-driven stance. Experts also possess more relevant content knowledge, which can then help them to construct an appropriate nonliteral interpretation. Further, once the reader has constructed a nonliteral interpretation, he or she can then turn to understanding how the linguistic aspects of the text guided the reader to that interpretation. While minimal empirical data speaks to this aspect of the process, identifying such linguistic aspects, and connecting them to particular themes and interpretations, likely comes from the activation of prior knowledge.

Novice readers, on the other hand, are less likely to adopt an interpretive stance, and even though they notice stylistic variation, they struggle with making sense of it. If novice readers are able to construct a satisfactory plot-level representation, they may be oblivious to the foregrounding of the text, and not move to a global, point-driven stance. Given no instruction, it seems readers initially approach a text with a generic reading strategy, as this strategy is most familiar. In theory, a heavily-foregrounded text that does not sustain local coherence should drive the novice to switch strategies, but, in practice, it seems that the novice reader merely gives up. It seems the only effective way to get novice readers to adopt an interpretive reading strategy is to explicitly ask them to do so. Further, even if the novice reader employs such a strategy, the

construction of a nonliteral interpretation is dependent upon how much related knowledge he or she can bring to the text.

Thus, this review suggests an integrated view of literary interpretation, relying on aspects of both the reader and the text. In short, a reader may approach a text with the intention of constructing a deeper meaning, but if the text does not have foregrounding and literary stylistic variation, it may be difficult to produce an abstract interpretation. Conversely, a text with abundant rules of notice and foregrounding will not yield interpretation if the reader (a) does not have a literary reading goal in mind or (b) does not possess enough domain-related prior knowledge to construct an interpretation.

Considerations and Future Directions

As education and achievement standards put an increased emphasis on not only reading, but reading for understanding (Council of Chief State School Officers [CCSSO], 2010; National Assessment of Educational Progress (NAEP) 2008), it is becoming more and more important to understand how readers engage in higher-order, interpretive reading processes. Yet, we lack a unifying framework or theory to explain these interpretive behaviors in literary reading. Below, I explore how models of text comprehension might explain literary reading behaviors, as well as how other literary reading behaviors may highlight the need for expansion or modification of the extant models.

Interpretation as Level of Representation

The most obvious issue in cognitive models of text comprehension, as it relates to literature, is that interpretation has not been explicitly accounted for. The most widely accepted model, the construction-integration model (van Dijk & Kintsch, 1978), proposes three levels of representation: (a) the surface code, or the specific words of the text, (b) the textbase that reflects

the propositions from the text, or the “gist” of the meaning, and (c) the situation model that is a representation of the text that incorporates prior knowledge to understand the situation to which the text refers. Typically, having an accurate, well-integrated, situation model representation of a text reflects the highest level of comprehension. In most text comprehension work, comprehension is assessed by accuracy on inference verification questions regarding characters, setting, and causal chains. As the research has shown, literary readers can readily generate these inferences about the story, but literary readers can also construct interpretations that are not addressed in these types of assessments. The presence of interpretive inferences suggests a representation of the text that goes beyond traditional notions of the situation model.

Kintsch (1998) makes brief mention of a *communicative model* that he proposes may connect the textbase and situation model. Graesser, Millis, and Zwaan (1997) outline a *communication context* that exists in a fourth level above the situation model. Graesser and McNamara (2010) subsequently refer to it as the *pragmatic-communicative model*, and advanced it even higher to a fifth level of representation above the text’s genre/rhetorical structure. While such considerations are certainly aimed in the right direction, researchers have been unable to include these additional levels in computational analyses because they are unsure of the appropriate parameters (Graesser & McNamara, 2010). Consequently, this level of representation is often overlooked in empirical studies of discourse comprehension. Further work that assesses all four levels (surface code, textbase, situation model, interpretive) must be conducted to understand how it is related to, but distinct from, the other levels of representation.

One important move that literary researchers must make is in establishing an accepted way of assessing interpretations and interpretive inferences. Thus far, the most effective way we have been able to identify evidence of interpretation is through the coding of open-ended

responses. In our previous work, we have differentiated between *text-based inferences* that stay within the story world and *interpretive inferences* that abstract beyond the story to speak to the world at large (McCarthy & Goldman, 2015). Some other research has used the selection or ranking of appropriate morals or themes from a list (i.e., Kurtz & Schoeber, 2001; Pfaff & Gibbs, 1997). This is beneficial, as it allows for a quantifiable method of assessing a readers' interpretation. However, there are two problems with this method as a stand-alone assessment. First, in directly asking a reader for an interpretation, it is unclear if this interpretation was generated spontaneously by the reader or if the question prompt required the reader to revisit his or her understanding of the text to generate an interpretation that he or she had not previously constructed, but this problem exists as well for open-ended responses that explicitly ask for an interpretation. Thus, this method may be best used in following more indirect, open-ended responses that can assess spontaneous production of interpretative inferences (McCarthy & Goldman, in press). Second, this is a forced-choice task. It is possible and quite likely that readers could generate an interpretation of the text that is not one of the options. This problem stems from the issue that there is no one "right" answer in literary interpretation (Hillocks & Ludlow, 1984; Langer, 2010). One promising approach to begin assessing an interpretive representation is categorizing interpretations in an essay as either *supported* or *unsupported* by the information in the text (Levine & Horton, 2013). Such evaluation avoids the inherent problem of pinpointing "the correct answer", but yields more information about the nature of the reader's representation than merely assessing the presence of an interpretation. Better understanding of what makes an interpretation viable or not, perhaps through expert analysis, or through comparison to a concept map of the characters and events in the text, may better help researchers to understand how information from other levels of representation informs the

interpretive, communicative level. Conducting research of this nature will help to define the parameters of this interpretive representation and to understand how it is related to, but distinct from, the more story-specific situation model representation.

Surface Structure Puzzle

Another issue that literature problematizes is the surface code representation. Discourse comprehension research has indicated that the surface form (the exact words) of the text decay rapidly in memory (Kintsch, Welsch, Schmalhofer, & Zimny, 1990), and yet, in literature, the specific language of the text is important for the interpretive process. Bortolussi and Dixon (2013) refer to this as the *surface structure puzzle*. Is it possible that the stylistic variations of literary text are immune to this decay? Based on the fact that readers are spending more time with foregrounded passages, one might predict that they would have better memory for the specific language of the text. When readers' beliefs about the text are manipulated, believing a story is literary does lead to better surface level retention (Zwaan, 1994). However, other work has shown no difference in memory for the surface features of the text, whether the text is in its original literary form, or if it has been modified to be more mundane (Bortolussi & Dixon, 2013). If readers are attending to this information more carefully, but not showing memory benefits, how else might this information be affecting their representation of the text? Further work must be conducted to more thoroughly investigate the effects of literary language on how the text is processed, retained, and subsequently, interpreted.

One way we might consider assessing readers' attention to the surface features of the text and how this affects interpretation would be to use an eye-tracking paradigm. Dixon and Bortolussi (2011) noted that no eye-tracking studies have been conducted using literary texts. Eye-tracking is a relatively non-invasive, on-line measure, when compared to presenting a text

line-by-line, as is typical in reading time protocols, or when asking readers to engage in thinking aloud. However, the fine-grained data eye-tracking can provide is overwhelming unless investigators have a priori predictions regarding word or phrase-level regions of interest. Another shortcoming of this methodology is that, as others have noted, in authentic works, inferences may span several pages, or even an entire novel (Rapp, Komeda, & Hinze, 2011). Eye-tracking technology limits the length of text that can be presented on a single screen. Consequently, tracking regression across long spans of text would be problematic. One interesting point of entry might be to use short-form poetry. In addition to brevity, poetry makes heavy use of foregrounding. Researchers could identify foregrounded phrases, and, in addition to replicating the previous reading time results, researchers could also explore if readers regress back to these parts of the text more frequently than the background sentences. Longer fixations at or near the foregrounded text, slowed reading time on spillover sentences, or regressions to key regions of interest may indicate a change in processing. It would be of value to compare the eye movements of novices and experts, as well as novices given different reading goals, to see if reading behaviors reflect these differences. It would also be interesting to explore how manipulating authentic literary text by removing, adding, or changing key textual features might affect eye movement. Altering a literary work presents some problems, in that changing even a single word in a literary work may affect the overall impact of the text, but as long as the results are interpreted with caution, this may be the most effective way to understand how textual features differentially affect processing.

In order to understand literary interpretation, we must not only know about what goes on during reading and what goes on after reading, but also how processing of both are related. This highlights another limit to the existing work. There seems to be a disconnect between what

happens during reading and the subsequent interpretations readers produce. Studies indicate that literary texts are read differently than nonliterary text (Hanauer, 1996; Zwaan, 1994), but how these differential on-line processes affect the resulting interpretation of the work has not been established. Conversely, other work has shown that differing epistemological beliefs do not affect comprehension of the text at a literal level, but affect the amount of interpretation produced (Mason, Scrica, & Salvi, 2006). Studies that have looked specifically at literary texts tend to focus on on-line measures, off-line memory-based tasks, or interpretation, but few have looked at all three of these measures within the same investigation. Few have attempted to connect what readers do during reading to their subsequent interpretation of the work. Measurement of on-line eye movements would be an excellent complement to off-line interpretive behaviors to explore if certain reading behaviors predict the quality or quantity of interpretive behaviors, and if those parts of the text that were most attended to become evident in the readers' interpretations. Research that investigates different points of the reading process is needed to better understand how interpretations are constructed.

Literary Reading as Goal-Directed Reading

Finally, one of the most important findings from the literary interpretation research is that literary reading yields different reading behaviors than when readers maintain a more domain-general, literal stance towards the text. It seems to be the case that these interpretations are not generated automatically, and are dependent on the particular task, text, and reader. It is important, then, to have a descriptive and predictive model of literary comprehension that does not assume an ideal reader, but instead reflects real readers with varying levels of knowledge and with particular reading goals. Kendeou, Bohn-Gettler, and Fulton (2011) noted that the “default assumption for reading comprehension models, has been *reading for general comprehension*.”

What seems to be apparent in this exploration of literary research is that engaging with a text in a literary way (i.e., constructing one or several interpretations of a text) is different than general comprehension in both processing and outcome. It is unlikely that literary reading is wholly different than other kinds of comprehension, but we have little empirical evidence that literary reading follows the same rules as the kinds of text and reading described in the existing models. Rather than treating literature as an exception to the rule, perhaps using literary texts may better inform our understanding of text comprehension at large. Much more work must be conducted to explore the idea of literary reading as a form of goal-directed reading. Understanding this will help us to modify or expand upon existing models of comprehension, in terms of what kinds of inferences are generated and when.

This more goal-directed view is consistent with the Constructionist (Graesser, Singer, & Trabasso, 1994) and Landscape models (van den Broek, Young, Tzeng, & Linderholm, 1999) of text comprehension. However, neither of these models has directly tested their predictions with literary text, or with interpretation as an outcome measure. The Constructionist model proposes that readers do not rely only on automatic reading processes, but also actively search for meaning to fulfill specific reading goals. The model centers around three assumptions: The *reader goal assumption* is that the reader's purpose informs the construction of the representation. The *explanation assumption* is that readers work to explain why the actions and events are in the text. The *coherence assumption* is that readers attempt to make the text make sense at both the global and local levels. The Constructionist model suggests that readers generate different inferences on a situation-by-situation basis as they help the reader to satisfy these goals. The Landscape model suggests that the reader's goals inform what *standards of coherence* are activated, and these

standards set the criterion for successful comprehension. This criterion then dictates which inferences are generated in order to meet that standard.

One particularly interesting notion in the Constructionist model that could be modified to explain literary reading is the balance between maintaining local and global coherence. One way that reading and discourse comprehension researchers might be able to account for the role of foregrounding on interpretation is thinking of these stylistic variations as a kind of coherence break (Kintsch, 1998). The unusual way of wording may force readers into more active processing in order to repair the break in local coherence. Thus, foregrounding demands readers immediately remediate their representations, which requires more active attention. Consistent with this notion are data showing that literary texts are read more slowly than non-literary texts (Hanauer, 1998), and that foregrounded sentences are read more slowly than non-foregrounded sentences (Miall & Kuiken, 1994a). Slower reading times are indicative of deeper processing and the activation of prior knowledge to supplement the construction of a representation (Graesser, Lu, Olde, Cooper-Pye, & Whitten, 2005).

Graesser, Singer, and Trabasso (1994) assume that readers are consistently trying to maintain local and global coherence, and this seems to operate orthogonally to the reader's goals. While the extant work supports the idea of maintaining coherence at both local and global levels, it may be that the reader's goals inform what kinds of coherence need be maintained for comprehension with literary text, putting precedence upon global, rather than local, coherence.

Interactivity of Text and Reader: A Strategy Activation Account

Instead of this foregrounding-as-coherence-break account mentioned above, I propose an alternative strategy-activation explanation for how textual features affect literary processing. This is grounded in approaches to comprehension that draws upon both the coherence and

explanation assumptions unique to the Constructionist model, as well as standards of coherence in the Landscape model. The explanation assumption suggests that readers search for meaning to explain why the particular elements of the text exist. This may account for the fact that even though readers of literature take a more holistic perspective on meaning-making, they still put more time and attention into metalinguistic aspects of the text than during nonliterary reading (Claassen, 2012; Graves & Frederiksen, 1991; Peskin, 1998). The explanation assumption would predict that the reader would work to understand why the author chose to write the text the way that he or she did, rather than in a more straightforward way. This would suggest that the reader reads the language more slowly (at a local level) in service of understanding and maintaining global coherence in terms of the deeper meaning of the work.

I propose that foregrounding signals to the reader that he or she should use a particular reading strategy. As literary texts have a different purpose than other kinds of texts, a more domain-specific reading strategy would be more effective. It has been suggested that literary devices are *rules of notice*, or conventionalized signals within the text, left by the author to draw attention to certain aspects of the text, particularly that the reader should not take the text at face value (Rabinowitz, 1987). These rules (such as repetition, a shift in tone, juxtaposition, privileged position, and deviations from norms or disruptions and discrepancies) indicate to the reader how to approach the text. Unlike previous accounts of foregrounding that put emphasis on trying to establish local coherence, this strategy-activation account suggests that foregrounding encourages effort to be put into constructing global coherence and nonliteral inferences. That is, when readers struggle to understand a text at a literal level because of defamiliarization, they shift their attention to making higher-order interpretive representations to make sense of the text. Given that interpretations represent ideas that are explicitly stated in the text, an emphasis on

global over local coherence seems to have some explanatory power. For example, a literary passage, such as “Who has not Congress, lobster, love, luau, - the Regency Room, the Statue of Liberty, - runs. A sloppy amalgamation.” from Gwendolyn Brooks’ *Boy Breaking Glass*, does not make sense at a semantic level. It would be unsurprising to find readers slowing down at such a passage as they worked to make sense of it. However, such a bottom-up approach would likely be in vain. A literary reader could bypass this gap in local coherence by shifting to understanding the text as a whole and the words’ aesthetic purpose instead of their semantic one. The foregrounded features of the text are signals that encourage the reader to adopt a new purpose for reading that focuses on the bigger picture of the text, rather than specific plot elements (Hillocks & Ludlow, 1984; Johnson-Laird, 1983; Vipond & Hunt, 1984). This strategy-activation account would be consistent with the standards of coherence of the Landscape model approach to comprehension, but importantly suggests that these standards could be activated from factors outside of explicit task instruction. Further, these standards of coherence are not preexisting structures, but are built through experience and establish what qualifies as adequate comprehension. Experts seem to be able to activate the relevant standards of coherence without explicit instruction, suggesting that these standards of coherence are part of their domain-specific schema and are an internal intention, as opposed to an external intention (McCrudden, Magliano, & Schraw, 2010; McCrudden & Schraw, 2007). A strategy-activation explanation suggests that understanding who adopts a literary stance, and when and how they do so, is particularly crucial to understanding literary reading. One potentially useful future direction would be to explore other means of inducing the adoption of an interpretive stance. That is, if it is the case that foregrounded text serves as an indicator that interpretation is appropriate, then texts that have more foregrounding should yield more evidence of interpretation than a mundane version of this

same text. Exactly how explicit the rules need to be, if all readers recognize the rules, and more detail on how this shift in strategy affects comprehension, have not been adequately explored.

While great progress has been made in understanding how readers construct meanings beyond the literal words in the text, the limitations and concerns I present highlight the need for continued empirically rigorous work in this complex domain. It is my hope that this review encourages a broader discussion amongst those interested in literary comprehension.

References

- Abrahamsen, E. P., & Sprouse, P. T. (1995). Fable comprehension by children with learning disabilities. *Journal of Learning Disabilities, 28*, 302-308. doi: 10.1177/002221949502800506
- Anderson, R. C., & Pichert, J. W. (1978). Recall of previously unrecallable information following a shift in perspective. *Journal of Verbal Learning & Verbal Behavior, 17*, 1-12. doi:10.1016/S0022-5371(78)90485-1
- Bortolussi, M., & Dixon, P. (1996). The effects of formal literary training on literary reception. *Poetics, 23*, 471-487. doi: 10.1016/0304-422X(95)00014-B
- Bortolussi, M., & Dixon, P. (2013). Minding the text: Memory for literary narrative. In L. Bernaerts, L. Herman, B. Vervaeck, & D. de Geest (Eds.) *Stories and minds: Cognitive approaches to literary narrative* (pp. 23-37). Lincoln: University of Nebraska Press.
- Bråten, I., & Strømsø, H. I. (2006). Constructing meaning from multiple information sources as a function of personal epistemology. *Information Design Journal, 14*, 56-67. Doi: 10.1075/idj.14.1.07bra
- Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology, 4*, 55-81. doi: 10.1016/0010-0285(73)90004-2
- Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5*, 121-152.
- Claassen, E. (2012). *Author representation in literary reading*. Utrecht, Netherlands: John Benjamins.
- Comer Kidd, D., & Castano, E. (2013). Reading literary fiction improves theory of mind. *Science, 342*, 377-380. doi: 10.1126/science.1239918

- Culler, J. (1994). Structuralism in literature. In D. Keeseey (Ed.), *Contexts for criticism* (pp. 280-289). Mountain View, CA: Mayfield.
- Dixon, P., & Bortolussi, M. (2011). The scientific study of literature: What can, has, and should be done. *Scientific Study of Literature, 1*, 59-71. doi: 10.1075/ssol.1.1.06dix
- Dixon, P., Bortolussi, M., Twilley, L. C., & Leung, A. (1993). Literary processing and interpretation: Towards empirical foundations. *Poetics, 22*, 5-33. doi: 10.1016/0304-422X(93)90018-C
- Fish, S. (1980). *Is there a text in this class? The authority of interpretive communities*. London, United Kingdom: Harvard University Press.
- Djikic, M., Oatley, K., & Moldoveanu, M. C. (2013). Opening the closed mind: The effect of exposure to literature on the need for closure. *Creativity Research Journal, 25*, 149-154. doi: 10.1080/10400419.2013.783735
- Dorfman, M. H., & Brewer, W. F. (1994). Understanding the points of fables. *Discourse Processes, 17*, 105-129. doi: 10.1080/01638539409544861
- Elfenbein, A. (2006). Cognitive science and the history of reading. *PMLA, 121*, 484-502.
- Gibbs, R. (2001). Authorial intentions in text understanding. *Discourse Processes, 32*, 73-80. doi: 10.1207/S15326950DP3201_04
- Goldman, S. R. (2004). Cognitive aspects of constructing meaning through and across multiple texts. In N. Shuart-Faris & D. Bloome (Eds.), *Uses of intertextuality in classroom and educational research* (pp. 317-351). Greenwich, CN: Information Age.
- Goldman, S. R. (2012). Adolescent literacy: Learning and understanding content. *The Future of Children, 22*, 89-116. doi: 10.1353/foc.2012.0011

- Goldman, S. R., & Bisanz, G. (2002). Toward a functional analysis of scientific genres: Implications for understanding and learning processes. In J. Otero, J. A. Leon, & A. C. Graesser (Eds.), *The psychology of science text comprehension*. Mahwah, NJ: Erlbaum.
- Goldman, S. R., McCarthy, K. S., & Burkett, C. (2015). Interpretive inferences in literature. In E. O'Brien, A. Cook, & R. Lorch (Eds.), *Inferences during reading*. New York, NY: Cambridge University Press.
- Graesser, A. C., & McNamara, D. S. (2011). Computational analyses of discourse comprehension. *Topics in Cognitive Science*, 3, 371-398. doi: 10.1111/j.1756-8765.2010.01081.x
- Graesser, A. C., Millis, K. K., & Zwaan, R. A. (1997). Discourse comprehension. In J. T. Spence, J. M. Darley, & D. J. Foss (Eds.), *Annual review of psychology*. Palo Alto, CA: Annual Reviews. doi: 10.1146/annurev.psych.48.1.163
- Graesser, A. C., Lu, S., Olde, B., Cooper-Pye, E., Whitten, S. N. (2005). Question asking and eye tracking during cognitive disequilibrium: Comprehending illustrated texts when the devices breakdown. *Memory and Cognition*, 33, 1235-1247. doi: 10.3758/BF03193225
- Graesser, A. C., Pomeroy, V. J., & Craig, S. D. (2002). Psychological and computational research on theme comprehension. In M. Louwerse & W. Van Peer (Eds.). *Thematics: Interdisciplinary studies* (pp. 19-34). Amsterdam, Netherlands: John Benjamins.
- Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, 101, 371-395. doi: 10.1037/0033-295X.101.3.371
- Graves, B., & Frederiksen, C. H. (1991). Literary expertise in the description of fictional narrative. *Poetics*, 20, 1-26. doi: 10.1016/0304-422X(91)90031-J

- Hanauer, D. (1998). The genre-specific hypothesis of reading: Reading poetry and encyclopedic items. *Poetics*, 26, 63-80. doi: 10.1016/S0304-422X(98)00011-4
- Harker, W. J. (1996). Toward a defensible psychology of literary interpretation. In R. J. Kreuz & S. M. MacNealy (Eds.), *Empirical approaches to literature and aesthetics* (pp. 645-658). Norwood, NJ: Ablex.
- Hillocks, G., & Ludlow, L. H. (1984). A taxonomy of skills in reading and interpreting fiction. *American Educational Research Journal*, 21, 7-24.
- Hoffstaedter, P. (1987). Poetic text processing and its empirical investigation. *Poetics*, 16, 75-91. doi: 10.1016/0304-422X(87)90037-4
- Johnson, D. F., & Goldman, S. R. (1987). Children's recognition and use of rules of moral conduct in stories. *The American Journal of Psychology*, 100, 205-224. doi: 10.2307/1422404
- Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge, MA: Harvard University Press.
- Kendeou, P. Bohn-Gettler, C.M., & Fulton, S. (2011). What we have been missing: The role of goals in reading comprehension. In M. T. McCrudden, J. P. Magliano, & G. Schraw (Eds.), *Text relevance and learning from text*. Greenwich, CT: Information Age.
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge, England: Cambridge University Press.
- Kinstch, W., Welsch, D., Schmalhofer, F., & Zimny, S. (1990). Sentence memory: A theoretical analysis. *Journal of Memory and Language*, 29, 133-159. doi: 10.1016/0749-596X(90)90069-C

- Kotovsky, K., Hayes, J. R., & Simon, H. A. (1985). Why are some problems hard? Evidence from Tower of Hanoi. *Cognitive Psychology*, *17*, 248-294. doi: 10.1016/0010-0285(85)90009-X
- Kreuz, R. J., & Roberts, R. M. (1993). The empirical study of figurative language in literature. *Poetics*, *22*, 151-169. doi: 10.1016/0304-422X(93)90026-D
- Kurtz, V., & Schober, M. F. (2001). Readers' varying interpretations of theme in short fiction. *Poetics*, *29*, 139-166. doi: 10.1016/S0304-422X(01)00040-7
- Langer, J. A. (2010). *Envisioning literature: Literary understanding and literature instruction, 2nd edition*. New York, NY: Teachers College Press.
- Lea, R. B., Rapp, D. N., Elfenbein, A., Mitchel, A. D., & Swinburne Romine, R. (2008). Sweet silent thought: Alliteration and resonance in poetry comprehension. *Psychological Science*, *19*, 709-716. doi: 10.1111/j.1467-9280.2008.02146.x
- Lee, C. D. (2007). *Culture, literacy, and learning: taking bloom in the midst of the whirlwind*. New York, NY: Teachers College Press.
- Lee, C. D. (2011). Education and the study of literature. *Scientific Study of Literature*, *1*, 49-58.
- Levine, S., & Horton, W. S. (2013). Using affective appraisal to help readers construct literary interpretations. *Scientific Study of Literature*, *3*, 105-136. doi: 10.1075/ssol.3.1.10lev
- Louwerse, M. M., Benesh, N., & Zhang, B. (2008). Computationally discriminating literary from non-literary texts. In S. Zyngier, M. Bortolussi, A. Chesnokova, & J. Auracher (Eds.), *Directions in empirical literary studies* (pp. 175-192). Amsterdam, Netherlands: John Benjamins.
- Magliano, J. P., Baggett, W. B., & Graesser, A. C. (1996). A taxonomy of inference categories that may be generated during the comprehension of literary texts. In R. J. Kreuz & S. M.

- MacNealy (Eds.), *Empirical approaches to literature and aesthetics* (pp. 201-220). Norwood, NJ: Ablex.
- Magliano, J. P., & Graesser, A. C. (1991). A three-pronged method for studying inference generation in literary text. *Poetics*, *20*, 193-232. doi: 10.1016/0304-422X(91)90007-C
- Mason, L., Scirica, F., & Salvi, L. (2006). Effects of beliefs about meaning construction and task instructions on interpretation of narrative texts. *Contemporary Educational Psychology*, *31*, 411-437. doi: 10.1016/j.cedpsych.2005.12.001
- McCarthy, K. S., & Goldman, S. R. (in press). Comprehension of short stories: Effect of task instruction on literary interpretation, *Discourse Processes*. doi: 10.1080/0163853X.2014.967610
- McCarthy, P.M., Myers, J.C., Briner, S.W., Graesser, A.C., & McNamara, D.S. (2009). A psychological and computational study of sub-sentential genre recognition. *Journal for Language Technology and Computational Linguistics*, *24*, 23-55.
- McCrudden, M. T., Magliano, J. P., & Schraw, G. (2010). Exploring how relevance instructions affect personal reading intentions, reading goals and text processing: A mixed methods study. *Contemporary Educational Psychology*, *35*, 229-241. doi: 10.1016/j.cedpsych.2009.12.001
- McCrudden, M. T., & Schraw, G. (2007). Relevance and goal-focusing in text processing. *Educational Psychology Review*, *19*, 113-139. doi: 10.1007/s10648-006-9010-7
- McNamara, D. S., & Magliano, J. P. (2009). Toward a comprehensive model of comprehension. *Psychology of Learning and Motivation*, *51*, 297-384. doi: 10.1016/S0079-7421(09)51009-2
- Means, M. L., & Voss, J. F. (1985). Star Wars: A developmental study of expert and novice

- knowledge structures. *Journal of Memory and Language*, 24, 746-757. doi: 10.1016/0749-596X(85)90057-9
- Miall, D. S., & Kuiken, D. (1994a). Beyond text theory: Understanding literary response. *Discourse Processes*, 17, 337-352. doi: 10.1080/01638539409544873
- Miall, D. S., & Kuiken, D. (1994b). Foregrounding, defamiliarization, and affect response to literary stories. *Poetics*, 22, 389-407. doi: 10.1016/0304-422X(94)00011-5
- Miall, D. S., & Kuiken, D. (1998). The form of reading: Empirical studies of literariness. *Poetics*, 25, 327-341. doi: 10.1016/S0304-422X(98)90003-1
- Narvaez, D., Bentley, J., Gleason, T., & Samuels, J. (1998). Moral theme comprehension in third graders, fifth graders, and college students. *Reading Psychology*, 19, 217-241. doi: 10.1080/0270271980190203
- Peskin, J. (1998). Constructing meaning when reading poetry: An expert-novice study. *Cognition and Instruction*, 16, 135-263. doi: 10.1207/s1532690xci1603_1
- Peskin, J. (2007). The genre of poetry: Secondary school students' conventional expectations and interpretative operations. *English in Education*, 41, 20-36. doi: 10.1111/j.1754-8845.2007.tb01162.x
- Rabinowitz, P. (1987). *Before reading: Narrative conventions and the politics of interpretation*. Columbus: Ohio State University Press.
- Rapp, D. N., Komeda, H., & Hinze, S. R. (2011). Vivifications of literary investigation. *Scientific Study of Literature*, 1, 123-135. doi: 10.1075/ssol.1.1.13rap
- Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology*, 82, 498-504. doi: 10.1037/0022-0663.82.3.498

- Schraw, G. (1997). Situational interest in literary text. *Contemporary Educational Psychology*, 22, 436-456. doi: 10.1006/ceps.1997.0944
- Schraw, G. (2000). Reader beliefs and meaning construction in narrative text. *Journal of Educational Psychology*, 92, 96–106. doi: 10.1037/0022-0663.92.1.96
- Schraw, G., & Bruning, R. (1996). Readers' implicit models of reading. *Reading Research Quarterly*, 31, 290-305. doi: 10.1598/RRQ.31.3.4
- Snow, C., & Biancarosa, G. (2003). *Adolescent literacy and the achievement gap: What do we know and where do we go from here?* New York, NY: Carnegie Corporation of New York.
- Stathopoulou, C., & Vosniadou, S. (2007). Exploring the relationship between physics-related epistemological beliefs and physics understanding. *Contemporary Educational Psychology*, 32, 255-281. doi: 10.1016/j.cedpsych.2005.12.002
- Tobias, S. (1994). Interest, prior knowledge, and learning. *Review of Educational Research*, 64, 37-54. doi: 10.3102/00346543064001037
- van den Broek, P., Lorch, R. F., Linderholm, T., & Gustafson, M. (2001). The effects of readers' goals on inference generation and memory for texts. *Memory and Cognition*, 29, 1081-1087. doi: 10.3758/BF03206376
- van den Broek, P., Young, M., Tzeng, Y., & Linderholm, T. (1999). The landscape model of reading: Inferences and the on-line construction of a memory representation. In H. van Oostendorp & S. R. Goldman (Eds.), *The construction of mental representations during reading* (pp. 71-98). Mahwah, NJ: Erlbaum.

- Vipond, D., & Hunt, R. A. (1984). Point-driven understanding: Pragmatic and cognitive dimensions of literary reading. *Poetics, 13*, 261-277. doi: 10.1016/0304-422X(84)90005-6
- Warren, J. E. (2011). "Generic" and "specific" expertise in English: An expert/novice study in poetry interpretation and academic argument. *Cognition and Instruction, 29*, 349-374. doi: 10.1080/07370008.2011.607929
- Wiley, J., & Voss, J. F. (1999). Constructing arguments from multiple sources: Tasks that promote understanding and not just memory for text. *Journal of Educational Psychology, 91*, 301-311. doi: 10.1037/0022-0663.91.2.301
- Zeitz, C. M. (1994). Expert-novice differences in memory, abstraction, and reasoning in the domain of literature. *Cognition and Instruction, 4*, 277-312. doi: 10.1207/s1532690xci1204_1
- Zhang, H., & Hoosain, R. (2005). Activation of themes during narrative reading. *Discourse Processes, 40*, 57-82. doi: 10.1207/s15326950dp4001_3
- Zwaan, R. A. (1993). *Aspects of literary comprehension: A cognitive approach*. Utrecht, Netherlands: John Benjamins.
- Zwaan, R. A. (1994). Effect of genre expectation on text comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 920-933. doi: 10.1037/0278-7393.20.4.920
- Zwaan, R. A. (1996). Processing narrative time shifts. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 1196-1207. doi: 10.1037/0278-7393.22.5.1196
- Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin, 123*, 162-185.