Constructing Interpretive Inferences about Literary Text: The Role of Domain-Specific Knowledge

Kathryn S. McCarthy & Susan R. Goldman
University of Illinois at Chicago

To appear in Learning and Instruction
Accepted December 13, 2017
ABSTRACT

Student readers struggle to construct the interpretive inferences necessary for successful literary comprehension. Expert think-alouds were conducted to identify the kinds of domain-specific knowledge that were drawn upon when reading the short story *The Elephant*. These data were used to construct three reading instructions provided to student (novice) literary readers. These instructions informed the student about two types of literary conventions (Rules of Notice, Rules of Signification, Combined). Analysis of the students’ essays indicated having both types of domain-specific knowledge yielded the most interpretive inferences. Attention to language mediated the effect suggesting a means for domain-specific knowledge to be used to leverage student engagement in literary interpretation.

**Keywords**: literary text comprehension, interpretation, inferences, prior knowledge
Highlights

- Successful literary reading requires the generation of interpretive inferences.
- Students do not readily construct these interpretive inferences.
- Experts rely on domain-specific knowledge of literary conventions to generate these inferences.
- Providing expert-like literary convention knowledge increases production of interpretive inferences.
Constructing Interpretive Inferences about Literary Text: The Role of Domain-Specific Knowledge

1. INTRODUCTION

1.1. Domain-Specific Reading

Recent work in reading and discourse comprehension, in conjunction with changing literacy standards, has emphasized the importance of domain-specific reading (Goldman, 2012; Goldman, et al., 2016). Domain or discipline-specific reading emphasizes that different disciplines require different reading skills and strategies in order to make sense of and learn from texts. That is to say that a textbook for a science class is read differently than a novel for an English class (e.g., Hanauer, 1998). More specifically different amounts and types of inferences are generated to maintain comprehension given different types of texts and tasks. The majority of this work investigates expository texts within the domains of science and history. Less attention has been paid to literature and understanding how students read, process, understand, and learn from literary texts (McCarthy, 2015).

1.2. Literary Text and Interpretation

In literary reading, one of the reader’s goals is to identify a deeper meaning in the work and examine how the author conveyed this meaning through his or her manipulation of language (Langer, 2010; Lee, 2007, 2011; Levine & Horton, 2013). Thus, in addition to the processing that must occur to maintain a mental representation of the story world, successful literary readers must also construct interpretive inferences that go beyond the world of the story to speak to the world at large (Goldman, McCarthy, & Burkett, 2015; Magliano, Baggett, & Graesser, 1996; McCarthy & Goldman, 2015). Expert literary readers readily produce these interpretive inferences when reading and writing about literary works. Novices, on the other hand, do not generate these interpretive inferences. Rather they mainly restate elements of the plot (e.g.
Earthman, 1992; Graves & Frederiksen, 1991; Peskin, 1998; Zeitz, 1994). Novices in these studies (high school and college students) are likely expert readers in the general sense, but are novice literary readers because they do not have extensive experience in interpreting literary works. For the remainder of this paper, *expertise* refers to years of formal training in literary analysis. One possible reason novices tend not engage in interpretation is that they may be unaware of the interpretive purpose of literary works. Rather than adopting an *interpretive* stance, or reading goal, toward the text, they maintain a more domain-general *literal* stance (Goldman, McCarthy, & Burkett, 2015). Indeed, when novices are given task instructions that bias an interpretive stance they produce more interpretive inferences (McCarthy & Goldman, 2015). Of course, simply knowing that interpretation is a goal of literary reading does not mean that the novice will be able to construct interpretations or that the interpretations they produce will be as sophisticated as those constructed by the experts. In addition to knowledge of the interpretive goal of literary text, experts also have knowledge about common literary conventions and themes that they can draw upon. This knowledge may be what allows them to recognize that an interpretive stance is appropriate, construct interpretations, and justify the interpretation with textual evidence and appeals to cultural and literary norms (Lee & Goldman, 2015; McCarthy, 2015).

Rabinowitz (1987) refers to these literary conventions as *Rules of Notice* and *Rules of Signification*. He suggests that authors are able to convey particular meanings because author and reader have shared knowledge of these rules. *Rules of Notice*, such as repetition, tone shift, juxtaposition, privileged position, deviations from the norm, and disruptions or discrepancies, are stylistic variations that cue the reader to look for meaning beyond the literal and draw the reader’s attention to specific parts of the text. Importantly, knowledge of what to notice is only
the first step in constructing interpretations. Think-aloud studies show that novice readers do
attend to some of these stylistic variations, particularly disruptions or discrepancies that impede
their comprehension of the literal aspects of the text. However, they do not use the information to
make an interpretation. Instead, they tend to comment on the rule of notice as odd or confusing
and may attribute it to their own comprehension failure (Burkett & Goldman, 2016; Graves &
Frederiksen, 1991). This highlights the importance of the second aspect of convention
knowledge, Rules of Signification. Rules of Signification are the conventionalized ways of
making meaning from what is noticed. These rules reflect common themes (e.g. man’s
inhumanity to man, loss of innocence) and genres (e.g. political satire, romanticism) to construct
interpretations. There are stark contrasts between experts and novices use of these rules, as
illustrated by the following example from a think-aloud study that featured a passage from *The
Color Purple* (Graves & Frederiksen, 1991). The passage contained sentences that were
syntactically incorrect. A novice reader noted this was confusing and difficult to read. In
contrast, the expert reread and determined that the sentence likely reflected a southern black
dialect. From there the expert was able to contextualize the work and consider that an
interpretation of the text might speak to race relations.

Thus, the knowledge experts bring to bear when reading literary works encompasses
knowing the interpretive purpose of literature, what features of the text to attend to, as well as
how to integrate these features into an interpretation. One way in which researchers and
educators may be able to encourage novice readers to engage in interpretation may be to
familiarize them with this domain knowledge.

1.3. The Current Study
The purpose of the current study was to explore the effect of domain-specific knowledge on novice readers’ interpretative behaviors, both in terms of the production of interpretive inferences and an increased attention to the specific language in the text. Rather than guess at whether, where, and what kind of literary convention knowledge was relevant in a given text, we first collected expert think-aloud data to identify the kinds of behaviors experts engage in and the aspects of the text they attended to for this particular story (Study 1). This information was then used to construct three reading instructions that provided readers with information about 1) Rules of Notice, 2) Rules of Signification, or 3) a combination of both types of information (Study 2). We assessed how these different instructions affected the generation of interpretive inferences and how this was related to attention to language in the text.

2. STUDY 1: EXPERT THINK-ALOUD STUDY

Experts were recruited to identify the Rules of Notice and Rules of Signification relevant for this particular story. Consistent with previous research, expertise was defined by amount of formal training (e.g., English graduate students or faculty). In this study, our experts were English PhDs who were current faculty members in a college English department. In addition, these experts completed the Author Recognition Task (ART; Acheson, Welles, & MacDonald, 2008) which assesses readers’ familiarity and exposure to literature and fiction.

2.1. Method

2.1.1. Participants

Participants were four English PhDs, employed as full-time members of the English Department at a small mid-Atlantic liberal arts college. Experts were given a small monetary compensation ($10) for their time.

2.1.2. Materials
2.1.2.1. **Author Recognition Task**

The Author Recognition Task (ART; Acheson, Wells, & MacDonald, 2008) requires participants to read a list of 130 names and put an X by those names they are certain are names of authors. The ART score is calculated by subtracting the number of false alarm recognitions (those names marked with an X that are not famous authors) from correct recognitions. As there are 65 real authors, the highest score possible is 65.

2.1.2.2. **Think-aloud and reading instruction**

The think-aloud and reading instruction was adapted from Peskin (1998, 2007):

*Today you will be reading a story. When you are finished reading, you will be asked to answer a few questions about the text. As you read aloud, I’d like you to think aloud as you try to make sense of the text. Say everything you are thinking. It’s just as if you are turning up the volume on your associations, inferences, or any minor thoughts as they flit through your mind. Don’t censor anything.*

Graves and Frederiksen (1991), who used a similarly short and open-ended think-aloud instruction, note that such a task is “familiar and appropriate for anyone studying English literature” (p. 8). Consequently, it was decided that additional scaffolding would be unnecessary and could bias the experts into different behaviors than they would provide by themselves.

2.1.2.3. **Short Story**

The text was an English translation of the satirical short story *The Elephant* by Slowomir Mrozek (1972). The story is about a zoo director who, in order to cut costs, replaces a real elephant with a fake one made of a rubber skin filled with air. In the conclusion of the story, the fake elephant is carried away by a breeze, much to the astonishment of the schoolchildren at the exhibit. This story was selected because it has been used in previous work on literary interpretation. Students have little trouble understanding the plot events in the text, but do not readily generate interpretations of the text (McCarthy & Goldman, 2015).
2.1.3. Procedure

Expert participants completed the study individually in one-hour sessions. Participants first completed the ART. They were then given the think-aloud instruction on a sheet of paper that was also read aloud by the experimenter. To acquaint the participants with thinking aloud, the experimenter modeled the think-aloud process with the first half of a short example text. The participant then read the second half of this example to practice and get comfortable with the procedure. After any questions were addressed, participants read and thought aloud about the target text, *The Elephant*. The think-aloud was audio-recorded.

2.2. Results

2.2.1. Author Recognition Task

Previous work indicates that the average ART score for an undergraduate sample is $M = 22.70$ out of a possible 65 (Acheson, Wells, & MacDonald, 2008). The expert participants in this study produced a mean score of 54.50 ($SD = 8.29$), nearly three standard deviations above this average. Thus, the experts’ familiarity with literary works is consistent with their level of expertise.

2.2.2. Think-Aloud

The think-aloud protocols were transcribed and analyzed for references to Rules of Notice, Rules of Signification, and for any additional convergence amongst the experts’ comments.

2.2.3. Rules of Notice

Of Rabinowitz’s (1987) explicitly mentioned Rules of Notice (repetition, tone shift, juxtaposition, privileged position, deviations from the norm, and disruptions or discrepancies), three were present in the think-alouds of multiple experts: deviations from the norm,
juxtaposition, and disruption. Examples of each are provided below. Note that excerpts from the think-aloud protocols have been edited to remove disfluencies.

2.2.3.1. Deviations from the norm

In this story, all four experts indicated that the story must have a deeper meaning because of its absurdity. Indeed, one expert hypothesized that the text must be satiric because of its absurd nature. She commented, “…So I’m wondering if this is a satire or a farce? Because certainly, this does not make sense as a straight story because if it was supposed to be a straight story, this would be just ridiculous and wouldn’t make any sense.”

2.2.3.2. Juxtaposition

Three of the experts commented on the sentence Placed in front of a large real rock, [the fake elephant] looked fierce and magnificent. They acknowledged that it was both humorous and telling that the author made a point to identify the inanimate rock as real while discussing a fake elephant. This juxtaposition draws attention to the absurdity of the situation at hand.

2.2.3.3. Disruptions

Disruptions, or ruptures, are parts of the text that are strange or unusual and that slow comprehension. The disruptions in this text were related to specific word choice. Three of the four experts pondered the use of the word “carcass” to describe the hide of the rubber elephant. One noted that this was a particularly “ugly image”. Another noted that this was an unusual word to choose as a description. While they mentioned they were interested in understanding why the author chose this word instead of something more typical, none of them resolved this in the think-aloud.

All four experts commented that the choice of “the jackass” in the sentence Outside, human voices were stilled and only the cry of the jackass interrupted the silence was deliberate.
Three of them explicitly identified the duality of this word as a jab at either the characters or the system in general.

Three of the experts noted that the final line *And they no longer believe in elephants* was interesting or “profound”. One expert noted, “How could you believe in an elephant? Well, clearly, what the word *believe* is telling me is that the elephants are metaphorical and they mean more that just the actual existence of the animal, right?”

Though the story hints at government and bureaucracy, there is only one specific mention of political parties. The sentence reads, *Blowing up an elephant is not an everyday job. And it’s because our director is a leftist.* All four experts commented on this abrupt introduction of politics. Two experts noted that the sudden introduction of politics was quite surprising and, thus, something worthy of a second examination. The other two, who had already suspected a political agenda in the story, used this information to confirm their suspicions that the story was a criticism of socialism or communism.

### 2.2.4. Rules of Signification

Importantly, the experts were not given any information about the text prior to reading. Even without this information, two of the four experts used clues in the text to identify the text as “satire”. As previously mentioned, one expert explicitly identified a rule of notice (deviation from the norm) and then inferred the story was likely satiric. The other two experts did not use the term *satire*, but one noted that the text was “mocking” authority and the other commented “This is written in the style of absurdity […] which you see in a lot of Soviet literature, which shows the absurdity of the ways that the socialist authorities governed.”

### 2.3. Discussion
These experts, defined as such by their formal training and performance on the ART, yielded think-aloud data consistent with previous expert studies (e.g. Graves & Frederiksen, 1991; Peskin, 1998). During reading, they produced interpretations that drew on prior knowledge about conventions, themes, and the context of the story. Critical to the development of the reading instructions in the following experiment, the think-alouds indicated that this text relied on juxtapositions, disruptions, and deviations from the norm to cue the reader that this story has a deeper meaning or message and, that this deeper message was likely a satiric critique of government or authority.

3. STUDY 2: READING INSTRUCTION MANIPULATION

Three instructional conditions were developed based on the information provided by the experts in Study 1. In the Rules of Notice condition, the task instructions indicated that experts look at the language in the text to think about a deeper meaning and provided descriptions of the three Rules of Notice commonly identified by the experts. In the Rules of Signification condition, readers were provided a brief definition of satire and its purpose as well as examples drawn from contemporary culture. Finally, the Combined condition merged the other two instructions to provide both types of domain-specific knowledge.

It was predicted that that the participants in the Combined condition would produce to the most interpretive inferences and the participants in the Rules of Notice condition would produce the least. These predictions were based on prior research that showed that recognition of Rules of Notice alone is not sufficient for interpretation. That is, students often notice “odd” things in stories, but rarely leverage them to look for deeper meaning (Burkett & Goldman, 2016; Graves & Frederiksen, 1991). In addition, it was also predicted that the reading instruction manipulation
would affect the readers’ attention to language in the text and that this attention would mediate the relationship between reading instruction and interpretive inferences.

3.1. Method

3.1.1. Participants

Ninety-three undergraduates (Female: N = 62; \( M_{\text{age}} = 19.18, SD = 1.25 \)) from a large, urban university in the Midwestern United States were given course credit for their participation. Two participants’ data were omitted due to less than 10 years of English language experience. All other participants reported that they had been speaking English for at least 10 years. Additionally, one participant’s essay was off-topic and was removed from consideration, leaving a total of 90 participants in the analysis (Rules of Notice: 29; Satire: 32; Combined: 29).

3.1.2. Design & Materials

This study employed a 3-level (reading instruction: Rules of Notice, Rules of Signification, Combined) between-subjects design.

3.1.2.1. Author Recognition Test

The ART was used to compare these college students’ familiarity and exposure to literature and fiction to that of the experts from Study 1. This test was administered prior to the experimental session in a mass testing survey.

3.1.2.2. Short story

Participants read the same English adaption of *The Elephant* used for Study 1.

3.1.2.3. Reading instructions

The three reading instructions all provided information about the title of the story and the name of the author. The instructions differed with respect to the types of domain-specific
knowledge they conveyed: Rules of Notice, Rules of Signification, and the two combined. The full instructions can be found in Appendix B.

3.1.2.3.1.  **Rules of Notice instruction**

This instruction explained that experts pay attention to linguistic cues in the text and identified the three Rules of Notice (deviations, juxtapositions, disruptions) used by experts in Study 1. Importantly, the phrase *keep in mind that not everyone noticed the same signals or all of these* was included to discourage participants from approaching the activity as a search task rather than a reading comprehension task.

3.1.2.3.2.  **Rules of Signification instruction**

The Rules of Signification instruction indicated that the author of the text was known for his use of satire. It gave a definition of satire and a description of the function of satire. Examples of from popular media were also included to help the students understand and apply the rule.

3.1.2.3.3.  **Combined instruction**

The combined instruction included both the information about Rules of Notice and Rules of Signification. It largely used the same language as the previous two reading instructions, but changed one key phrase in the first sentence of the Rules of Notice paragraph (bolded and underlined in Appendix B) to avoid repetition and to provide the information that experts use Rules of Notice in recognizing and understanding texts as satiric.

3.1.2.4.  **Essay Prompt**

Participants were given the following essay instruction: *Please write an essay answering the following question: What do you think the author is trying to say about the world? Be sure to*
use evidence from the text to support your claims. This prompt explicitly encouraged the reader to engage in interpretation as well as to justify this interpretation with evidence from the text.

3.1.2.5. Comprehension Test

A twenty-question true/false comprehension test was constructed to assess participants’ understanding of the basic elements of the story. Participants were asked to indicate whether each of the 20 events had occurred (true) or not (false). The ten true statements were constructed by asking the experts from the previous study to identify main events in the story. The ten false statements were plausible, but incorrect distractor statements. Three versions of this test were constructed with the sentences presented in randomized orders.

3.1.2.5.1. Demographic Questionnaire

Participants were asked to provide information about their age, gender, year in school, and English fluency. They also indicated their familiarity with The Elephant from having never read the text before to having read the text and analyzed it in a class.

3.1.3. Procedure

The ART was distributed at the beginning of the semester in mass testing of the Psychology Subject Pool. The ART was distributed to approximately one-third of this mass-testing sample. Taking the ART was not prerequisite for participating in the current study.

This paper-and-pencil study was conducted in groups no larger than five, with each group being randomly assigned to a reading instruction condition. Participants were given a packet that included the appropriate reading instruction, story, and essay prompt. Participants were instructed to read the reading instruction and then continued on to the story. This reading was untimed. Most participants read the text in about 10-15 minutes. After all participants in the group finished reading the story, the experimenter instructed the participants to turn the page and
read the prompt out loud. Participants were then given up to twenty-five minutes to complete their essay. After completing the essay, the experimenter collected the packet. Participants completed the comprehension test. The comprehension test was administered after the essay task to prevent processing from the comprehension test affecting the central essay writing activity. Finally, participants completed the demographic questionnaire.

3.1.4. Essay Scoring

Essays were scored for 1) types of idea units (paraphrase, text-based inferences, interpretive inferences) and 2) evidence of attention to features of the language used in the story.

3.1.4.1. Types Of Idea Units

Using the procedure and codebook from McCarthy and Goldman (2015), essays were parsed into idea units \( n = 1058 \) across all essays and categorized as 1) verbatim, 2) paraphrase, 3) text-based inference, 4) interpretive inference, or 5) other. A verbatim statement is one that comes directly from the text. Paraphrase statement reflects the gist of a sentence from the text with little to no additional information. A text-based inference includes information from prior knowledge, but stays within the story world, whereas an interpretive inference goes beyond the story to speak to the deeper meaning. Other statements were those that did not fit into any of the other categories (e.g., “Eureka!”). Two raters who had previous experience with this codebook scored a random subset of 20 essays for verbatim/paraphrase, text-based inference units, interpretive inference units, or other, with good reliability (ICC = .93). Other idea units represented less than 2% of the total idea units and were not included in the analyses.

3.1.4.2. Attention to Language Features

The essays were scored for the presence (1) or absence (0) of four features: 1) one or more Rules of Notice, 2) use of the word satire or a synonymous phrase to characterize the story,
3) a specific target group at which the critique was directed, for the story’s criticism, and 4) the
author’s use of language as evidence to support claims about the texts meaning. The same two
raters who scored the idea units scored each essay for the presence or absence of each of these
four features. There were only three disagreements, which were resolved through discussion.
Quality scores for each participant were calculated by summing the number of features present in
the essay yielding a score from 0 to 4.

3.1.5. Results

3.1.5.1. ART

More than half (n = 55) of the participants completed the ART in mass testing. These
participants were evenly distributed across reading instruction conditions (Rules of Notice: n =
16; Satire: n = 17; Combined: n = 19). ART Scores ranged from -1 to 32, with a mean score of
7.18 (SD = 6.14), suggesting that these novice participants were, indeed, less familiar with
literary text than the experts (M = 54.50). There was no difference in ART scores across the three
reading instruction conditions, $F < 1.00$, nor was ART score related to the amount of interpretive
inferences in the essays, $r(55) = .07, ns$. Consequently, ART score was not included in the
subsequent analyses.

3.1.5.2. Comprehension Test

The comprehension test indicated that participants understood the plot of the story, with
an average of 16.84 questions correct out of 20. An ANOVA revealed no significant differences
across reading instruction condition, $F < 1.00$.

3.1.5.3. Types of Idea Units

A 3-level analysis of variance (ANOVA) indicated no effect of reading instruction
condition on the amount of verbatim/paraphrase idea units, $F < 1.00$ A similar ANOVA on the
amount of text-based inference idea units in the essays, $F(2, 87) = 1.40, ns$. An ANOVA indicated the predicted effect of reading instruction condition on the amount of interpretive inference idea units, $F(2,87) = 4.83, p < .01, \eta^2_p = .10$. Post-hoc Tukey HSD tests ($p < .01$) indicated participants in the Combined condition produced more interpretive inferences than participants in the Rules of Notice condition, but that the Rules of Signification instruction condition was not significantly different from either.

Table 3. Mean number of verbatim/paraphrase, text-based inference, and interpretive inference idea units as a function of reading information condition.

<table>
<thead>
<tr>
<th></th>
<th>Verbatim/Paraphrase</th>
<th>Text-based Inferences</th>
<th>Interpretive Inferences</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
<td>$M$</td>
</tr>
<tr>
<td>Rules of Notice Instruction</td>
<td>3.79</td>
<td>3.09</td>
<td>2.41</td>
</tr>
<tr>
<td>Rules of Signification Instruction</td>
<td>3.34</td>
<td>2.09</td>
<td>3.38</td>
</tr>
<tr>
<td>Combined Instruction</td>
<td>3.72</td>
<td>2.49</td>
<td>2.90</td>
</tr>
</tbody>
</table>

3.1.5.4. **Attention to Language (Quality Score)**

Table 4 shows the effect of reading instruction condition on Quality Score. These data indicated that majority of those in the Rules of Notice reading instruction condition produced essays with low Quality Scores. The distribution of essays in the other two conditions is more distributed in terms of Quality Score. These data suggest that the reading instruction may have had an effect on participants’ propensity to attend to the language in the story and discuss it in their essay.
Table 4.
*Distribution of Essays as a function of Quality Score and Reading Instruction*

<table>
<thead>
<tr>
<th>Instruction</th>
<th>Quality Score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Rules of Notice</td>
<td>23</td>
</tr>
<tr>
<td>Rules of Signification</td>
<td>11</td>
</tr>
<tr>
<td>Combined</td>
<td>14</td>
</tr>
</tbody>
</table>

There was a significant correlation between Quality Score and number of interpretive inferences \((r = .46)\) suggesting these two scores were related, but not redundant.

A mediation model was employed to test the relationship between reading instruction condition, attention to language in the text (as indicated by Quality Score), and number of interpretive inferences. We used the PROCESS macro for SPSS, which employs a bootstrapping method to test the effect of a mediating variable on the relationship between a multicategorical (i.e. non-continuous) independent variable and a continuous dependent variable (Hayes & Preacher, 2014). The model is presented in Figure 1. The Rules of Notice condition served as the referent group \((D0)\). D1 reflects the contrast between the Rules of Signification condition and the Rules of Notice condition and D2 reflects the contrast between the Combined condition and the Rules of Notice condition. The unstandardized regression coefficients outside the parentheses reflect the effect of reading instruction on the number of interpretive inferences. The coefficients inside the parentheses reflect when Quality Score is included in the model.
Without Quality Score, the overall model was significant, $R^2 = .10$, $F(2, 86) = 4.83$, $p < .01$. This replicates the findings from the ANOVA, indicating a significant effect of reading instruction on the amount of interpretive inferences in the essays. When Quality Score was included, the model was no longer significant, $R^2 = .05$, $F(2, 86) = 2.79$, $p = .07$. This reduction of the model from significant to non-significant indicates that there is not a direct relationship between the reading instruction and number of interpretive inferences, but rather that the reading instruction influences the Quality Score, which influences the amount of interpretive inferences produced. Sobel tests indicated both the Rules of Signification and Combined conditions significantly differed from the Rules of Notice referent group. Point estimates indicate the mean over 1000 samples. Confidence intervals that do not cross zero reflect a significant mediation effect. The Rules of Signification condition yielded a point of estimate of .85 and a 95% confidence interval of .24 and 1.91. The Combined condition yielded a point of estimate of .75
and a 95% confidence interval of .16 and 1.78. The model suggests that, relative to the Rules of Notice condition, both the Rules of Signification and Combined conditions affected the amount of attention paid to language in the text, which, in turn, affected the number of interpretive inferences in the essay.

As indicated by the comprehension test scores, the reading instructions did not affect the participants’ literal comprehension of the story. However, the instructions did yield differences in amount of interpretation. Importantly, the effect of reading instructions on the amount of interpretive inferences was mediated by to how much the reader attended to aspects of the language in the story.

4. DISCUSSION

This study showed that providing domain-specific knowledge about literary conventions increased novice readers’ inclusion of interpretive inferences and attention to language features in their essays. Participants who were given information about both Rules of Notice (stylistic variations in the text that signal to the reader to pay attention) and Rules of Signification (common meanings or themes that can be drawn upon from what is noticed) produced more interpretive inferences than those who were provided only the Rules of Notice information. This finding highlights that noticing is just the first step toward literary interpretation. Developing literary readers need support in connecting what they notice about the linguistic aspects of text to the potential interpretive significance. This literary convention knowledge provides a basis for moving beyond the literal text to explore how the text speaks to the world at large. The mediation analyses further revealed that it was the readers’ greater attention to language in the story (Quality Score) that mediated the relationship between the reading instruction and the production of interpretive inferences.
Two promising lines of future work are suggested by the present results. The first is the development of longer-term interventions that provide instruction on how to identify rules of notice and how to connect these stylistic variations to deeper meanings of the text. This would allow students time to practice these skills, receive feedback on their work, and to see how these themes and rhetorical choices manifest across a variety of texts. Such an intervention could also implement writing instruction to help students articulate and justify their interpretations in a domain-appropriate way.

A second line of work might explore other types of expert knowledge. As seen in previous expert-novice comparison studies (e.g. Graves & Frederiksen, 1991; Peskin, 1998) experts have prior knowledge about common themes and genres of works, but also knowledge about specific authors or groups of authors who repeatedly use a certain rhetorical device or often write about a certain set of themes. This knowledge may also include information about the time and place a story was written and, importantly, the knowledge that literary works should be contextualized. Others have noted that interpretive inferences are “difficult to construct without the pragmatic context of the text, such as who wrote the text, why it was written, who read the text, and why it was read” (Graesser, Singer, & Trabasso, 1994, p. 373). This is to say that, even if a reader adopts an interpretive stance, it may be difficult to generate (and, further, justify) interpretations if the reader has no knowledge about the author or the time in which the author was writing (Lee & Spratley, 2010). Understanding the context of the work is likely an important aspect of knowledge that experts can activate that sheds light upon why it was written, the messages the author may be trying to communicate, and the techniques the author might be using to communicate those messages (i.e. what literary conventions might be particularly relevant to this text) all of which are relevant to the construction of interpretations. Further work should be
conducted to encourage novice literary readers to consider where and when the text might have been written and how this might color their interpretation.

This study demonstrates the importance of interpretive inferences in literary text comprehension, but also speaks more generally to the importance of activation and integration of relevant prior knowledge for any domain. Moving students beyond basic comprehension to a deeper understand and more sophisticated reasoning about the text is an important aspect of reading in any discipline. The long-term goal of this kind of work is to support the development of readers who have the skills and knowledge base to move flexibly between domains and can draw upon a variety of strategies to engage in sophisticated reading behaviors.
REFERENCES


### Appendix A
Open-Ended Response Scoring adapted from McCarthy & Goldman (2015)

<table>
<thead>
<tr>
<th>Type</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Verbatim</strong></td>
<td>Copied directly from the text</td>
<td>The schoolchildren who had witnessed the scene in the zoo soon started neglecting their studies and turned into hooligans. It is reported they drink liquor and break windows. And they no longer believe in elephants.</td>
</tr>
<tr>
<td><strong>Paraphrase</strong></td>
<td>Rewording of the sentences from the text; Summary or combining of multiple sentences from the text</td>
<td>After seeing this the students gave up on education became drunks and stopped believing in elephants.</td>
</tr>
<tr>
<td><strong>Text-Based Inference</strong></td>
<td>Reasoning-based on information presented in the story, with some use of prior knowledge; connecting information from two parts of the text</td>
<td>After being deceived [sic] by the fake elephant, the children became poor students, and grew up behaving badly because they were lied to.</td>
</tr>
<tr>
<td><strong>Interpretive Inference</strong></td>
<td>Inferences that reflect nonliteral, interpretive interpretations of the text</td>
<td>The theme is that being lied to ends the innocence of the young boys and girls.</td>
</tr>
</tbody>
</table>
### Appendix B

**Reading instructions**

<table>
<thead>
<tr>
<th>Condition</th>
<th>Instruction</th>
</tr>
</thead>
</table>
| **Rules of Notice** | The story you are about to read is *The Elephant* by Slawomir Mrozek. When experts read literary texts, they pay attention to clues in the text that signal to them that the text has a deeper meaning. Here are three signals that experts often noticed in *The Elephant* but keep in mind that not everyone noticed the same signals or all of these.  
1) things that deviate from the norm (things or events differ from what you might expect in the real world),  
2) juxtaposition (things or events that seem to contrast each other),  
3) disruptions (unexpected word choices, things, or events).  
Consider this information as you read the story. |
| **Rules of Signification** | The story you are about to read is *The Elephant* by Slawomir Mrozek. Mrozek is known for his use of satire. Satire is a literary device that authors sometimes use to ridicule or criticize an individual or a society.  
You might be familiar with how satire operates from shows like *The Colbert Report*, *Saturday Night Live*, *Key & Peele*, and *Inside Amy Schumer*. These shows use exaggeration and humor to point out and criticize current social problems. When making sense of satiric works, you should consider what the target of this criticism might be. |
| **Combined** | The story you are about to read is *The Elephant* by Slawomir Mrozek. Mrozek is known for his use of satire. Satire is a literary device that authors sometimes use to ridicule or criticize an individual or a society.  
You might be familiar with how satire operates from shows like *The Colbert Report*, *Saturday Night Live*, *Key & Peele*, and *Inside Amy Schumer*. These shows use exaggeration and humor to point out and criticize current social problems. When making sense of satiric works, you should consider what the target of this criticism might be.  
**In order to recognize satire and to understand how it works in a particular story**, experts pay attention to clues in the text that signal to them that the text has a deeper meaning. Here are three signals that experts often noticed in *The Elephant* but keep in mind that not everyone noticed the same signals or all of these.  
1) things that deviate from the norm (things or events differ from what you might expect in the real world),  
2) juxtaposition (things or events that seem to contrast each other),  
3) disruptions (unexpected word choices, things, or events). |

*Bold added for emphasis*