# Making Interpretation Visible With an Affect-Based Strategy

Sarah Levine

*Northwestern University, Evanston, Illinois, USA*

## ABSTRACT

Experienced readers of literature are more likely than novices to identify aspects of text that are salient to literary interpretation and to construct figurative meanings and thematic inferences from literary texts. This quasi-experimental study explores the hypothesis that novice readers can be supported in constructing literary interpretations by drawing on and applying everyday interpretive practices to their readings. Specifically, an everyday affect-based practice can serve as an interpretive heuristic to support the move from a local summary to a range of figurative interpretations. The affect-based interpretive heuristic involves identifying language in a literary text that a reader feels is particularly affect-laden, ascribing valence to that language, and then explaining or justifying those ascriptions. In a four-week, classroom-based instructional intervention, a 12th-grade class from a high-poverty, low-achieving, urban high school practiced this interpretive heuristic as they read literary texts. A comparative class also engaged in a unit of literary interpretation but did not use the heuristic. Analysis of a pre- and poststudy interpretive writing task and clinical think-aloud protocols from both groups showed that students receiving the intervention made gains in interpretive responses, whereas the comparison group did not. The results suggest that explicit instruction in affect-driven interpretive heuristics can support novice readers in constructing interpretive readings of literary texts.

In a collection of essays (Coady & Johannessen, 2006), writers describe reading experiences that influenced their understandings of themselves and their worlds. One writer described her response to Toni Morrison's (2007) novel *The Bluest Eye*. The writer was stirred by Morrison's imagery of "'the garbage and the sunflowers'" and explained that reading the novel helped her build a new understanding of "the common cause of hurt and…the nature of love and forgiveness" (Coady & Johannessen, 2006, pp. 1–2).

This kind of literary reading—one in which both awareness of text and personal response contribute to the construction of general insights about human nature—is a goal that many English language arts teachers hold for their students (Agee, 2000; Yagelski & Leonard, 2002). Yet, those inexperienced in literary reasoning—herein termed *novice readers*—are instead more likely to engage in literal sense making, with a focus on surface features of short stories, poetry, and other literary texts[1] (Culler, 2002; Peskin, 1998; Rabinowitz, 1998; Scott & Huntington, 2007). Experienced readers, however, are more likely to engage in figurative sense making, connecting salient details of a text, and constructing thematic interpretations about human nature and the role of the individual in the world (Dorfman, 1996; Harker, 1994; Janssen, Braaksma, & Rijlaarsdam, 2006).

For example, in one study of literary interpretation, graduate students specializing in literature and college freshmen with no such specialization read a story about a "glittering" wedding ring that had fallen

into a "mess of leaves" at the bottom of a dirty swimming pool. During a think-aloud interview, one of the graduate students referred to that image, saying, "Now *that* sounds like a symbol…a symbol of purity rising up from corruption." In contrast, a freshman reader commented, "'This is one terrible pool. I can't believe they don't drain it'" (Earthman, 1992, pp. 366–367).

The comparison of these two responses is not meant to suggest that the expert's reading is more valuable than that of the novice, or that the novice, should aspire to the particular interpretation of the expert (Fairclough, 2001; Gee, 2008; Hull & Rose, 1990). However, the ability to construct meaningful connections between concrete imagery and abstract concepts is an important element of interpretive sense making. Constructing abstractions is necessary for interpreting metaphors and figurative language (Bowdle & Gentner, 2005; Cacciari & Glucksberg, 1994; Gibbs, 2008), as well as building thematic inferences (Brinker, 1995; Lehr, 1988). The process may also be an important part of the gap filling that occurs as readers construct models of texts (Kintsch, 1980), and the way readers project themselves into a story (Mar & Oatley, 2008) to make it both personally and globally meaningful. Teachers need accessible models and approaches to support novice readers in learning to connect the concrete to the abstract in salient ways, so that a diamond ring in a swimming pool might also be a representation of corrupted love, or "'the garbage and the sunflowers'" might also be an exploration of "the nature of love and forgiveness."

One way to do this is to leverage interpretive practices in which novice readers may already engage. For example, although the freshman reader in the swimming pool study did not construct figurative interpretations, he did move beyond pure summary in one significant way: He made an affective evaluation of the image of the pool based on imagery of the "mess of leaves" and similar details. He said, "This is one terrible pool." This study explores the idea that such affect-based evaluations, where a reader draws on context and personal knowledge to ascribe valence to text, is an everyday interpretive practice that can be made visible to novice readers. The study hypothesizes that when combined with explanation and reflection, this practice may act as a useful interpretive heuristic for novice readers, helping them move from literal to interpretive sense making.

## An Example of the Affective Evaluation Heuristic

Before examining research that informs the hypothesis about the affective evaluation heuristic, it will be helpful to describe the heuristic. In this study, students in an intervention group were taught to do the following:

1. Draw on their everyday affect-based interpretive practices to identify language in a literary text that they feel is particularly affect-laden

2. Ascribe valence (a range of positive and/or negative values) to that language

3. Explain or justify their ascriptions

In the framework of affective evaluation, valence, along with the related elements of tone and mood, is not an immutable property of a text, but instead is constructed, either automatically or strategically, by a person who is interacting with that text in a particular place and time (Brendl & Higgins, 1996; C.D. Lee, 2001; Rosenblatt, 1982; Smagorinsky, 2011). Additionally, readers' appraisals of valence are not binary. Texts can be interpreted as slightly negative, extremely positive, or some combination. The more complex texts are, the more likely they are to invite nuanced affective evaluations.

The final line of Morrison's *The Bluest Eye* can be used to illustrate the heuristic in action. The line describes a young girl searching for something "among the garbage and the sunflowers of [a] town." In the first step of the heuristic, readers identify words or phrases that they find to be affect-laden; for example, in context, a reader might find the words *garbage* and *sunflowers* to be strongly emotionally evocative, especially in comparison with the surrounding text. In the second step, readers ascribe valence to that emotionally evocative text. Probably, although not certainly, a reader might ascribe a negative value to *garbage* and a positive value to *sunflowers*, or negative *and* positive values in response to the juxtaposition of the two words. In the third step, readers explain or justify their ascriptions of valence to the affect-laden phrases. For example, a reader might explain a negative evaluation of *garbage* by constructing abstract associations of filth or being devalued. Likewise, a reader might justify a positive evaluation of *sunflowers* by drawing on concepts of light, life, beauty, or hope.

## Affective Evaluation as an Everyday Interpretive Practice

The study's hypothesis about the affective evaluation heuristic draws on three general areas of theory and research. First, it assumes that affective evaluation is an everyday interpretive practice. This assumption is supported by research suggesting that affective evaluation acts as an everyday control structure with which to make sense of the world. People readily evaluate everyday phenomena on a scale anchored by good and bad, or benefit and harm (Bower, 1992; Ortony, Revelle, & Zinbarg, 2006; Osgood, 1964; Zajonc, 1980, 1984). For

example, several studies showed that even young children tended to evaluate others' emotional states and attitudes on a positive/negative scale (Babad & Taylor, 1992; Laplante & Ambady, 2003; Nelson, 1987; Walker-Andrews, 1997).

Some models of affective response propose that such evaluations are drawn from a set of more abstract personal beliefs, goals, and cultural expectations (Clore & Ortony, 2000; Lerner, Miller, & Holmes, 1976), suggesting that individuals might be able to explain their affective evaluations in terms of concepts and abstractions. For example, in a controversy over the removal of the Confederate flag from the South Carolina state capitol, a white South Carolinian senator explicitly evaluated the flag as positive and then explained his evaluation in abstract terms: "I see courage, I see honor." Meanwhile, an African American citizen of the same state evaluated the flag negatively, explaining, "I see oppression. I see racism" (Smagorinsky, 2001, p. 135).

## Affect in Literary Processing

A second assumption of the heuristic is that the process of affective evaluation can be applied in both the real world and literary worlds. The graduate student in the swimming pool study, for example, may have drawn on positive cultural associations with wedding rings when interpreting the concrete image of the ring as a symbol of purity. Studies of theme development in fables found that readers made positive or negative appraisals of characters based on the extent to which the characters' actions aligned with readers' real-world values or expectations (Williams et al., 2002; Zillmann, 1995). If a character in a fable stole, a reader was likely to evaluate him or her negatively. If that character then encountered punishment, a reader was more likely to construct a global inference that approved such an outcome, such as "stealing is wrong" (Dorfman & Brewer, 1994). Other research similarly suggests that readers use context and their own world knowledge to evaluate fictional characters as sympathetic or unsympathetic (Gerrig & Rapp, 2004; Mar & Oatley, 2008) and to construct abstract thematic points that correspond to the positive or negative outcome of the characters' goals (Dyer, 1983; Seifert, Robertson, & Black, 1985; Zhang & Hoosain, 2001).

Much of this research may involve automatic affective evaluations of language, characters, and outcomes in both the real world and literary texts. However, the current study proposes that potentially automatic evaluations can be combined with strategic justification or explanation, and that this process can be taught explicitly as a heuristic for novice readers engaged in interpretive sense making.

## Interpretive Apprenticeships

The third assumption in this study has to do with explicit teaching of cognitive processes. The proposed heuristic assumes that to become independent interpreters of literature, students must identify and understand often implicit models of the processes of interpretation. Making implicit sense-making practices visible is part of an instructional model that has been called the cognitive apprenticeship (Collins, Brown, & Holum, 1991). In such an apprenticeship, an expert models ways of thinking in a particular domain. A few studies in literary interpretation have successfully taken up this model, explicitly instructing students in some of the interpretive practices of experienced readers, such as reading Western poetry with attention to dominant Western tropes (Levine & Horton, 2013; Peskin, Allen, & Wells-Jopling, 2010; Zyngier, Fialho, & do Prado Rios, 2007).

Another body of research has developed a framework of cultural modeling, where novices engage familiar texts to gain a meta-level knowledge of their own everyday, culturally based interpretive practices, and then apply those practices to more formal domains (C.D. Lee, 1995, 2007; Nasir, Rosebery, Warren, & Lee, 2006). For example, in one intervention, high school students were made aware of ways that they already identified and interpreted symbolism in popular songs. The students then articulated the knowledge and interpretive heuristics they used to make such everyday interpretations, and were taught to apply similar processes to more canonical texts. Empirical results in that intervention showed gains in interpretive reasoning (C.D. Lee, 2006).

The current study builds on the research and theory previously described to support novice readers in the complex processes of connecting concrete literary images and language with abstract connotations and interpretations. The following section describes the implementation of the study.

## Methods

The study explores a quasi-experimental instructional intervention in which a teacher was trained in the process of affective evaluation. The teacher then made this process visible to students by showing them ways in which they currently practiced affective evaluation, and then taught them to use affective evaluation as a heuristic for literary interpretation. To assess the efficacy of the intervention in supporting the move from literal to interpretive sense making, the study evaluated interpretive writing and think-aloud protocols tasks from both an intervention and a comparison group. In the writing task, students were asked to construct a thematic statement about a poem. In the think-aloud task,

students were asked to construct oral thematic statements in response to a one-page excerpt of a novel. They were also asked to make online interpretations of the excerpt, to allow for more fine-grained examination of use of the affective evaluation heuristic (Andringa, 1990; van Someren, Barnard, & Sandberg, 1994). The prediction was that, on the written task, students who had been taught affective evaluation would construct more thematic statements and fewer summaries in response to a poem. In the think-aloud protocols, the prediction was that students would construct more abstract interpretations in response to concrete details in the body of the novel excerpt.

## *Participants*

The study took place at an urban, public high school serving several thousand students. Approximately 86% of the students came from low-income households. At the time of the study, 29% of the students met state reading standards, and 1% exceeded standards. The majority (approximately 80%) of students were Latino, and about half of those students had learned English as a second language. Approximately 11% of students were African American, 4% Asian American, 4% Caucasian, and 1% American Indian or Alaskan Native. Students were required to take daily, 45-minute English literature classes during all four years of high school.

At the time of this study, there were approximately 170 full-time classroom teachers at the school. About 65% were women. About 62% were Caucasian, 16% African American, 14% Latino, and about 7% Asian American.

## Intervention and Comparison Teachers

A request for teacher participants for a classroom study on literary interpretation was sent to the school's English Department, and several teachers responded. Two teachers of untracked senior classes were asked to participate because both were interested in and had already begun to plan units of study in textual interpretation, with a focus on poetry and short fiction. One teacher, whose schedule aligned with that of the researcher, was asked to participate in the instructional intervention. The other teacher's class formed the comparison group. Each teacher was paid a small stipend to participate. Both teachers were female and Caucasian. The intervention teacher was 40 years old at the time of the study, with nine years of teaching experience. The comparison teacher was 33, with seven years of experience.

Prestudy interviews were conducted via e-mail to compare the teachers' pedagogical goals and philosophies (Franke, Carpenter, Levi, & Fennema, 2001; Grossman & Shulman, 1994; Schoenfeld, 2011). Both teachers stated that literary interpretation was not only for school but also an important real-world tool, and both noted that their goals included helping students understand the connection between literary techniques and themes.

## Intervention and Comparison Students

The participating students were all 12th graders. Although the total number of students in the intervention class was 28, only 19 consented to participate in the study; likewise, there were 29 students in the comparison class, but only 18 consented to be in the study. In both classes, about 75% of the students who returned consent forms were over 18 and therefore able to sign and return consent forms themselves. About 25% of students who needed a parent or guardian's signature on their consent forms returned them.

The following information refers to the participating students only. Both the intervention and comparison students were 12th graders in untracked classes. An untracked class is one that teaches a regular—as opposed to advanced or honors—curriculum. The classes were matched along two measures: class grades and a standardized reading test.[2] The standardized reading scores were separated into four categories: warning, below standards, meets standards, and exceeds standards. The average scores of both classes hovered on the border between below standards and meets standards. The intervention group had 10 boys and 9 girls and was about 80% Latino and 20% African American. The comparison group had 10 boys and 8 girls and was about 85% Latino and 15% African American. The two classes showed a similar range of English grades for the previous quarter, with similar numbers of failing (five or six), average (between seven and 10), and top grades (two or three). Both classes also saw similar levels of truancy.

Selected students from the intervention group ($n = 5$) and the comparison group ($n = 5$) participated in semi-structured reading protocols of a novel excerpt. Initially, the goal of selection was to identify students who were high-, middle-, and low performing, according to both grades and test scores. However, only a few students earned grades that correlated with their standardized test scores, making this kind of selection impossible. Instead, each group contained three students whose standardized reading scores lay in the "below standards" range, and three students who scored in the "meets standards" range. All of those students had earned either Bs or Cs during the previous quarter of the school year.

Unfortunately, before the end of the study, one student from each group dropped out of school, leaving the two groups unmatched in terms of reading scores. The intervention group ultimately included three students who achieved "meets standards" scores on the standardized reading test, whereas the comparison group included only two "meets standards" scorers. However,

as study results will demonstrate, pretest results of both the written and oral interpretation tasks showed no significant difference between the interpretive levels of "meets standards" and "below standards" students, either within or between groups, suggesting that the standardized reading scores were not a factor in preintervention interpretive performance. In addition, during the four weeks of the study, the grades of students in both groups wavered between Bs and Cs, suggesting some additional similarity in skills and engagement levels both within and between groups.

The comparison and intervention groups reflected the ethnic and racial makeup of the classes as a whole. Each group contained three female students and two male students. All students were either 17 or 18 years old.

## Materials

### Assessments

Two kinds of assessments were given to students. A poem was used for a pre–post writing task, given to all students in each group. An excerpt from a novel was given to a representative subgroup from each class for a pre–post think-aloud.

### Writing Task Materials

Li-Young Lee's (1986) poem "I Ask My Mother to Sing" was used for the pre–post writing task given to the full intervention ($n = 19$) and comparison groups ($n = 18$). This poem portrays bittersweet images of a family remembering a lost loved one as they sing an old song about their home country. Even though the singing makes the family weep, they continue to sing. The poem was chosen because it was widely anthologized and offers a range of interpretations (Gillan & Gillan, 1994; Xiaojing, 1996). It is also short (105 words) and readable, measuring at the Flesch–Kincaid fourth-grade level. Both teachers judged the poem to be accessible to their classes. The text was unfamiliar to both groups of students.

For the purposes of comprehension, a note was added to the poem to identify some of the place names in the poem. Also, in a line describing "picnickers running away in the grass," the word *picnickers* was defined.

### Think-Aloud Materials

An adapted excerpt from the novel *Prisoner's Dilemma* by Richard Powers (1996) was used for both pre- and post-study reading protocols with the selected students from the intervention ($n = 5$) and comparison groups ($n = 5$; see Appendix A). The excerpt describes a father and his children looking at the stars on a cold night. The father wants to teach his children facts about the constellations. The children long for warmth and human closeness, which, ultimately, they do not get. This particular excerpt was first used on an International Baccalaureate literature exam several years earlier, which indicated that it was complex enough to warrant in-depth literary interpretation.

For the purposes of the think-aloud task, the International Baccalaureate excerpt was shortened from 739 words to 554. In addition, a few words that students in both groups were unlikely to know were revised in order to limit student distraction or concern about unknown vocabulary. For example, the text references a "six-volt flashlight beam" (Powers, 1996, p. 13) that the father uses to point to the stars. This phrase was changed to "cheap flashlight" because both teachers felt that few of their students would understand the original phrase. The edited excerpt measured at the Flesch–Kincaid seventh-grade reading level and was judged by the intervention and comparison group teachers to be accessible to their students. The story in all its forms was unfamiliar to both classes of students.

### Instructional Unit Materials

Both teachers drew from popular and canonical texts during instruction, including songs, pictures, and short stories, and in the comparison group's case, a short novel. See Table 1 for the titles of every text used in each teacher's unit, along with brief descriptions of the activities designed for each text. (See Appendix B for a more detailed description of each unit.)

### Matching Texts

Because the choice of texts is an important factor in student learning (Friese, Alvermann, Parkes, & Rezak, 2008; Ivey & Broaddus, 2001; C.D. Lee, 2007), and because the two groups used different texts, all texts used in each group were compared in terms of reading difficulty and potential for supporting literary interpretation.

In terms of reading difficulty, reading levels were analyzed with the widely used Flesch–Kincaid readability test (Gunning, 2003).[3] Although this test cannot account for many facets of text complexity (C.D. Lee, 2011), it provides one a useful measure of comparison. With one exception, all texts in both groups ranged from the third- to seventh-grade reading levels.

The texts were also compared in terms of their capacity to support literary interpretation, including the extent to which they supported multiple interpretations. In the intervention group, Richard Wright's *Black Boy*, Sandra Cisneros's *The House on Mango Street*, and Junot Díaz's short stories have been the subject of much scholarly literary interpretation, and presented many opportunities for students to consider literary choices and themes (Bloom, 2009; Gaskill, 1973; Hakutani, 1985; Wissman,

**TABLE 1**
**Texts and General Activities in Each Unit of Study**

| Timeline | Unit elements | Intervention classroom | Comparison classroom |
|---|---|---|---|
| Prestudy | Prestudy assessment | • Written task in response to the poem "I Ask My Mother to Sing" by Li-Young Lee[a]  <br>• Think-aloud in response to an excerpt from the novel *Prisoner's Dilemma* by Richard Powers[b] | |
| Day 1 | Introduction | • Discussion of the connection between everyday and school interpretation  <br>• Examination of language that creates specific effects | • Discussion of the power of literature  <br>• Examination of language that creates specific effects |
| Days 2 and 3 | Discussion about effects of language | • Headlines and print ads  <br>• Short story: Short version of Cinderella (created by the teacher) | • Article: "Justin Bieber: Giving Teens the Wrong Idea" (created by the teacher)  <br>• Pop song: "Baby" sung by Justin Bieber[c] |
| Days 4 and 5 | Writing argument about effects of language | Argument about symbolism and effects in the short version of Cinderella | Argument about the effects of language in pop songs |
| Day 6 | Constructing interpretations with a focus on symbolism | Novel excerpt: *Black Boy* by Richard Wright[d] | Poem: "Harlem Sweeties" by Langston Hughes[e] |
| Days 7 and 8 | Constructing interpretations by comparing two texts with similar themes but different styles | Pop songs:  <br>• "I Will Follow Him" sung by Little Peggy March[f]  <br>• "I Will Follow You Into the Dark" by Death Cab for Cutie[g] | Poems:  <br>• "Delight in Disorder" by Robert Herrick[h]  <br>• "The Shirt" by Jane Kenyon[i] |
| Days 9 and 10 | Writing argument about literary effects | Argument about symbol and themes in love songs | Argument about symbol and themes in love poems |
| Days 11–13 | Discussion and activities for interpretation of theme of fiction | Novella excerpt: "Linoleum Roses" from *The House on Mango Street* by Sandra Cisneros[j] | Novella: *Like Water for Chocolate* by Laura Esquivel[k] |
| Days 14–16 | Discussion and activities for interpretation of theme of fiction | Short story: "Ysrael" by Junot Díaz[l] | Novella: *Like Water for Chocolate* by Laura Esquivel |
| Days 17 and 18 | Writing argument about literary effects | Argument about symbol and theme in "Ysrael" | Argument about symbol and theme in *Like Water for Chocolate* |
| Poststudy | Poststudy assessment | • Written task in response to the poem "I Ask My Mother to Sing"  <br>• Think-aloud in response to the excerpt from the novel *Prisoner's Dilemma* | |

[a]Lee, L.-Y. (1986). I ask my mother to sing. In *Rose* (p. 50). Rochester, NY: BOA Editions.
[b]Powers, R. (1996). *Prisoner's dilemma*. New York, NY: HarperPerennial.
[c]Nash, T., Stewart, C., Bieber, J., Milian, C., & Bridges, C. (2010). Baby [Recorded by J. Bieber]. On *My world 2.0* [CD]. Atlanta, GA: Island RBMG.
[d]Wright, R. (2005). *Black boy*. New York, NY: HarperCollins.
[e]Hughes, L. (1995). Harlem sweeties. In *The collected poems of Langston Hughes* (pp. 245–246; A. Rampersad & D. Roessel, Eds.). New York, NY: Knopf.
[f]Plante, J. (1963). I will follow him [Recorded by Little Peggy March; N. Gimbel, Trans.]. On *Wind up doll*. Nashville, TN: RCA Victor.
[g]Gibbard, B. (2006). I will follow you into the dark [Recorded by Death Cab for Cutie]. On *Plans* [CD]. New York, NY: Atlantic Records.
[h]Herrick, R. (2004). Delight in disorder. In M.W. Ferguson, M.J. Salter, & J. Stallworthy (Eds.), *The Norton anthology of poetry* (5th ed., p. 355). New York, NY: W.W. Norton.
[i]Kenyon, J. (2005). The shirt. In *Collected poems* (p. 34). Saint Paul, MN: Graywolf.
[j]Cisneros, S. (1991). Linoleum roses. In *The house on Mango Street* (2nd ed., p. 101). New York, NY: Vintage Contemporaries.
[k]Esquivel, L. (1995). *Like water for chocolate: A novel in monthly installments with recipes, romances, and home remedies* (C. Christensen & T. Christensen, Trans.). New York, NY: Anchor.
[l]Díaz, J. (1997). Ysrael. In *Drown* (pp. 1–20). New York, NY: Riverhead.

2007). In the comparison group, the poems by Langston Hughes and Robert Herrick are widely anthologized (e.g., Ferguson, Salter, & Stallworthy, 2004; Lehman, 2006) and have also been the subject of academic study and literary interpretation (Davis, 1952; Ingram, 1998; Schultz, 2002), as has Laura Esquivel's *Like Water for Chocolate* (de Valdés, 1995; Dobrian, 1996). Hence, although the texts were not matched in terms of content or style, both sets of texts offered many avenues for student interpretation of techniques and effects.

## Procedures

### Participant Protection

When students were invited to participate in the study, they were told that participation was voluntary. They were assured that their grades and evaluations would not be affected by their participation status and that all students, regardless of participation, received the same instruction and assignments.

Students who did not consent to be part of the study were not videotaped during the study, and none of their discussions or written work was used as data in the study. Students who consented to participate were filmed, and their work was collected. All work was de-identified for scoring, and no names were used in this study.

### Writing Task

The writing task was given before and after the study period to all students in the intervention ($n = 19$) and comparison classes ($n = 18$), including the students selected to participate in the think-alouds. Using the same text allowed for precise within- and between-group comparisons, and if there were carryover effects from one reading to the next, they would occur for both groups.

After receiving "I Ask My Mother to Sing," students were told that they should read the poem as many times as they wished, write on it if they wished, and then respond to the two questions at the bottom of the page. The first question was "Please summarize the action in this poem. What is happening in this poem?" The second question, adapted from the 2009 College Board Advanced Placement English Literature exam, was interpretive in nature: "Using what you know about understanding literature, please discuss your ideas about the themes or underlying meanings of the poem as a whole. Please write at least two sentences." Students had 45 minutes to complete this task at both pre- and poststudy. The text was not discussed during the study period.

### Think-Aloud Task

The think-aloud interviews were semistructured, lasting from 30 to 35 minutes. All interviews were conducted by the researcher. At both pre- and poststudy, the participating students were told that they would be reading a story aloud, and as they read, they should "use any strategies or ways you know to make sense of the story and make interpretations of deeper meanings, symbolism, or other meanings beyond the literal." The students were asked to share all of their thoughts aloud as they read. They were told that if they were uncertain about a word, the interviewer would define it for them. They were also told that they should feel free to reread and write on the text if they wished. The researcher then modeled a think-aloud using an algebra problem as an example.

Next, the students were told that at the end of the think-aloud, they would be asked to summarize the story and talk about its themes. They were reminded that *theme* could be defined as their ideas about a story's underlying meanings, big ideas, or worldviews about society or human nature. These definitions were adapted from those used in studies of thematic inference (Kurtz & Schober, 2001; Lehr, 1988) and were similar to the definitions that intervention and comparison teachers used in their classes.

At prestudy, students were given no instructions about when to comment in order to minimize interference with their thought processes or attention to particular details (Ericsson & Simon, 1984). However, if students read more than one paragraph without comment, indicated special attention to a detail by pausing or annotating, or indicated general uncertainty about how to proceed, they were encouraged to speak their thoughts out loud (Ericsson & Simon, 1984). The interviewer made this encouragement via a general prompt ("What are you thinking?") or an interpretive prompt ("Use strategies or ways you've learned in class or anywhere else to help you interpret what you're reading").

At the end of the pre- and postintervention think-alouds, both groups were asked to describe the literal sequence of events in the story. Then, they were asked to construct a thematic statement in response to the same prompt used in the written task: "Using what you know about understanding literature, please discuss your ideas about the themes or underlying meanings of the story as a whole."

### Training for the Intervention

The intervention teacher was trained in and practiced affective evaluation in three meetings (two hours each) that took place over a two-week period before the start of the school year. In this training, the researcher first described some of the everyday ways that individuals use affective evaluation. Then, the researcher and teacher identified the components of affective evaluation, and the teacher identified instances in which she used affective evaluation in everyday life. Next, the teacher was taught to use the steps of affective evaluation to examine positive and negative connotations of newspaper headlines, short poems, and then short stories. The texts used during training were not the same as those used in the intervention (simply because the teacher had not yet chosen them). Building on the principles of the cognitive apprenticeship (Collins et al., 1991) and cultural modeling (C.D. Lee, 1995, 2007), the teacher then practiced selecting texts that lent themselves to affective evaluation so students could more easily "see" the practice. Finally, the teacher practiced

the explicit modeling and meta-level discussion of this interpretive process, with the researcher playing the role of student.

The comparison teacher did not receive training in explicit modeling or affective evaluation. However, the researcher and comparison teacher met three times before the start of the study period to discuss this the teacher's selection of texts that lent themselves to literary interpretation.

## Instructional Design and Implementation

Both teachers designed units to accommodate their own goals as well as departmental and school requirements. As described earlier, both were interested in teaching skills in literary interpretation and, specifically, symbolic interpretation. In addition, the teachers wanted to integrate those skills with exploration of a specific topic. The intervention teacher wanted her students to read texts that explored different aspects of gender and coming-of-age, and the comparison teacher wanted to do the same with the topics of gender and romantic love. Finally, both needed to fulfill departmental requirements for the teaching of argumentative writing, using an argument template that had been adopted schoolwide. To accommodate these requirements, both teachers had their students write arguments about ways that symbolism in some of the chosen texts shed light on the themes of those texts.

Each unit lasted 4½ weeks, with a total of 18 class periods of instruction. In both groups, the lessons generally stretched across several days of class; for example, a lesson involving interpretation of symbolism in a short story might take two or three days. The comparison teacher provided the researcher with weekly lesson plans in advance of the actual lessons, and to the extent possible, the intervention unit followed the structure of the comparison unit. For example, over the course of her unit, the comparison teacher asked her students to write three short arguments about symbolism's effects on theme, so the intervention teacher did the same.

## Assessing Fidelity to Treatment

Classroom videos, field notes, and lesson plans from the intervention class were examined by the author for the presence or absence of use of the affective heuristic. This examination showed that the teacher introduced the affective evaluation heuristic to her students in much the same way that she herself was introduced to it: by modeling the process of affective evaluation, identifying its components, and having students identify instances where they might already use that interpretive practice outside of school. In 14 out of 18 days, the intervention teacher prompted her students to name the steps of the affective evaluation heuristic or, after

reading a passage in class, asked or quizzed the students about how they might begin interpretation.

The videos, notes, and daily lesson plans for the comparison teacher were examined for the presence or absence of teaching related to literary interpretation. The comparison teacher led whole-class discussions about symbolism and overall thematic meaning 13 out of 18 days, including the tracking of several symbols and their effects in poetry and the short novel *Like Water for Chocolate*. On two days, the teacher explicitly modeled an interpretive strategy: looking for patterns in texts and then asking what those patterns could represent. However, as mentioned earlier, students were given no explicit instruction or support in *how* to construct representations once they had identified patterns.

Although students discussed their emotional responses to several of the unit texts, and likewise discussed their interpretations of moods or tones, they did not use the affective evaluation heuristic.

## Coding of Student Responses

There are three sources of data in this study: (1) students' ($N = 37$) written thematic statements in response to a question about overall themes of the poem "I Ask My Mother to Sing," (2) representative students' ($n = 10$) oral thematic statements in response to a question about the overall themes of the excerpt from *Prisoner's Dilemma*, and (3) those same representative students' concurrent think-aloud responses to individual propositions in the excerpt.

In the reading protocols, both the students and the teachers made their responses to the text almost exclusively at the ends of sentences or independent clauses. To examine and compare readers' responses, each independent clause in the story was marked and numbered as a separate proposition, along with two dependent clauses to which several readers responded, for a total of 61 propositions. A reader's response was defined as the sum of statements made after any one proposition from *Prisoner's Dilemma*. One response could be a single phrase or several sentences.

The author and an independent rater with a doctorate in literature and high school teaching experience coded student responses. They used data not included in the actual study to practice and align coding according to the criteria subsequently described.

### Use of Affective Evaluation

The intervention group's reading protocol transcripts were examined for evidence of use of affective evaluation. Responses were coded as having used the strategy if students explicitly responded to propositions using the terms *positive* or *negative* (or synonymous terms used in class, e.g., *thumbs-up* or *thumbs-down, happy* or

*sad*) and then explained their ascriptions of valence. Although the comparison students had not been trained in the strategy, their transcripts were also examined for any instances of affective evaluation.

Because the intervention students were taught to use affective evaluation in class, it is unsurprising that they did so. Measuring use of affective evaluation acted as an additional data point in determining efficacy of the heuristic.

### Interpretive and Literal Responses

Because this study explores the extent to which students moved from summary to theme, or literal to interpretive sense making, coders used an a priori set of codes that characterized the written and oral statements as either literal or interpretive (see Table 2). These codes were adapted from several studies of students' interpretations of poetry and fiction (Eeds & Wells, 1989; Lehr, 1988; Svensson, 1987; Vipond & Hunt, 1984). In a literal response, a reader offered a summary or rephrasing of parts or all of the text in question. In an interpretive response, a reader derived or constructed at least one abstract or conceptual inference that could be supported in some way by the features of the text. Such interpretations might include inferences about characters' motivations or feelings, symbolic interpretations, or construction of connotations of concrete imagery. If a response included a local or global abstraction, it was coded as interpretive.

All written thematic statements (74) and oral thematic statements (20) were coded by the author and the independent rater. All online think-aloud responses to individual propositions (333) were coded by the author; to test reliability, 30% of the online think-aloud responses were coded by the independent rater. Both the author and the independent rater were blind to the condition associated with the responses. However, a few student responses included words specific to the intervention such that raters were made aware of the condition. Inter-rater agreement was 84% (weighted Cohen's Kappa = .70).

### Other Responses

For coding of the online think-aloud protocols, a third category called "other" was added to the coding scheme.

Responses coded as other included generic comments about the text itself, such as, "That word is spelled wrong"; questions about word meaning; and comments that could not be linked to either interpretive or literal processing, such as, "I don't know what this could be." Four student responses identified literary devices without further analysis (e.g., "That's a simile"). These responses were also coded as "other".

There were no responses to the writing task that were coded as "other", so this category was only used in coding the think-aloud responses to individual propositions from *Prisoner's Dilemma*.

## Data Analyses

### Affective Evaluation

The frequency of use of affective evaluation in the online think-aloud responses was measured with an ANCOVA. The use of affective evaluation was not measured for the written or oral thematic statements because they did not make clear the use of the heuristic. For example, whereas the online think-aloud responses included words such as *thumbs-up* or *negative*, the written responses did not.

### Writing Task Thematic Statements

The thematic statements (74) from the writing task were coded categorically as either literal or interpretive in nature. Because the number of literal and interpretive responses were similar at pretest, and because the frequency of responses in any one category was small, a Fisher's exact test was used to compare the frequencies of interpretive responses within each group. The comparative frequencies of interpretive statements were then submitted to ANCOVA. Note that participants in the writing task only received one score at pretest and one score at posttest.

### Think-Aloud Thematic Statements

Because there were only 20 thematic statements from the think-aloud protocols (one for each student at prestudy and one at poststudy), the statements were examined for trends only, and no quantitative data were analyzed.

TABLE 2
Examples of Codes for Responses to an Adapted Excerpt From *Prisoner's Dilemma* by Richard Powers[a]

| Proposition from *Prisoner's Dilemma* | Student response coded as literal | Student response coded as interpretive |
| --- | --- | --- |
| "I think I can make out the Milky Way." (p. 14) | "It's a constellation he can see." | "This shows he's trying really hard to please his dad by seeing what he sees." |
| "My father points a [cheap flashlight] at the holes in the enclosing black shell." (p. 13) | "OK, so he's shining the flashlight up at the sky." | "It's like light, which is like knowledge. He's trying to give knowledge, but he doesn't actually have very much." |

[a]Powers, R. (1996). *Prisoner's dilemma*. New York, NY: HarperPerennial.

## Think-Aloud Responses to Individual Propositions

In the think-alouds, each student made multiple responses and, therefore, had multiple literal, interpretive, or "other" codes. Each student's literal, interpretive, and "other" think-aloud responses were counted and then converted into proportions of that student's total responses. Because the raw proportions violated assumptions about the normality of the distribution of data required for the statistical measures used in this study, an arcsine transformation was applied to the proportions prior to analysis (Judd, McClelland, & Ryan, 2009; Olson, 1976). The statistical results that follow are based on the transformed data.[4] For ease of interpretation, however, untransformed frequencies, means, and standard deviations are reported throughout.

# Results

The results below address the overarching research question of the study: To what extent does the affective evaluation heuristic support the move from literal to interpretive sense making? This question is examined by measuring the number and proportion of interpretive responses in both the whole-class writing task and the small-group think-aloud interviews.

## Use of Affective Evaluation

Before examining the results of the pre–post written and oral interpretation tasks, it was necessary to determine whether the intervention students were in fact using affective evaluation and also whether the comparison students used it, either automatically or strategically. Before the intervention, no responses in the intervention group and 1% of all responses in the comparison group were coded as having used the affective heuristic. After the intervention, 63% of all responses in the intervention group and 1% of all responses in the comparison group were coded as such. Mean scores were calculated based on proportions of affective responses for the intervention and comparison groups and submitted for ANCOVA.

After controlling for prestudy performance, the analysis revealed a significant main effect of condition, $F(1, 9) = 39.93$, $p < .001$, $\eta_{partial}^2 = 0.85$, with the proportion of poststudy affective evaluation responses higher in the intervention group than in the comparison group. Follow-up paired sample $t$-tests, which examined each group's interpretive gains across timepoints, showed that the comparison group's gain in affective responses from pre- to poststudy was not significant ($t[4] = 0.71$, $p = .519$), whereas the intervention group gain was significant ($t[4] = 6.34$, $p = .003$). Thus, the results showed that after the intervention, the intervention group made frequent use of affective evaluation during their reading protocols, whereas the comparison group did not. Again, as the intervention students were explicitly taught to use the affective heuristic, this finding is unsurprising.

## Whole-Class Writing Task: Thematic Statements in Response to the Poem

Students completed a writing task before and after the intervention. At each task point, each student's response was coded as either literal or interpretive. No responses were coded as "other," so that category was not reflected in the reporting of results. Thus, a total of 74 written responses were coded. Table 3 shows the frequency and percentages of literal and interpretive responses for students in the intervention and control classrooms by time (pre–post). The results of the whole-class writing task indicate that the intervention and comparison groups made similar numbers of literal and interpretive thematic statements at prestudy as indicated by a two-tailed Fisher's exact test: $P = 0.737$. At poststudy, however, a two-tailed Fisher's exact test yielded $P = 0.008$, indicating a significant difference in literal and interpretive responses between students in the intervention and comparison conditions, where the intervention group showed significant gains in interpretive statements, but the comparison group did not (see Table 3). This result suggests that in terms of constructing an overall thematic statement for a text, the intervention class as a whole benefited from use of the affective evaluation heuristic.

From pre- to poststudy, the intervention students showed a shift from local, concrete summaries to more global, abstract thematic statements. For example, at pretest, an intervention student wrote, "I think the meaning of this poem is that they remembered certain places, enjoyable places. They want to revisit those places." At posttest, the student wrote, "Sharing good

**TABLE 3**
**Frequency, Percentages, and Types of Responses to the Poem "I Ask My Mother to Sing" by Li-Young Lee[a] by Condition and Time**

| Responses | Intervention group (n = 19) | | Comparison group (n = 18) | |
|---|---|---|---|---|
| | Prestudy | Poststudy | Prestudy | Poststudy |
| Literal | 8 (42%) | 1 (5%) | 6 (33%) | 8 (44%) |
| Interpretive | 11 (58%) | 18 (95%) | 12 (67%) | 10 (56%) |
| Total | 19 | 19 | 18 | 18 |

*Note.* The table includes the number of literal and interpretive thematic statements written in response to the poem at pre- and poststudy.
[a]Lee, L.-Y. (1986). I ask my mother to sing. In *Rose* (p. 50). Rochester, NY: BOA Editions.

memories is beneficial to your children. This helps them have a better understanding of who you are."

It is worth noting two other trends that emerged in the students' thematic statements. First, when comparison and intervention students constructed global responses at pretest, they tended to formulate statements that took the form of inspirational clichés or adages that could be applied to many literary texts, regardless of content. For example, one intervention student whose prestudy thematic statement was coded as interpretive wrote, "This [poem] confused me but I think they are trying to say that…life goes on." Although this thematic statement can find support in the poem "I Ask My Mother to Sing," it does not necessarily indicate understanding of the poem's particular themes. However, at posttest, the same student, along with several other intervention students, moved from a generic thematic statement to a more nuanced interpretation. For example, the same student wrote, "This text celebrates those who never give up regardless of what they've been through. It celebrates those who are united and act like a family, who let no one break them apart." This thematic statement is universal while being more attuned to the specific themes of this poem.

## Small-Group Think-Aloud Responses

### Think-Aloud Thematic Statements

Additionally, two subgroups of five students were selected in the intervention and comparison classrooms to participate in semistructured think-aloud interviews in response to the excerpt from *Prisoner's Dilemma*. Each student made one oral thematic statement at prestudy and one at poststudy, for a total of 20 oral thematic responses. In this case, almost all intervention and comparison students constructed abstract, universal thematic responses at prestudy. As was true in the written responses, the prestudy thematic statements in both groups could be characterized as inspirational clichés or moral lessons that could arguably be supported by at least some important features of the text but sometimes did not account for contradicting events in or resolutions of the text. For example, one student articulated that a theme in *Prisoner's Dilemma* was that "your love for your family will never die." Such an interpretation takes into account a depiction of potential closeness between the family at the beginning of the piece but not of the distance and coldness portrayed at the conclusion. Other prestudy thematic statements were simply very vague and, therefore, vaguely applicable to *Prisoner's Dilemma*, such as a student's statement that "the story shows learning about life."

At poststudy, the comparative students' thematic statements remained at the general level, with one exception. In the intervention group, four out of the five

intervention students' thematic statements could be characterized as having shifted away from the application of clichés or rousing moral lessons and toward interpretations that accounted for more of the connotations and abstract associations made with specific features of the text. For example, at prestudy, one intervention student offered this theme for *Prisoner's Dilemma*: "You have all the time in the universe, but don't waste it…if you want to say something now, say it." At poststudy, the student formulated a more specific interpretation of a possible worldview of the text: "This text creates a mood of loneliness and rejection based on [the idea] that we are alone out here 'on a sliver of rock' in the 'black vacuum.'" Although it is difficult to draw more general conclusions from this small sample, it is worth further study to determine whether affective evaluation might support the construction of richer thematic inferences.

## Think-Aloud Responses to Individual Propositions

The bulk of the data in this study comes from the 333 online think-aloud responses of the selected students from each condition. In the intervention group, there were a total of 73 responses at prestudy (mean [$M$] = 14.60, standard deviation [$SD$] = 7.50, range = 9–27) and 111 responses at poststudy ($M$ = 22.02, $SD$ = 8.40, range = 14–35); in the comparison group, there were 75 responses at prestudy ($M$ = 15.00, $SD$ = 7.58, range = 6–27) and 74 responses at poststudy ($M$ = 14.80, $SD$ = 7.66, range = 7–27). The results of analysis of this data show that at prestudy, the groups were similar both in terms of number of responses and proportions of interpretive responses. Specifically, before the intervention, 21% of all responses in the intervention group and 24% in the comparison group were coded as interpretive. The rest of the responses were coded as follows: In the intervention group, 64% of the responses were coded as literal and 15% as "other"; and in the comparison group, 61% were literal responses, and 15% were "other."

At poststudy, however, the intervention group made significant gains both in the number of overall responses and the proportion of interpretive responses. Specifically, 80% of all responses in the intervention group were interpretive. In the comparison group, 46% were interpretive. Interestingly, the comparison group showed a decrease in "other" responses at poststudy (5%), whereas the intervention group did not (3%).

Mean scores were calculated based on the proportions of interpretive, literal, and "other" responses for the intervention and comparison groups (see Table 4).

### Interpretive Responses

The proportions of interpretive think-aloud responses of the small group of students were submitted to ANCOVA,

**TABLE 4**
**Interpretive and Literal Responses: Means and Standard Deviations for Proportions**

| | Intervention group (*n* = 5) | | | Comparison group (*n* = 5) | | |
|---|---|---|---|---|---|---|
| | Prestudy | Poststudy | | Prestudy | Poststudy | |
| Responses | *M* (*SD*) | *M* (*SD*) | Hedge's *G* | *M* (*SD*) | *M* (*SD*) | Hedge's *G* |
| Interpretive | .25 (.19) | .83 (.16) | 2.98 | .26 (.18) | .38 (.30) | 0.44 |
| Literal | .60 (.17) | .13 (.14) | 2.73 | .60 (.15) | .59 (.29) | 0.04 |
| "Other" | .15 (.10) | .05 (.03) | 1.22 | .14 (.11) | .04 (.06) | 1.02 |

*Note.* Means and standard deviations are presented for the proportions of responses out of total responses. *G* values between 0.2 and 0.5 constitute a small effect, between 0.5 and 0.8 a medium effect, and 0.8 or greater a large effect. Some totals exceed 1 because of rounding.

with the proportions of interpretive responses at prestudy as the covariate. After controlling for prestudy performance, the analysis revealed a significant main effect of condition, $F(1, 9) = 10.54$, $p = .014$, $\eta_{partial}^2 = 0.60$, indicating large effects. Specifically, 66% of the between-subject variance is accounted for by condition. As shown in Table 4, there were higher proportions of interpretive responses in the intervention group than in the comparison group. Follow-up paired sample *t*-tests, which examined each group's interpretive gains across timepoints, showed that the comparison group's gain in interpretive responses from pre- to poststudy was not significant ($t[4] = 1.08$, $p = .343$), whereas the intervention group gain was significant ($t[4] = 11.36$, $p < .001$). These results indicate that the heuristic supported intervention students in the move from literal to figurative sense making *during* reading as well as in the thematic statements made *after* reading.

### "Other" Responses
The results of ANCOVA for "other" responses showed that after controlling for prestudy performance, no effects were found for condition: $F(1, 9) = 0.06$, $p = .808$, $\eta_{partial}^2 = 0.01$. Paired sample *t*-tests measuring changes in the proportion of each student's "other" responses showed a significant decrease for the comparison group ($t[4] = 3.36$, $p = .028$) but not for the intervention group ($t[4] = 10.91$, $p = .129$). See Table 4 for means and standard deviations. Because the numbers of "other" responses in each group are so small, it is difficult to pinpoint a relationship between affective evaluation and changes in "other" comments. As conjecture, it is possible that reading through an affect-based lens at poststudy led the intervention students to focus on different details, which led to new questions or comments.

### Student-by-Student Responses
To know if only one or two intervention students were responsible for the group gains, the data were also examined on a student-by-student basis (see Figure 1). Further, students were identified by standardized test score (either "below standards" or "meets standards"). This identification provided information about whether mismatch between standardized test scores in the groups was responsible for the differences in levels of interpretive response.

The prestudy proportions of interpretive responses were comparable across the comparison and intervention groups, both for the "below standards" students ($t[3] = 0.14$, $p = .907$) and for the "meets standards" students ($t[3] = 0.61$, $p = .585$), suggesting that students in each group were comparable in interpretive skills at prestudy. Interestingly, there was also no significant difference in the "below standards" and "meets standards" responses *within* the comparison group ($t[3] = 1.02$, $p = .418$) or the intervention group ($t[3] = 0.20$, $p = .870$).

In terms of poststudy interpretive responses, the five intervention students made comparable gains regardless of their standardized test scores. In the comparison group, the two "meets standards" students made gains in interpretive response, whereas the "below standards" students did not. Again, it is difficult to make any inferences given the small sample size, but it is worth noting the distinction in performance between the "meets standards" and "below standards" groups. To the extent that standardized tests are reflections of reading skills, the differences suggest (a) that more competent readers are more likely to show growth regardless of teaching method and (b) that readers who struggle the most may benefit the most from lessons in which everyday interpretive processes are made visible.

### Frequency of Responses
Although the study focuses on increases in interpretive responses, the changes in the number of total responses were calculated as well. Each intervention student made more total responses at posttest than at pretest, with a Wilcoxon signed-rank test showing the change to be significant ($z = .04$). In the comparison group, the number of responses increased for one "below standards" student, decreased for another, and remained stable for the remaining three, including the two "meets standards" students.

FIGURE 1
Proportions of Interpretive Responses at Pre- and Poststudy

FIGURE 1
Proportions of Interpretive Responses at Pre- and Poststudy
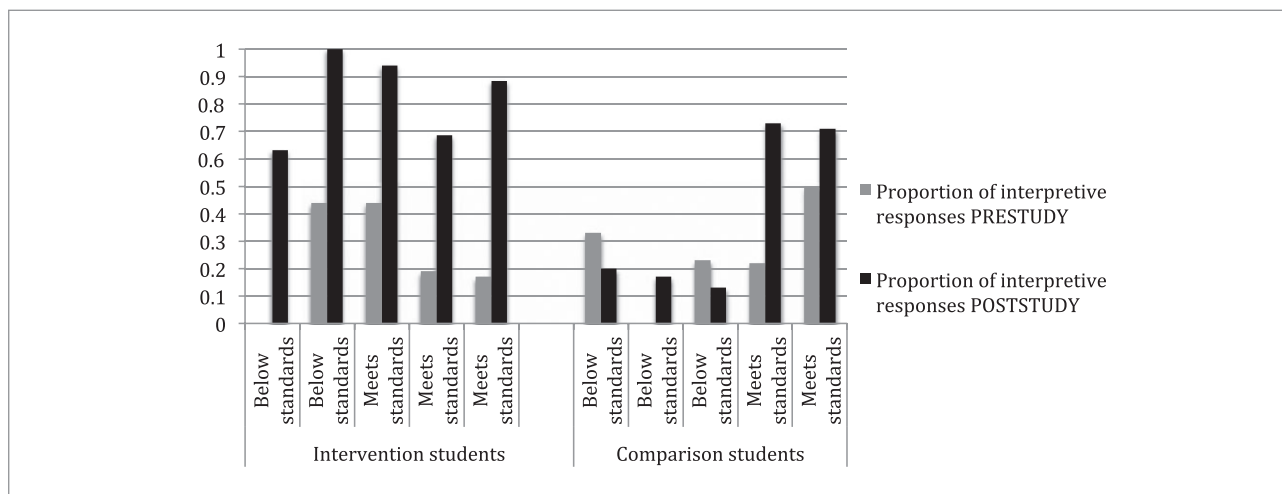
*Note.* Each student is identified by standardized reading score.

Overall, the pattern of results suggests that the affective evaluation strategy was successful in helping students in the intervention group construct interpretive responses to *Prisoner's Dilemma* regardless of their performance on a standardized reading test.

*Interviewer Prompts*

Ratios of general prompts (e.g., "What are you thinking?") to student responses were calculated to identify any correlations. In addition, interpretive prompts— those that included explicit encouragement to use strategies for interpretation—were compared with the number of interpretive responses in each group (see Table 5). This analysis was done in part to determine whether the interviewer, who knew the condition of each student during the think-aloud protocols, had unwittingly made more interpretive prompts to the intervention group.

GENERAL PROMPTS. At prestudy, the interviewer gave general prompts to the comparison group more often than to the intervention group; however, the groups had similar numbers of total responses. At poststudy, the interviewer prompted the comparison group at a slightly higher rate than the intervention group, but that higher rate did not correlate to a higher number of responses from the comparison group.

INTERPRETIVE PROMPTS. Similarly, the comparison group received a higher rate of interpretive prompts than did the intervention group in both pre- and poststudy think-aloud protocols. However, the intervention group showed a similar number of interpretive prompts at pretest and a much higher proportion of interpretive responses at posttest. Both of these results indicate that interviewer prompting was not a factor in the gains made by the intervention group.

# Discussion

The results of this study indicate that affective evaluation was a useful interpretive strategy for these novice readers of literature. Novice readers in this study made gains in constructing abstract connotations in relation to specific, concrete details while reading, as well as in overall thematic inferences after reading. It is worth exploring why the heuristic may have worked and how it can be further explored.

TABLE 5
Ratio of Prompts to Responses in Think-Aloud Interviews for the Intervention and Comparison Groups

| Item | Intervention group (*n* = 5) | | Comparison group (*n* = 5) | |
|---|---|---|---|---|
| | Prestudy | Poststudy | Prestudy | Poststudy |
| Ratio of general prompts to total responses | 1:8 | 1:5 | 1:5 | 1:4 |
| Ratio of interpretive prompts to interpretive responses | 1:8 | 1:7 | 1:5 | 1:4 |

*Note.* This table shows the number of prompts across pre- and poststudy for the comparison and intervention groups, and the ratio (rounded to the nearest whole number) of prompts to student responses across time and group.

Research indicates that unless prompted otherwise, readers tend to build representations of texts by strategically drawing from easily available associations (McKoon & Ratcliff, 1992; Paivio, 1990; Schwanenflugel, Akin, & Luh, 1992). For novice readers, those associations are more likely to relate to surface features of the text, and further, the associations may not be useful in building coherent models (Rapp, van den Broek, McMaster, Kendeou, & Espin, 2007).

That research is supported by the prestudy think-aloud protocols of the intervention students. For example, one intervention student read a line from *Prisoner's Dilemma* in which the father tries to teach his children about space by pointing a "cheap flashlight" at the sky. The flashlight's beam only goes "a few feet before it is swallowed up in the general black" (Powers, 1996, p. 13). In response to the line, the student seemed to focus on the word *cheap*, saying, "Oh, that's like the Dollar Store or whatnot." Here, the link between "cheap flashlight" and "Dollar Store" is reasonable but does not necessarily add to an interpretive model of the text.

After the intervention, in his poststudy reading protocols, the student again responded to this line, this time using affective evaluation. The heuristic may have constrained his response so that he found associations such as "Dollar Store" less salient. At poststudy, the student said, "It [the image] is negative. The flashlight isn't making an impact. It seems to be like a guide because [the family] thought it can reach all the way to the stars, but it's really not all that powerful." In cognitive terms, it may be that a different set of semantic nodes was activated when this image was encoded through an affective lens (Collins & Loftus, 1975; Collins & Quillian, 1969). Using affective evaluation, easily available associations like "Dollar Store" may not have created a good fit between the word *flashlight* and the student's evaluation of the line as negative (Kintsch, 1998). Affective evaluation may have helped this novice reader move to a level of abstraction to articulate the connection between the concrete image and the reader's ascription of valence.

It is also possible that the process of affective evaluation supported the development of literary schemata or control structures (Kintsch, 1980, 1998) for the intervention students. As one student noted in her prestudy think-aloud, "I never know what I'm supposed to be paying attention to." A flexible heuristic such as affective evaluation may have acted as a guide or constraint for intervention students' readings, which in turn led them to respond to textual propositions more frequently and with interpretive purpose.

## Limitations and Future Work

The primary limitation of this study is the messiness of its real-world context. Differences in participants and texts may have acted as confounds in this study. For example, differences in students' background knowledge or orientations toward literature undoubtedly affected their responses. One comparison student revealed in casual conversation that he read Shakespeare's sonnets "for fun" and that he took pleasure in looking for "deeper meanings" in those poems. It seems likely that this student brought a set of interpretive skills to the table that other students did not. In fact, of all the intervention and comparison students, this student made the highest percentage of interpretive responses at pretest. This example points to the socially situated nature of students' experiences and school performance, both generally speaking and in this study. At the same time, it bears repeating that everyday experience with interpretation did not lead to high interpretive scores for either group before the study began, suggesting that even if students had strong interpretive skills, they did not recruit them for classroom reading. This result can be seen as another argument for the importance of making everyday interpretive practices visible to students.

In a field in which personal interactions and local environment influence student motivation and performance (Klem & Connell, 2004; Patrick, Hisley, & Kempler, 2000; Skinner & Belmont, 1993), inevitable differences in the teachers' relationships with the students may also have influenced the results of this study. It is also worth noting that a meta-analysis of quasi-experimental interventions like this one shows that when intervention and comparison curricula are taught by different teachers, the results for the intervention group are higher on average than if both curricula are taught by the same teacher (Swanson & Hoskyn, 1998). However, the teachers were similar in many important ways, including level of education and experience and epistemological commitments. In addition, it remains common practice to design studies with separate intervention and comparison teachers and classrooms (Bortolussi & Dixon, 1996; Morrow, 1992; Peskin & Wells-Jopling, 2012). This practice suggests that this aspect of the design did not undermine the results.

Finally, although the major texts used in both classes were widely anthologized and similar in reading levels, suggesting similarities in accessibility, richness of literary language, and general capacity for interpretation, the texts themselves were different. They explored different themes and employed different styles. Those differences may have influenced the activation of students' interpretive resources or students' levels of interest.

Next steps in this work should address some of these limitations, particularly the potential confounds of different teachers and texts. The next iteration of this study

could recruit a teacher who teaches more than one section of the same course (e.g., four sections of senior English) and uses the same texts for each section. The teacher could then implement the intervention in half of the sections (see Smith, 1989). Although it may be difficult or inconvenient for a teacher to compartmentalize his or her teaching in this fashion, this is probably the cleanest design for this type of study. An alternative iteration involves asking intervention and comparison teachers to use all or many of the same texts (see Janssen, Braaksma, & Rijlaarsdam, 2005) in their instructional units. This design would require additional planning and coordination between individual teachers, which might make for an additional set of interesting comparisons.

Another step in this work is the further exploration of the process by which students moved from identifying valence-laden details, to connecting salient details, to constructing full-fledged thematic statements. Such a study must begin to explore the constellation of classroom practices that support students in interpreting texts, as well as an exploration of the many ways teachers make such practices visible to students. Hopefully, such research will help teachers and students understand and access the richness of literary reading and interpretation.

## NOTES

[1] For the purposes of this study, literary texts are defined as those that are created intentionally by an author and that offer some potential for duplicity of code or displacement of meaning (Scholes, 1977). Literary texts here include songs, poems, short stories, novels, speeches, and advertisements.

[2] The standardized reading test was part of the ACT, a nationally accepted college entrance exam.

[3] In order to use the Flesch–Kincaid readability test accurately, the poems were restructured as prose, so important qualities and complexities of the poems were lost. To compare the short stories read in the intervention class and the novella read in the comparison class, representative excerpts were submitted to the Flesch–Kincaid test.

[4] The same statistical tests were performed on the original, untransformed data, and the results were unchanged.

## REFERENCES

Agee, J. (2000). What is effective literature instruction? A study of experienced high school English teachers in differing grade- and ability-level classes. *Journal of Literacy Research*, *32*(3), 303–348. doi:10.1080/10862960009548084

Andringa, E. (1990). Verbal data on literary understanding: A proposal for protocol analysis on two levels. *Poetics*, *19*(3), 231–257. doi:10.1016/0304-422X(90)90022-W

Babad, E., & Taylor, P.J. (1992). Transparency of teacher expectancies across language, cultural boundaries. *The Journal of Educational Research*, *86*(2), 120–125. doi:10.1080/00220671.1992.9941148

Bloom, H. (2009). *Hispanic-American writers*. New York, NY: Bloom's Literary Criticism.

Bortolussi, M., & Dixon, P. (1996). The effects of formal training on literary reception. *Poetics*, *23*(6), 471–487. doi:10.1016/0304-422X(95)00014-B

Bowdle, B.F., & Gentner, D. (2005). The career of metaphor. *Psychological Review*, *112*(1), 193–216. doi:10.1037/0033-295X.112.1.193

Bower, G. (1992). How might emotions affect learning? In S.A. Christianson (Ed.), *The handbook of emotion and memory: Research and theory* (pp. 3–31). Hillsdale, NJ: Erlbaum.

Brendl, C., & Higgins, E. (1996). Principles of judging valence: What makes events positive or negative? *Advances in Experimental Social Psychology*, *28*, 95–160. doi:10.1016/S0065-2601(08)60237-3

Brinker, M. (1995). Theme and interpretation. In C. Bremond, J. Landy, & T.G. Pavel (Eds.), *Thematics: New approaches* (pp. 33–44). Albany, NY: SUNY Press.

Cacciari, C., & Glucksberg, S. (1994). Understanding figurative language. In M.A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 447–477). San Diego, CA: Academic.

Clore, G.L., & Ortony, A. (2000). Cognition in emotion: Always, sometimes, or never? In R.D. Lane, & L. Nadel (Eds.), *Cognitive neuroscience of emotion* (pp. 24–61). New York, NY: Oxford University Press.

Coady, R.J., & Johannessen, J. (Eds.). (2006). *The book that changed my life: 71 remarkable writers celebrate the books that matter most to them*. New York, NY: Gotham.

Collins, A., Brown, J.S., & Holum, A. (1991). Cognitive apprenticeship: Making thinking visible. *American Educator*, *15*(3), 6–11, 38–46.

Collins, A., & Loftus, E.F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*(6), 407–428. doi:10.1037/0033-295X.82.6.407

Collins, A., & Quillian, M.R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, *8*(2), 240–247. doi:10.1016/S0022-5371(69)80069-1

Culler, J.D. (2002). *Structuralist poetics: Structuralism, linguistics and the study of literature* (2nd ed.). London, UK: Routledge.

Davis, A.P. (1952). The Harlem of Langston Hughes' poetry. *Phylon*, *13*(4), 276–283.

de Valdés, M.E. (1995). Verbal and visual representation of women: *Como agua para chocolate/Like Water for Chocolate*. *World Literature Today*, *69*(1), 78–82. doi:10.2307/40150861

Dobrian, S.L. (1996). Romancing the cook: Parodic consumption of popular romance myths in *Como agua para chocolate*. *Latin American Literary Review*, *24*(48), 56–66. doi:10.2307/20119730

Dorfman, M.H. (1996). Evaluating the interpretive community: Evidence from expert and novice readers. *Poetics*, *23*(6), 453–470. doi:10.1016/0304-422X(96)00004-6

Dorfman, M.H., & Brewer, W.F. (1994). Understanding the points of fables. *Discourse Processes*, *17*(1), 105–129. doi:10.1080/01638539409544861

Dyer, M.G. (1983). The point of thematic abstraction units. *Behavioral and Brain Sciences*, *6*(4), 599–600. doi:10.1017/S0140525X00017611

Earthman, E. (1992). Creating the virtual work: Readers' processes in understanding literary texts. *Research in the Teaching of English*, *26*(4), 351–384.

Eeds, M., & Wells, D. (1989). Grand conversations: An exploration of meaning construction in literature study groups. *Research in the Teaching of English*, *23*(1), 4–29.

Ericsson, K.A., & Simon, H.A. (1984). *Protocol analysis: Verbal reports as data*. Cambridge, MA: MIT Press.

Fairclough, N. (2001). *Language and power* (2nd ed.). Harlow, England: Longman.

Franke, M.L., Carpenter, T.P., Levi, L., & Fennema, E. (2001). Capturing teachers' generative change: A follow-up study of professional development in mathematics. *American Educational Research Journal*, *38*(3), 653–689. doi:10.3102/00028312038003653

Friese, E.E.G., Alvermann, D.E., Parkes, A., & Rezak, A.T. (2008). Selecting texts for English language arts classrooms: When assessment is not enough. *English Teaching: Practice and Critique*, 7(3), 74–99.

Gaskill, G. (1973). The effect of black/white imagery in Richard Wright's *Black Boy*. *Negro American Literature Forum*, 7(2), 46–48. doi:10.2307/3041268

Gee, J.P. (2008). *Social linguistics and literacies: Ideology in discourses* ( 3rd ed.). London, UK: Routledge.

Gerrig, R., & Rapp, D.N. (2004). Psychological processes underlying literary impact. *Poetics Today*, 25(2), 265–281. doi:10.1215/03335372-25-2-265

Gibbs, R.W., Jr. (Ed.). (2008). *The Cambridge handbook of metaphor and thought*. New York, NY: Cambridge University Press. doi:10.1017/CBO9780511816802

Gillan, M.M., & Gillan, J. (1994). *Unsettling America: An anthology of contemporary multicultural poetry*. New York, NY: Penguin.

Grossman, P.L., & Shulman, L.S. (1994). Knowing, believing, and the teaching of English. In T. Shanahan (Ed.), *Teachers thinking, teachers knowing: Reflections on literacy and language education* (pp. 3–22). Urbana, IL: National Council of Teachers of English.

Gunning, T.G. (2003). The role of readability in today's classrooms. *Topics in Language Disorders*, 23(3), 175–189. doi:10.1097/00011363-200307000-00005

Hakutani, Y. (1985). Creation of the self in Richard Wright's *Black Boy*. *Black American Literature Forum*, 19(2), 70–75. doi:10.2307/2904009

Harker, W.J. (1994). "Plain sense" and "poetic significance": Tenth-grade readers reading two poems. *Poetics*, 22(3), 199–218. doi:10.1016/0304-422X(94)90005-1

Hull, G., & Rose, M. (1990). "This Wooden Shack Place": The logic of an unconventional reading. *College Composition and Communication*, 41(3), 287–298. doi:10.2307/357656

Ingram, R. (1998). Robert Herrick and the makings of *Hesperides*. *Studies in English Literature, 1500–1900*, 38(1), 127–147. doi:10.2307/451084

Ivey, G., & Broaddus, K. (2001). "Just plain reading": A survey of what makes students want to read in middle school classrooms. *Reading Research Quarterly*, 36(4), 350–377. doi:10.1598/RRQ.36.4.2

Janssen, T., Braaksma, M., & Rijlaarsdam, G. (2005, August). *Flexibility in reading literary texts: Differences between weak and strong adolescent readers*. Paper presented at the 11th European Conference for Learning and Instruction, Nicosia, Cyprus.

Janssen, T., Braaksma, M., & Rijlaarsdam, G. (2006). Literary reading activities of good and weak students: A think aloud study. *European Journal of Psychology of Education*, 21(1), 35–52. doi:10.1007/BF03173568

Judd, C.M., McClelland, G.H., & Ryan, C.S. (2009). *Data analysis: A model comparison approach* (2nd ed.). New York, NY: Routledge.

Kintsch, W. (1980). Learning from text, levels of comprehension, or why anyone would read a story anyway. *Poetics*, 9(1–3), 87–98. doi:10.1016/0304-422X(80)90013-3

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge, UK: Cambridge University Press.

Klem, A.M., & Connell, J.P. (2004). Relationships matter: Linking teacher support to student engagement and achievement. *Journal of School Health*, 74(7), 262–273. doi:10.1111/j.1746-1561.2004.tb08283.x

Kurtz, V., & Schober, M.F. (2001). Readers' varying interpretations of theme in short fiction. *Poetics*, 29(3), 139–166. doi:10.1016/S0304-422X(01)00040-7

Laplante, D., & Ambady, N. (2003). On how things are said: Voice, tone, voice intensity, verbal content, and perceptions of politeness. *Journal of Language and Social Psychology*, 22(4), 434–441. doi:10.1177/0261927X03258084

Lee, C.D. (1995). A culturally based cognitive apprenticeship: Teaching African American high school students skills in literary interpretation. *Reading Research Quarterly*, 30(4), 608–630. doi:10.2307/748192

Lee, C.D. (2001). Is October Brown Chinese? A cultural modeling activity system for underachieving students. *American Educational Research Journal*, 38(1), 97–141. doi:10.3102/00028312038001097

Lee, C.D. (2006). 'Every good-bye ain't gone': Analyzing the cultural underpinnings of classroom talk. *International Journal of Qualitative Studies in Education*, 19(3), 305–327. doi:10.1080/09518390600696729

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(1), 49–58. doi:10.1075/ssol.1.1.05lee

Lehr, S. (1988). The child's developing sense of theme as a response to literature. *Reading Research Quarterly*, 23(3), 337–357. doi:10.2307/748046

Lerner, M.J., Miller, D.T., & Holmes, J.G. (1976). Deserving and the emergence of forms of justice. *Advances in Experimental Social Psychology*, 9, 133–162. doi:10.1016/S0065-2601(08)60060-X

Levine, S., & Horton, W.S. (2013). Using affective appraisal to help readers construct literary interpretations. *Scientific Study of Literature*, 3(1), 105–136. doi:10.1075/ssol.3.1.10lev

Mar, R.A., & Oatley, K. (2008). The function of fiction is the abstraction and simulation of social experience. *Perspectives on Psychological Science*, 3(3), 173–192. doi:10.1111/j.1745-6924.2008.00073.x

McKoon, G., & Ratcliff, R. (1992). Inference during reading. *Psychological Review*, 99(3), 440–466. doi:10.1037/0033-295X.99.3.440

Morrow, L.M. (1992). The impact of a literature-based program on literacy achievement, use of literature, and attitudes of children from minority backgrounds. *Reading Research Quarterly*, 27(3), 251–275. doi:10.2307/747794

Nasir, N.S., Rosebery, A.S., Warren, B., & Lee, C.D. (2006). Learning as a cultural process: Achieving equity through diversity. In R. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 489–504). New York, NY: Cambridge University Press.

Nelson, C.A. (1987). The recognition of facial expressions in the first two years of life: Mechanisms of development. *Child Development*, 58(4), 889–909. doi:10.2307/1130530

Olson, C.L. (1976). On choosing a test statistic in multivariate analysis of variance. *Psychological Bulletin*, 83(4), 579–586. doi:10.1037/0033-2909.83.4.579

Ortony, A., Revelle, W., & Zinbarg, R. (2006). Why emotional intelligence needs a fluid component. In G. Matthews, M. Zeidner, & R.D. Roberts (Eds.), *Science of emotional intelligence: Knowns and unknowns* (pp. 288–304). Oxford, UK: Oxford University Press.

Osgood, C. (1964). Semantic differential technique in the comparative study of cultures. *American Anthropologist*, 66(3), 171–200. doi:10.1525/aa.1964.66.3.02a00880

Paivio, A. (1990). *Mental representations: A dual coding approach*. New York, NY: Oxford University Press. doi:10.1093/acprof:oso/9780195066661.001.0001

Patrick, B.C., Hisley, J., & Kempler, T. (2000). "What's everybody so excited about?": The effects of teacher enthusiasm on student intrinsic motivation and vitality. *Journal of Experimental Education*, 68(3), 217–236. doi:10.1080/00220970009600093

Peskin, J. (1998). Constructing meaning when reading poetry: An expert-novice study. *Cognition and Instruction*, 16(3), 235–263. doi:10.1207/s1532690xci1603_1

Peskin, J., Allen, G., & Wells-Jopling, R. (2010). "The educated imagination": Applying instructional research to the teaching of symbolic interpretation of poetry. *Journal of Adolescent & Adult Literacy*, 53(6), 498–507. doi:10.1598/JAAL.53.6.6

Peskin, J., & Wells-Jopling, R. (2012). Fostering symbolic interpretation during adolescence. *Journal of Applied Developmental Psychology*, *33*(1), 13–23. doi:10.1016/j.appdev.2011.08.002

Rabinowitz, P. (1998). *Before reading: Narrative conventions and the politics of interpretation*. Columbus: Ohio State University Press.

Rapp, D.N., van den Broek, P., McMaster, K.L., Kendeou, P., & Espin, C.A. (2007). Higher-order comprehension processes in struggling readers: A perspective for research and intervention. *Scientific Studies of Reading*, *11*(4), 289–312. doi:10.1080/10888430701530417

Rosenblatt, L. (1982). The literary transaction: Evocation and response. *Theory Into Practice*, *21*(4), 268–277. doi:10.1080/00405848209543018

Schoenfeld, A.H. (2011). *How we think: A theory of goal-oriented decision making and its educational applications*. New York, NY: Routledge.

Scholes, R. (1977). Toward a semiotics of literature. *Critical Inquiry*, *4*(1), 105–120. doi:10.1086/447926

Schultz, E. (2002). Natural and unnatural circumstances in Langston Hughes' *Not Without Laughter*. *Callaloo*, *25*(4), 1177–1187. doi:10.1353/cal.2002.0171

Schwanenflugel, P.J., Akin, C., & Luh, W.-M. (1992). Context availability and the recall of abstract and concrete words. *Memory & Cognition*, *20*(1), 96–104. doi:10.3758/BF03208259

Scott, V.M., & Huntington, J.A. (2007). Literature, the interpretive mode, and novice learners. *Modern Language Journal*, *91*(1), 3–14. doi:10.1111/j.1540-4781.2007.00506.x

Seifert, C.M., Robertson, S.P., & Black, J.B. (1985). Types of inferences generated during reading. *Journal of Memory and Language*, *24*(4), 405–422. doi:10.1016/0749-596X(85)90037-3

Skinner, E.A., & Belmont, M.J. (1993). Motivation in the classroom: Reciprocal effects of teacher behavior and student engagement across the school year. *Journal of Educational Psychology*, *85*(4), 571–581. doi:10.1037/0022-0663.85.4.571

Smagorinsky, P. (2001). If meaning is constructed, what is it made from? Toward a cultural theory of reading. *Review of Educational Research*, *71*(1), 133–169. doi:10.3102/00346543071001133

Smagorinsky, P. (2011). Vygotsky's stage theory: The psychology of art and the actor under the direction of *perezhivanie*. *Mind, Culture, and Activity*, *18*(4), 319–341. doi:10.1080/10749039.2010.518300

Smith, M.W. (1989). Teaching the interpretation of irony in poetry. *Research in the Teaching of English*, *23*(3), 254–272.

Svensson, C. (1987). The construction of poetic meaning: A developmental study of symbolic and non-symbolic strategies in the interpretation of contemporary poetry. *Poetics*, *16*(6), 471–503. doi:10.1016/0304-422X(87)90014-3

Swanson, H.L., & Hoskyn, M. (1998). Experimental intervention research on students with learning disabilities: A meta-analysis of treatment outcomes. *Review of Educational Research*, *68*(3), 277–321. doi:10.3102/00346543068003277

vanSomeren, M.W., Barnard, Y.F., & Sandberg, J. (1994). *The think aloud method: A practical guide to modelling cognitive processes*. London, UK: Academic.

Vipond, D., & Hunt, R.A. (1984). Point-driven understanding: Pragmatic and cognitive dimensions of literary reading. *Poetics*, *13*(3), 261–277. doi:10.1016/0304-422X(84)90005-6

Walker-Andrews, A.S. (1997). Infants' perception of expressive behaviors: Differentiation of multimodal information. *Psychological Bulletin*, *121*(3), 437–456. doi:10.1037/0033-2909.121.3.437

Williams, J.P., Lauer, K.D., Hall, K.M., Lord, K.M., Gugga, S., Bak, S.-J., …deCani, J.S. (2002). Teaching elementary school students to identify story themes. *Journal of Educational Psychology*, *94*(2), 235–248. doi:10.1037/0022-0663.94.2.235

Wissman, K. (2007). "Writing will keep you free": Allusions to and recreations of the fairy tale heroine in *The House on Mango Street*. *Children's Literature in Education*, *38*(1), 17–34. doi:10.1007/s10583-006-9018-0

Xiaojing, Z. (1996). Inheritance and invention in Li-Young Lee's poetry. *Melus*, *21*(1), 113–132. doi:10.2307/467810

Yagelski, R., & Leonard, S.A. (2002). *The relevance of English: Teaching that matters in students' lives*. Urbana, IL: National Council of Teachers of English.

Zajonc, R.B. (1980). Feeling and thinking: Preferences need no inferences. *The American Psychologist*, *35*(2), 151–175. doi:10.1037/0003-066X.35.2.151

Zajonc, R.B. (1984). On the primacy of affect. *American Psychologist*, *39*(2), 117–123. doi:10.1037/0003-066X.39.2.117

Zhang, H., & Hoosain, R. (2001). The influence of narrative text characteristics on thematic inference during reading. *Journal of Research in Reading*, *24*(2), 173–186. doi:10.1111/1467-9817.00140

Zillmann, D. (1995). Mechanisms of emotional involvement with drama. *Poetics*, *23*(1/2), 33–51. doi:10.1016/0304-422X(94)00020-7

Zyngier, S., Fialho, O., & do Prado Rios, P.A. (2007). Revisiting literary awareness. In G. Watson & S. Zyngier (Eds.), *Literature and stylistics for language learners: Theory and practice* (pp. 194–209). New York, NY: Palgrave Macmillan.

## LITERATURE CITED

Cisneros, S. (1991). Linoleum roses. In *The house on Mango Street* (2nd ed., p. 101). New York, NY: Vintage Contemporaries.

Díaz, J. (1997). Ysrael. In *Drown* (pp. 1–20). New York, NY: Riverhead.

Esquivel, L. (1995). *Like water for chocolate: A novel in monthly installments with recipes, romances, and home remedies* (C. Christensen & T. Christensen, Trans.). New York, NY: Anchor.

Ferguson, M.W., Salter, M.J., & Stallworthy, J. (Eds.). (2004). *The Norton anthology of poetry* (5th ed.). New York, NY: W.W. Norton.

Gibbard, B. (2006). I will follow you into the dark [Recorded by Death Cab for Cutie]. On *Plans* [CD]. New York, NY: Atlantic Records.

Herrick, R. (2004). Delight in disorder. In M.W. Ferguson, M.J. Salter, & J. Stallworthy (Eds.), *The Norton anthology of poetry* (5th ed., p. 355). New York, NY: W.W. Norton.

Hughes, L. (1995). Harlem sweeties. In *The collected poems of Langston Hughes* (pp. 245–246; A. Rampersad & D. Roessel, Eds.). New York, NY: Knopf.

Kenyon, J. (2005). The shirt. In *Collected poems* (p. 34). Saint Paul, MN: Graywolf.

Lee, L.-Y. (1986). I ask my mother to sing. In *Rose* (p. 50). Rochester, NY: BOA Editions.

Lehman, D. (2006). *The Oxford book of American poetry*. Oxford, UK: Oxford University Press.

Morrison, T. (2007). *The bluest eye*. New York, NY: Vintage.

Nash, T., Stewart, C., Bieber, J., Milian, C., & Bridges, C. (2010). Baby [Recorded by J. Bieber]. On *My world 2.0* [CD]. Atlanta, GA: Island RBMG.

Plante, J. (1963). I will follow him [Recorded by Little Peggy March; N. Gimbel, Trans.]. On *Wind up doll*. Nashville, TN: RCA Victor.

Powers, R. (1996). *Prisoner's dilemma*. New York, NY: HarperPerennial.

Wright, R. (2005). *Black boy*. New York, NY: HarperCollins.

**SARAH LEVINE** is a postdoctoral fellow in the School of Education and Social Policy at Northwestern University in Evanston, Illinois, USA; e-mail sarahlevine2013@u.northwestern.edu.

# Modified Novel Excerpt Used in the Think-Aloud Interpretive Task

The source for the excerpt used in this study is an International Baccalaureate exam, which in turn was excerpted from the novel *Prisoner's Dilemma* (Powers, 1996). The text shown here represents a shortened version of the excerpt, with a few difficult words replaced for ease of reading. Modifications are in brackets. The version presented to students did not include the brackets, information about the length of the original text, or the credit line.

Somewhere, my father is teaching us the names of the constellations. We lie in the cold, out in the dark backyard, on our backs against the hard November ground. We children distribute ourselves over his enormous body like so many spare handkerchiefs. He does not feel our weight. My father points a [cheap flashlight] at the holes in the enclosing black shell. We lie on the frozen earth while all in front of us spreads the illustrated textbook of winter sky. The six-volt beam creates the one weak warm spot in the entire world.

My father is doing what he does best, doing the only thing he knew how to do in this life. He is quizzing us, plaguing his kids with questions. Where is the belt of Orion [in the constellations]? What is the English for Ursa Major? […] How big is a magnitude? He talks to us only in riddles. […]

He points his way with the flashlight, although the beam travels only a few feet before it is swallowed up in the general black. Still, my father waves the pointer around the sky map as if the light goes all the way out to the stars themselves. "There," he says to us, to himself, to the empty night. "Up *there*." We have to follow him, find the picture by telepathy. We are all already expert at second-guessing. […] We lie all together for once, learning to see [the constellations] Taurus and Leo as if our survival depends on it. […]

"Here; [that] dim line [up there]. Imagine a serpent, a dragon: can you all see [it]?" My older sister says she can, but the rest of us suspect she is lying. I can see the Dipper, the big one, the obvious one.

And I think I can make out the Milky Way. The rest is a blur, a rich, confusing picture book of too many possibilities. But even if we can't see the [clusters of stars], all of us, even my little brother, can hear in my father's quizzes the main reason for his taking us out under the winter lights: "If [there is] one thing the universe excels at, it's empty space." We are out here alone, on a sliver of rock under the black vacuum, with nothing but his riddles for our thin atmosphere. He seems to tell us that the more we know, the less we can be hurt. But he leaves the all-important corollary, the how-to-get-there, up to us, the students, as an exercise. […]

We have a few questions of our own to ask him in return [before he flicks off the beam]. What are we running from? How do we get back? Why are you leaving us? What happens to students who fail? […] But I have already learned, by example, to keep the real questions for later. I hold [them] until [it's] too late.

I feel cold, colder than the night's temperature, a cold that carries easily across the following years. Only the sight of my mother in the close glow of kitchen window, the imagined smell of cocoa, blankets, and hot lemon dish soap, keeps me from going stiff and giving in. I pull closer to my father, but something is wrong. He has thought himself into another place. He has already left us. He is no longer warm.

(Adapted from *Prisoner's Dilemma* (pp. 13–14), by R. Powers, 1996, New York, NY: HarperCollins. Copyright 1988 by Richard Powers. Adapted with permission.)

# Description of the Intervention and Comparison Instructional Units

## *Intervention Group*

### *Days 1–3*

#### Making Affective Evaluation Visible

At the beginning of the unit, the intervention teacher introduced the affective evaluation heuristic with the concept of connotations. She presented students with a list of approximate synonyms for the word *house*, such as *domicile*, *shack*, *crib*, and *cottage*, and asked students where they would prefer to live. Students identified and explained their preferred terms by articulating associations with those terms. Students had negative reactions to *shack* and *domicile* but positive reactions to *crib* and *cottage*. For example, one student preferred the word *cottage* because it sounded more "pure and innocent." The teacher then defined what the students had done as "constructing connotations" of language. She pointed out to them that even though all the words they looked at were synonyms for *house*,

students knew that each word had different connotations and could have different effects on an audience. Next, the teacher pointed out that part of what students were doing when they created connotations was making an evaluation of the valence, or positive/negative effects, of the words. She explained that individuals often make a very quick evaluation of the positive or negative impact of a thing first and then figure out why they made such an evaluation. She explained that such a process was a great way to examine the effects of language in all sorts of texts.

This exercise was repeated in small groups with synonyms for *woman*, such as *lady* and *chick*, along with other terms that the students called out. Students first evaluated the positive or negative valence of each word and then explained their evaluations, generally using abstract terms. Comparisons of synonyms for *woman* led to a discussion about the way concrete words embodied abstract concepts and, therefore, had powerful effects. The teacher then reintroduced the idea of connotation and wrote down the three steps of the affective evaluation on the blackboard:

1. Identify language that you feel has strong valence.
2. Evaluate valence (positive, negative, or both).
3. Explain or justify your evaluations.

## Practicing Affective Evaluation on Short Texts

Then, the students engaged in exercises in which they practiced affective evaluation of other popular texts, such as the comparative effects of the stage name Marilyn Monroe and Monroe's given name, Norma Jeane Mortenson. They examined the tones created by different word choice in newspaper headlines (e.g., "Police Officer Controls Riot" vs. "Cop Suppresses Gathering") and articulated the connotations that they felt were created by specific words and the headline as a whole. Next, they did similar work with advertisements, making text-based evaluations of mood and tone and explaining those evaluations through abstractions. The teacher led several class discussions in which she reminded students that the process they were using was one they used in their everyday lives; now they were carrying that practice into school, taking it apart, and trying it on other kinds of texts.

## *Days 4 and 5: Exploring Gender and Coming-of-Age*

As discussed earlier, the intervention teacher also designed the unit to support students' exploration of the topic of gender and coming-of-age. She did this by introducing and asking students to define coming-of-age and compare the experience for young women and young men.

Students then practiced the affective evaluation heuristic with a series of literary texts that could be characterized as exploring both gender and coming-of-age. The teacher began with simpler, more familiar texts, such as a version of Cinderella. The teacher asked students to practice the first step of the heuristic by identifying details in the text that seemed particularly affect-laden. For example, after reading the line "It was afternoon. Cinderella was left at home, weeping by the ashes," students identified text that they felt was especially valence-laden. In this case, students felt that the second sentence ("Cinderella was left at home…") was more affect-laden than the first, and that it was negative in tone. Within that second sentence, students pointed to the words "left," "weeping," and "ashes" as being especially salient to their negative evaluations. Then, the students justified their evaluations, agreeing that in context, those words created a sense of abandonment and hopelessness. In this way, students moved beyond tracking the action of the story to identifying details that were salient to interpretation and constructing abstract connotations. Students repeated this process with every phrase or detail that they felt was valence-laden.

After finishing the story, the students practiced affective evaluation on the cumulative effects of this version of Cinderella. Most judged the story to have positive effects, explaining their evaluations with simple thematic inferences, such as "Hope will win out in the end," and "Love conquers all." Notably, a few students judged the overall effect to be negative, justifying their evaluation by saying that in Cinderella, "the only way a woman succeeds is through a man."

As the students formulated interpretations, the teacher provided them with sentence stems designed to support the use of affective evaluation. For instance, one stem read, "This text *condemns* a world in which…," and another read, "This text *celebrates* a world in which…." Students then wrote a one-paragraph argument about the effects of Cinderella.

## *Day 6: Continued Practice With the Heuristic*

The students continued their practice of affective evaluation with an excerpt from Richard Wright's novel *Black Boy*. This text also addressed issues of coming-of-age. First, students made arguments about whether the narrator had in fact come of age, and then returned to the text to use affective evaluation to construct connotation and thematic inferences of details and overall effects.

## Days 7–11: Comparing Texts and Writing Arguments

Students compared two popular love songs, "I Will Follow Him" and "I'll Follow You Into the Dark" that addressed the same theme, devoted love, but used different styles to do so. Students practiced affective evaluation with both texts. Students then used the school's common argument template to make claims about which kind of love was more mature. They incorporated some of their interpretations of the effects of language into those arguments.

## Days 12–18: Continued Practice With the Heuristic

Students read more contemporary short stories from Sandra Cisneros's *The House on Mango Street* and Junot Díaz's *Drown*. In a teacher-led discussion, students practiced the affective evaluation on a vignette from Cisneros's novel. They then worked as a whole class to identify and construct connections between connotations that they felt were created by the text. Students used the sentence stems to develop thematic statements based on connections between abstractions. Students then read a story from Díaz's collection; about 25% was read in class, and the rest was read for homework. To account for students who had not read for homework, the teacher asked class members to review important sections. Next, students worked in small groups to use affective evaluation with one motif from the text. Finally, students used the school argument template to make an argument about the effect of that motif on the themes of the story as a whole.

The intervention teacher had also planned to work with several poems but ran out of time.

## Comparison Group

As described earlier, the comparison teacher also combined teaching of literary interpretation with exploration of a topic (in her case, love and gender). Her students also used the schoolwide argument template to support claims about their literary interpretations.

## Days 1–3: Debating the Power of Literature and Language

The comparison group teacher began her unit on literary interpretation with a class discussion on the power of language. To do this, she designed an exercise in which the class debated the role of literature in everyday life. Specifically, the class debated the extent to which literary texts could change the way people felt or behaved and whether certain texts, such as some contemporary pop songs, were too provocative for young readers. Students then read the lyrics to "Baby," a Justin Bieber pop song, focusing on the effects of particular language choices in

the song. Students were asked to point to particular details in the song that they felt were especially provocative or innocent and explain why.

## Days 4 and 5: Argument About the Effects of Language

Students used the school argument template to make claims about how the song by Justin Bieber or a song of students' own choosing used language to provoke or in other ways affect their audience.

## Days 6–10: Exploring Love and Gender

Like the intervention teacher, the comparison teacher introduced and asked students to define *romantic love* and then asked them to debate whether the definition was the same for women and men. Then, as was done in the intervention group, the students read two poems that addressed the same themes, love and lust, but with very different styles: "The Shirt" by Jane Kenyon and Robert Herrick's "Delight in Disorder." The teacher asked the students to compare how each poem presented ideas about love and lust and determine which was a "better" representation. The teacher additionally asked students to identify as many literary devices as they could in each text. Students worked in small groups to identify literary devices such as metaphors, symbols, allusions, and alliteration.

Students then performed readings of the poems. During these exercises, the teacher defined, pointed out, or asked students to point out literary devices, particularly imagery and symbols. She led several class discussions about the meaning of the symbols in relation to the themes of the poems.

Additionally, the teacher presented an explicit heuristic for literary interpretation: looking for patterns of imagery in poetry. She modeled a search for patterns with the first stanza of Langston Hughes's "Harlem Sweeties." She identified patterns of imagery of fruit and color. She then explained that to her, the imagery of fruit helped suggest lusciousness and sexuality and that the colors, mostly reds and purples, might represent passion. The teacher encouraged the students to look for other kinds of patterns in "Harlem Sweeties" and consider what they might represent.

The teacher did not connect this practice to any everyday experiences that students might have involving finding patterns in texts or real-world events. She did not model the kind of thinking that students might engage in to construct abstract, symbolic meanings from concrete imagery.

Finally, students used the school argument template to support their claims about which poem best represented love and lust.

## Days 11–18: Reading and Interpreting a Novel

After studying and writing about poems, students began their reading of *Like Water for Chocolate*, a short novel by Laura Esquivel. Students were assigned to read about 40% of the novella in class and about 60% for homework. To account for students who did not read for homework, the teacher reviewed or had other students review the action and symbols of each night's reading. During in-class reading, the teacher often stopped students to ask or write down what kinds of images stood out for them. However, she did not frame this activity as a way to read literary texts or as something the students should do on their own. The teacher also continued to point out and ask students to identify literary devices used in the text. In at least one other lesson during the unit, the teacher encouraged students to look for patterns in the text. As a group, students tracked the motif of fire and heat in *Like Water for Chocolate*. The teacher encouraged the students to make associations with the images, but the *how* of making associations was not made explicit, and the emphasis on this or other strategies was not consistent.

During discussions, students sometimes voiced their feelings about characters and events in the fiction they read, and discussed and wrote personal responses to the texts. Along with other questions and prompts, the teacher asked students to attend to the mood and tone of the text. However, the affective evaluation heuristic was not used during instruction, nor were affect-based strategies for interpretation made explicit.