

JOLLE@UGA[®]

JOURNAL OF LANGUAGE & LITERACY EDUCATION

Playful Images and Truthful Words: Eighth-Graders Respond to Shaun Tan's *Stick Figures*

Stephanie F. Reid

Abstract: This interpretive study examined how eighth-grade students responded to a multimodal short story introduced to them by their English Language Arts teacher as part of a multimodal literacies curriculum unit. By analyzing fieldnotes constructed from observations, the classroom teacher's voice-recorded reflections, and the students' work, the author investigated how students articulated and described the relationships between words and images in multimodal texts. This study showed how students were open to multiple interpretive possibilities when reading images. However, when reading a text comprised of both words and images, students looked to written language as the most significant mode of representation and communication. Furthermore, most students in the class sought word-image coherence and alignment. The discussion suggests that multimodal literacies instruction might benefit from further alignment with critical literacy pedagogy, enabling students to question and challenge texts produced through any modal combination. Literacy education researchers and classroom practitioners have the potential to ensure that students are supported in becoming critical thinkers able to explore issues of representation within the multimodal texts that saturate their social worlds.

Keywords: Critical Literacy, Multimodality, Multimodal Literacies, Visual Literacy



Stephanie F. Reid is a PhD Candidate in the Learning, Literacies, and Technologies program in the Mary Lou Fulton Teachers College at Arizona State University. Previously, Stephanie taught secondary students English Language Arts for 15 years. Stephanie's research interests include multimodal literacies pedagogy and multimodal text analysis. Her work has appeared in *The Journal of Children's Literature*, *Voices from the Middle*, *The Reading Teacher*, and *Visual Communication*. Contact the author at Stephanie.Reid@asu.edu

Introduction¹

Closure will definitely be provided by the words. I want to know which ideas we had yesterday are closest. Which one was the true story. We might have gone past the true meaning of the story.

Mako, Eighth-Grade Student

Scholarship in multimodality (Jewitt, 2017; Kress, 2010) has suggested that all modes of representation and communication have “distinct potential to contribute equally” to meaning-making activity (Kress, 2010, p. 96). Therefore, while modes such as written and spoken language are deemed significant, so, too, are other modes, such as images and design features. More than one mode of representation and communication comprise multimodal texts (Kress & van Leeuwen, 2001), and text-users should consider paying attention to how combined modes interact with each other to offer interpretive possibilities (Unsworth & Cléirigh, 2017). Each mode has the potential to modify the meanings offered by different modes (Lemke, 1998). The meaning potential of any given multimodal text is not located in one single mode. Instead, readers construct meaning across the modes available to them within a text (Serafini, 2015). Thus, when interpreting an illustrated short story, readers can construct their interpretations using both image and written language.

The purpose of this article is to share the findings of a four-day interpretive study that investigated how 22 eighth-grade students responded to Shaun Tan’s *Stick Figures* (2009), a short multimodal narrative text. Six images and five written language paragraphs occupy the six pages of Tan’s illustrated short story. Tan’s story does not follow a typical

story arc. To interpret this story, readers must draw upon the interconnected contribution of the words and images that comprise this narrative. I wished to understand how students made meaning using modes of representation and communication and how students navigated the different perspectives on the storyworld offered by the images and words. The following overarching research question guided this study: How did eighth-grade students respond to Shaun Tan’s multimodal short story? Sub-questions included: (a) How did students talk or write about the words and images during meaning-making activities? and (b) How did students describe and understand the relationship between words and images in multimodal texts?

This article documents students’ modal preference (Smith, 2017) for written language due to their equation of language with truth. My study shows how, even though Tan’s (2009) text disrupts the notion that image serves written language in visual narrative texts, students defaulted to a preferred understanding of word-image relationships as complementary (Nikolajeva & Scott, 2000). My work is relevant to researchers and educators interested in the exploration of multimodal texts and multimodal concepts in secondary English Language Arts contexts because it suggests that interpreting multimodal texts is a complex process that is further complicated when modal components do not cohere.

Research Context

To ensure attention to both modalities, the class teacher, Ms. Scarlett (all student and teacher names are pseudonyms), presented her students with only the images from Tan’s (2009) story on the first day of the study. The images show a suburban town devoid of people. No humans are visually

pronouns to refer to individuals that correspond with the pronouns that they use to refer to themselves.

¹ I acknowledge that there is a gender spectrum and that myriad pronouns exist that I can use when referring to individuals in this report. Throughout this article I use

represented, but cars, houses, and streets indicate the existence of human society. Instead, the images depict a series of stick figures, entities with tumbleweeds for heads and branch-like limbs that resemble arms and legs. The figures are shown standing on a street, waiting by a bus shelter, and lurking at the end of an alleyway or at the edge of a used car lot. In the final image, four stick figures are featured. The stick figures appear to be reaching for each other's hands, suggesting that the stick figures are sentient beings able to love and forge connections with other stick people. Students initially viewed the images in Tan's visual narrative as a site of play and possibility. They generated multiple storylines, themes, and even worlds in response to viewing them.

On the second day of the study, students read the story in its original format with the written language included. As students read the complete *Stick Figures* (Tan, 2009), they soon realized that the words did not explain what was shown in the pictures. In contrast to the images, the words describe the human inhabitants of the suburban town shown in the illustrations. The humans treat the stick people abhorrently. Small children are instructed to ignore them. House-owners install sprinkler systems to keep them off their property. Older boys beat and dismember stick figures for their amusement. The stick figures are represented in language as a group whose lives the human characters do not understand (or seek to understand).

The words disappointed many of the students. Instead of providing the "true story" for which many of the students had hoped, the words posed more questions and opened more gaps for the reader to fill (Iser, 2000). The words did not provide the "closure" that many students, like Mako, predicted.

Theoretical Framework

Multimodality and Meaning-Making

My theoretical framework is primarily informed by multimodality scholarship (Jewitt, 2017; Kress, 2010). From a social semiotics perspective on multimodality, human acts of representation and communication are understood in terms of *modes*. A mode is a "socially and culturally shaped resource for making meaning" and has social, cultural, and material aspects (Kress, 2010, p. 79). Once meaning has been made material, the created text becomes available to other people to interpret and use for meaning-making activity. Examples of modes are visual images, written language, spoken language, music, numbering systems, and gesture. In a multimodal text or *ensemble* (Kress, 2010), multiple

"Shaun Tan's *Stick Figures* (2009)... is a multimodal text comprised of images, written language, and design features."

modes of communication and representation are brought together to form one semiotic whole or *syntagm* (Barthes, 1977). Shaun Tan's *Stick Figures* (2009) is a multimodal text comprised of images, written language, and design features.

Because all modes of representation and communication have the potential to contribute to meaning-making activity, all modes are worthy of study (Cowan & Kress, 2017). However, multimodality scholarship moves beyond examining modes in isolation (Kress & van Leeuwen, 2001). To study meaning-making and understand how texts work is to examine all modes used within a text or moment of communication and analyze how they work in combination to offer meaning potential (Jewitt, Bezemer, & O'Halloran, 2016). Essentially, multimodal scholarship strives to understand the intersemiotic (Jewitt, 2017) or intermodal (Unsworth & Cléirigh, 2017) relationship between modes in multimodal texts or interactions. Multimodal semiotic wholes are more than the sum of their

parts (Bezemer & Kress, 2016; Hull & Nelson, 2005; Lemke, 1998).

In formal schooling contexts, written and spoken language continue to be privileged, their dominant position in school curricula secured by standardized assessments and teacher accountability measures that focus on logocentric texts (Davis & Willson, 2015). Given the visual and multimodal nature of the world within which students exist, the absence of multimodality and multimodality metalanguage in US national standards for English Language Arts is notable (Mills & Exley, 2014). Although literacy organizations (National Council of Teachers of English, 2005) and education researchers (Jiménez, Roberts, Brugar, Meyer, & Waito, 2017; Pantaleo, 2019; Seglem & Witte, 2009; Serafini, 2014; Siegel, 2012) have advocated for multimodal literacies and texts in English Language Arts curricula, these arguments have yet to gain traction in academic contexts (Khadka & Lee, 2019). Old literacies (O'Brien, 2012) and canonical written language texts continue to prevail in school reading and writing classrooms (Renaissance Learning, 2019; Stotsky, 2010).

Furthermore, a social semiotic approach to multimodality conflicts with prevalent close reading and text-dependent classroom reading practices that take their inspiration from a New Criticism approach to reading literature (Fisher & Frey, 2014). From the New Criticism perspective, interpretation is confined to the “four corners of the text” (Calkins, Ehrenworth, & Lehman, 2012, p. 39) and a reader’s context, life experiences, and background knowledge are deemed largely irrelevant to interpretive work. This take on reading has been advanced through the Common Core State Standards and is ensconced in standardized reading assessment (Peel, 2017). In contrast, multimodality emphasizes that meaning is always remade anew with each act of representation and communication (Kress, 2010; Bezemer & Kress, 2016). The individual

meaning-maker and their situation within a particular social context impacts all semiotic work.

Thiel and Kuby (2018) asked that we consider carefully the consequences (both positive and negative) of the literacy practices for which we advocate as researchers and educators. Three possible consequences arise from a focus on written and spoken language texts in school. First, students may associate power and authority with the modes of representation and communication to which teachers ask students to attend. Second, students may become limited in their abilities to interpret multimodal texts because they have little experience in determining how different modes work together to offer meaning potential. Third, students may not fully realize their agency in the meaning-making process and the power they hold as both interpreters and makers of texts. None of these consequences empower our students as readers or composers.

Multimodal Narratives and Word-Image Relationships

Multimodality scholars have argued that each mode of representation and communication does different semiotic work than do other modes (Kress, 2010). Thus, in Tan’s (2009) multimodal narrative text, the images and the words carry a different semiotic load. The meaning potential offered through images cannot merely mirror the linguistic information, and vice versa. Although Nikolajeva and Scott (2000) described how the *symmetrical* word-image relationship marked one end of their word-image interaction spectrum, Lewis (2001) challenged their proposition by arguing that their notion of such symmetry is “illusory” (p. 39). Connections between words and images might be made by a reader-viewer (Serafini, 2012), but images and words will always offer meaning potential not offered through the other mode. Therefore, both modes are worthy of analysis and further investigation.

Tan's (2009) word-image interaction might be defined as a *complementary relationship* (Nikolajeva & Scott, 2000) because they do not overlap in many places and serve to strengthen the narrative by offering different narrative information, each mode filling gaps left by the other mode. However, this kind of complementarity does not quite capture the dissonance the reader-viewer might experience when reading Tan's (2009) narrative. In this sense, the word-image relationship in Tan's text could also be described as *perspectival counterpoint* (Nikolajeva & Scott, 2000) because the words and images show different perspectives: The words provide insight into human actions; the images depict only the stick figures and omit the human actors. Although both perspectives are shown from a third-person narratorial standpoint, this matching viewpoint does not resolve the dissonance between the images and words, particularly as the precise persona of the shower or teller is never revealed.

Iser (2000) addressed the "gaps" in literary texts as blanks that the reader seeks to fill. For Iser, the task of interpretation involves examining absence and presence; the absent and the present continually point at each other, forcing the reader to acknowledge the incompleteness of the text. When we consider dissonance between words and images, the kind found in Tan's (2009) multimodal short story, another modal layer of absence and presence is added, further complicating the interpretive work the reader-viewer of such texts undertakes. This dissonance between words and images in Tan's text highlights how images add nuance to the notion of text complexity. Although texts with images might be viewed as less scholarly (Jiménez & Meyer, 2016), children's and young adult literature scholarship has demonstrated that multimodal narrative texts are complex because they require students to draw from

both words and images in order to make sense of the story (Callow, 2018; Serafini, 2010, 2014; Pantaleo, 2012).

Importantly, it is not enough to explore each mode of communication and representation in isolation. From within comics scholarship, McCloud's (1993) work on comics discussed how the separation of literary and artistic regimes into separate fields has led to students studying and interpreting either words or images. McCloud's scholarship illuminated an important issue: Students exist in a world where words and language are brought together within recognizable mediums such as comics and print advertisements, as well as within increasingly common formats such as multimodal novels and transmedia texts made possible through advances in technology (see Garcia, 2017; Overstreet, 2018; Slota, Young, O'Byrne, & Ballestrini, 2016). English Language Arts educators could look to incorporate strategies to help students understand the orchestration of different modes (Kress, 2010) and how different modes interanimate each other (Meek, 1992).

“Importantly, it is not enough to explore each mode of communication and representation in isolation.”

Methods

Setting

The eighth-grade students in this study attended a grade 4-12 charter school. The charter school received an A grade from the state department of education following the publication of standardized assessment results, and the school's curriculum was constructed to help their students achieve content area mastery a grade-level ahead of peers attending other schools. Across its grade 4-12 English curriculum, the school tended to promote essayist and test literacy orientations (Gee & Hayes, 2011) designed to enhance students' abilities to perform

well on the standardized tests that act as gatekeepers to higher education institutions. This study, therefore, also sought to understand how students well-versed in test-taking strategies responded to a multimodal literary text that does not pursue a traditional structure, offer a neat conclusion, or provide illustrations that merely reflect what the language says (Nikolajeva & Scott, 2000).

Participants

This article focuses on four days within a more extensive five-week study that explored how Ms. Scarlett, a White English Language Arts teacher, introduced multimodal concepts and multimodal texts to students in four eighth-grade classes. When I met Ms. Scarlett, she was working toward finishing her Master of Arts degree in English Education and had just started planning a concluding project on new literacies. Ms. Scarlett's research on new literacies had sparked an interest in the notion of multiple modes of representation and communication (New London Group, 1996). She readily agreed to participate in my study because it supported her interest in incorporating multimodal concepts and texts into classroom instruction. Before this unit, Ms. Scarlett had discussed and analyzed images with students in the context of rhetoric and persuasion but had not explored narratives comprised of words and images. To make space for the multimodal text unit, Ms. Scarlett removed a classic novel from her teaching agenda.

My role in the study was as consultant. I provided my theoretical perspective, the texts used in the study, and teaching ideas, but Ms. Scarlett designed and implemented the lesson plans towards the close of the school year during March and April.

I acquired parental permission and student assent from all 22 students in Ms. Scarlett's seventh-hour class. State records identified 47.74% of the school's

student population as Asian, 34.67% of the students as White, 8.41% of students as Hispanic, 4.92% of the students as African American, and 3.52% of the students as Multiple Races. The demographic information for students in Ms. Scarlett's class reflected the data reported for the broader school community.

Instructional Sequence

The multimodal text unit lasted five weeks. This article draws upon fieldwork conducted over the first four days of the curriculum unit. During the first lesson Ms. Scarlett showed students the six images from Tan's (2009) short story. She intended to help students focus on the narrative aspects of visual modalities. Ms. Scarlett posed one question to the class: *What's the story?* Implicit in this question was the idea that the images are connected by a storyline. Students had time to write in response to each image, talk about each image with a partner, and share their thoughts with the rest of the class. Students' ideas regarding what the narrative might be about shifted and evolved with the reveal of each picture.

Ms. Scarlett opened the second lesson by revealing that students had not seen the complete story and invited students to reflect upon how they thought words would change (or not) their reading experience. Students initially documented their thinking in writing before sharing their thoughts in the whole-class forum. Students then read *Stick Figures* (Tan, 2009) in its complete form. Talk regarding students' reactions to the complete story took up the rest of the available class time.

Although the third and fourth lesson did not specifically focus on Tan's (2009) text, the work students undertook provided them with further opportunities to consider word-image relationships in multimodal texts. Ms. Scarlett introduced students to metalanguage from Kress and van

Leeuwen's (2006) visual grammar to support students in their interpretations and discussions of image, as well as advance the notion that images offer meaning potential. Ms. Scarlett also used images from Pam Smy's *Thornhill* (2017), a multimodal novel that represents one character's experiences in written language and a second character's story in images, and she modeled how knowledge of these components of visual grammar could be used to visually analyze images. Ms. Scarlett knew that students needed support in visual analysis because written language text analysis and composition dominated their school's English Language Arts curriculum.

During the fourth lesson, students chose a multimodal text from a library of multimodal narrative texts that included graphic novels, web comics, multimodal novels, and picturebooks. Although the majority of the fourth lesson was reserved for reading, during the last twenty minutes of class, students shared significant pages with each other and were asked specifically to think about how the words and images worked together. To conclude this portion of the multimodal text curriculum, students were asked: *What do you think images should do when they are paired with words?* Ms. Scarlett then challenged students to draw and/or write similes or metaphors that could be used to describe word-image relationships in multimodal texts.

Data Collection

Interpretive research (Erickson, 1986) draws from the family of research methods that includes both ethnography and case study research. Thus, fieldnotes, participant interviews, and participant-created artifacts were central to my data collection. I turned my jottings, notes taken during lessons, into detailed written fieldnotes after each classroom observation (Emerson, Fretz, & Shaw, 2011). I conducted short, informal interviews (Olson, 2011)

with Ms. Scarlett after each lesson, and Ms. Scarlett also recorded brief audio reflections using Voxer, a mobile software application. I captured my initial analytical thinking as asides in my fieldnotes (Emerson et al., 2011) and expanded these thoughts into theoretical memos (Bazeley, 2013).

Data Analysis

In line with Saldaña's (2016) recommendations, I conducted a first-cycle analysis of my field notes and students' written responses. I employed line-by-line open coding as I read through the entire data corpus (Emerson et al., 2011) and sought, where possible, to tag segments of text with in vivo codes (Saldaña, 2016) to ensure I was paying attention to students' words and voices. I coded the data collected in chronological order, my fieldnotes helping to contextualize and situate the student work. During this stage of analysis, I generated 293 codes. Examples of in vivo codes created during this stage of analysis included, *Images Correspond and Match Words*, *Mutual Amplification*, *Words Will Shape Meaning* and *Images Should Explain More*. Coding memos helped me think through the codes and patterns I noticed.

After conducting this initial coding, a second cycle (Saldaña, 2016) system of pattern coding, together with writing integrative memos (Emerson et al., 2011), enabled me to analyze the relationships between codes and categories and explore how students were representing their thinking regarding word-image relationships across the four days. Throughout this cycle of coding, I clustered codes together and assigned interpretations to that cluster of codes. The three core pattern codes I established were: *Image Play*, *Truthful Words*, and *Defining Word-Image Relationships*. I then developed sub-codes (Miles, Huberman, & Saldaña, 2019) to identify data that would help me describe the nuance within each main code.

The assertions (Erickson, 1986; Miles et al., 2019) that comprise my findings section center on the data analysis and interpretations constructed during second cycle coding. Those assertions are: (a) images alone were viewed by students as a site of possibility and play; (b) when paired with words, students reduced images to a subordinate and supplementary role to words; and (c) students defaulted to an understanding of word-image relationships that emphasized symmetry or coherence. The sub-codes I developed for each pattern code enabled me to explain and explore each assertion.

Findings

Nodelman (1988) explored the relationship between words and pictures in children's picturebooks, and described how audiences became frustrated when presented with just the images from a narrative because they generated a myriad of different story versions (p. 193). The eighth-grade students in this study did not demonstrate frustration while viewing the images in their narrative sequence. Instead, the students in Ms. Scarlett's classroom constructed their own interpretive directions free from the potential constraints that written language might place upon their meaning-making. Students seemed to view image as a site of possibility and play. Their frustration with the short story only became evident once students accessed the text in its entirety with images and language combined, when students' need for closure and truth took precedence.

Image as a Site of Perceptual and Interpretive Possibility

When students enjoyed images as sites of possibility, they restricted their imaginings to the boundaries of the storyworld depicted in the images and kept their noticings and interpretations to the stick figures, the setting, and the atmosphere or mood depicted in the story. Through my coding, I constructed two different categories of possibilities: *Perceptual* and *Interpretive* possibilities. Guided by the perceptual component of Serafini's (2010) multimodal text analysis framework, I determined that students noted perceptual possibilities when their written responses or verbal contributions to class

discussions around the sequence of images, contained an account of what they were noticing (Serafini, 2011). These accounts tended to log the visual details they encountered in the image without any reflection regarding what these visual components might mean. I coded students' written and verbal responses as *interpretive* when students wrote or talked about what visual elements might mean or when they articulated ideas regarding possible story themes.

Perceptual Possibilities.

Serafini (2011) described naming and identifying visual elements as an initial part of the meaning-making process. What is noted by a viewer tends to be congruent with things and objects seen and experienced in the viewer's social reality. Similarly, Panofsky (cited in Serafini, 2010) identified three levels of meaning, the first of which Panofsky described as the pre-iconographic stage of meaning-making. At this stage in the meaning-making process, the reader-viewer (Serafini, 2012) recognizes people, actions, objects, and places based on their lived experience. This act of naming does not extend

“When students enjoyed images as sites of possibility, they restricted their imaginings to the boundaries of the storyworld depicted in the images and kept their noticings and interpretations to the stick figures, the setting, and the atmosphere or mood depicted in the story.”

to interpretation on the iconographical level (Panofsky, cited in Serafini, 2010), where ideas and concepts attached to the elements inventoried are also identified.

Olivia used writing to record the ideational aspects of the image: A stick figure walking toward a house, the darkness, and the brown scenery (see Figure 1). There were a few moments when Olivia's comments began to veer toward interpreting salient elements, when Olivia talked about the stick figures "avoiding reality" or "the awkward placement of [the] figure," for example. Olivia did not link the images into a coherent narrative or write about why the details noted down might matter or what they might mean. Olivia's choice of words, even though they provided a largely perceptual account, also indicated that pre-iconographical accounts or inventories of an image's components will still vary from person to person. Olivia's use of the phrase "abandoned road" and "figure" to name what she saw were not used by all students who named the visual components they saw.

Similarly, Carl modified the inventoried components noted down (see Figure 2). What Carl

meant by "deformed figure" or a "small neighborhood" requires a level of interpretation beyond merely naming what is seen. In addition, Carl added to their inventory notes about frequency. They thought it salient to note that "a wooden pole seems to appear in all of the images," and they pointed out, too, that "the three people from picture one are now at a store called More." Significantly, any act of naming things observed within a visual image is an interpretive act and a transduction of information (Kress, 2010) from one mode of representation and communication (Tan's image) to another mode (Carl's written communication). This movement of information across modes will never result in a precise replica of the original symbol. Meaning is always (re)made anew (Kress, 2010). Like Olivia, Carl did not create a narrative thread that united the image components and the characters written about with story.

Interpretive Possibilities. Other students developed their understanding of image by considering why certain details or visual components might matter. Serafini (2011) described how interpretation involves students applying their knowledge and experience to the visual components

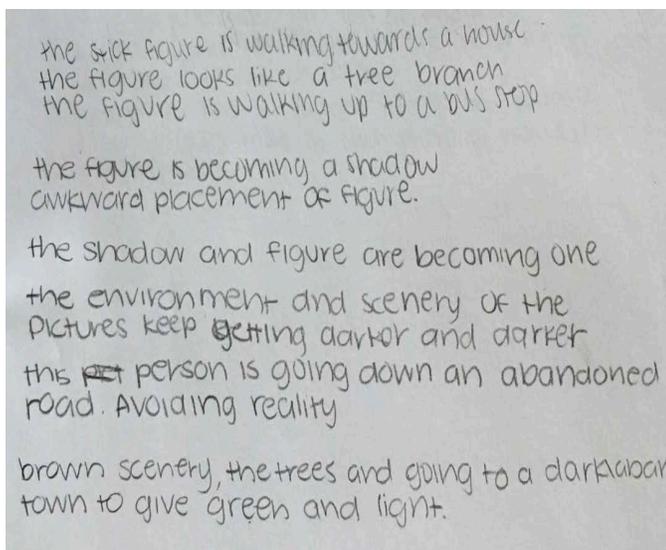


Figure 1. Olivia's written response to the images in Tan's (2009) narrative

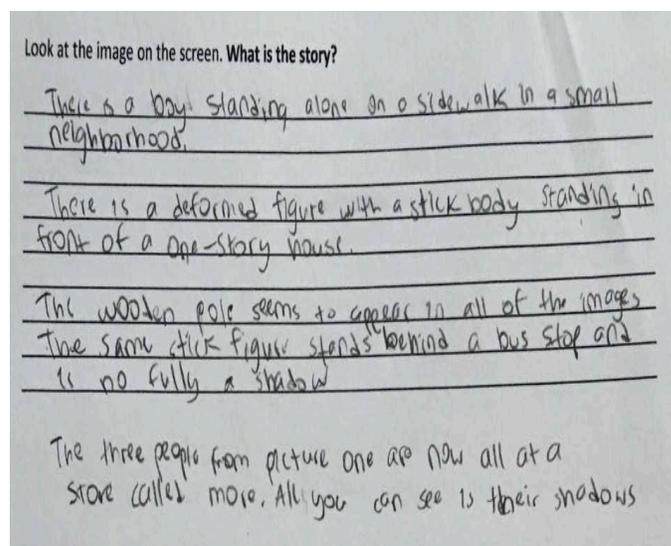


Figure 2. Carl's written response to the images in Tan's (2009) narrative

or elements that they have noticed. Students whose written and verbal texts I coded as interpretative wrote about setting, character behavior and motivation, mood, and atmosphere, and made narrative predictions that linked different images and image components together.

In their written response to the image sequence in Tan's (2009) work (see Figure 3), Brooklyn formulated several possible plot details, such as the notion that the figure at the bus stop has decided to leave home. Brooklyn also determined that the figure is homeless and writes about how the figure's "sadness is growing and spreading." In a concluding comment on the depiction of the figures in various suburban locations, they added that a rebellion is possible as the figures "spread their sadness and grievance everywhere." Brooklyn also briefly considered what the characters might be thinking and wrote that "They' are wondering what is going on and where they are." Brooklyn's use of punctuation around *they* identified the stick figures as a significant group, albeit a group for which Brooklyn had no name or familiar designation. Brooklyn recognized that the images are part of a narrative and, thus, developed noticings into comments on plot, character, and mood possibilities.

Image as a Site of Play

In their scholarship on play, Wohlwend (2017) built upon the work of the New London Group (1996) in describing how educational Discourses and formal school literacy practices "make some materials, modes, and meanings unavailable in classrooms" (Wohlwend, 2017, p. 164). Wohlwend added that "play is an unruly literacy, flexible and ambiguous, whimsical and

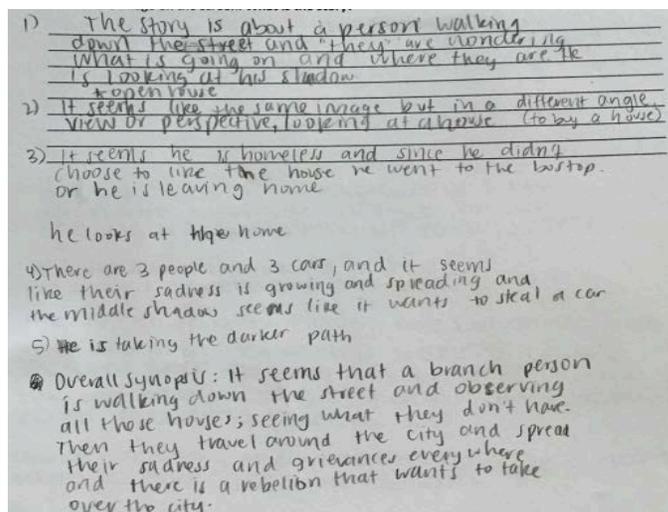


Figure 3. Brooklyn's written response to the images in Tan's (2009) narrative

capricious...Its fluidity of meaning creates a productive tension with pedagogical aims such as the need for cohesion in storytelling" (p. 166). However, although Wohlwend's scholarship focused on collaborative play involving multiple modes, a range of materials, and physical space through which students can move, I argue here that some of the students in Ms. Scarlett's class played as they engaged with the images in Tan's (2009) work. These students played with the images, dreamed up unlikely storylines, and took creative plot risks.

Mako, for example, took this classroom opportunity to dream up a dystopian world, *Utensiland*, which was inhabited by "daring, sporky souls." Mako's story depended upon readers understanding the parallel drawn between the shape of a spork and the shape of the stick figures' tumbleweed heads. When Mako shared this with the class, Mako's peers were entertained. They took pleasure in Mako's account of these figures as "mobile, sentient grass – the new dominant lifeform." Mako took pleasure in sharing their version of the story told across images.

Another student, Arvind, spun a story with flashbacks, flashforwards, and a dream sequence. David concocted a tale about a cowboy tumbleweed,

scratching across the ground of the Southwest. In Jennifer's dystopian world, the focal character was Stick Figure VIIA0012X, who spends their life waiting at the town bus stop for a bus that never arrives. Elmo experimented with intertextuality, blending the world of the stick figures with Dr. Seuss' *The Lorax*. There were clear links between the student-crafted stories and the images. Still, the students also took creative liberties, moving the plot and modified characters into fantastical worlds of their design that existed momentarily within the boundaries of this single class period.

The texts the students created did not resemble the kinds of formal writing or assessment performances that they regularly encountered in this high-stakes, test-focused school environment. Ms. Scarlett and I had placed parameters on this literacy event by choosing Tan's (2009) text, designing a response sheet, and delineating time spent constructing the response. However, despite the framing of the response activities by the teacher, the students who engaged in this kind of play worked within the planned structures of the lesson to bring to life unexpected characters and humorous stories in a manner that does not reflect the literary traditions of formal schooling. These eighth-grade students saw image as a site of play, and the classroom atmosphere during the sharing of these stories seemed exuberant and joyful. There was plenty of laughter. Their classmates readily joined in and supported this playfulness.

Words as Truth: Image as Subordinate to Words

From the moment that Ms. Scarlett notified students that *Stick Figures* (Tan, 2009) also contained words, students in the class began to assert their beliefs that words would (a) eradicate confusion and (b) provide a definitive answer regarding what the story meant. Ms. Scarlett's students connected written language with truth and

closure, reflecting their modal trust in written language to reveal the "true story" (Mako's words).

Written language as anchorage and truth. In the second lesson, when Ms. Scarlett stated that there were words to the story, there was an immediate shift in the attitude toward the text, the visual part of which they had enjoyed the lesson before. Ms. Scarlett began the lesson by asking the students to think about how words might affect their experience of reading the story and the understandings they had created during the previous lesson. Students reflected in writing first before sharing their ideas with the class.

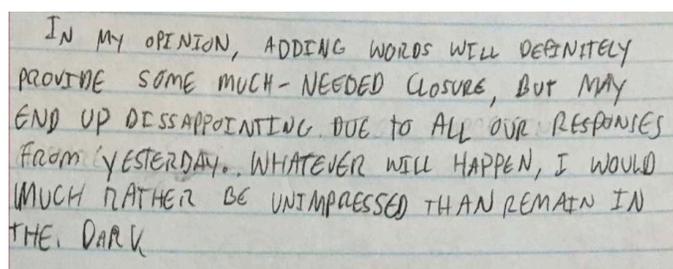
The class' responses were very much in line with Barthes's (1977) concept of *anchorage*. Barthes explained anchorage as a technique used by text-makers to direct readers to a preferred and ideologically-endowed meaning by encouraging attendance to certain aspects of a text. In our society, the linguistic message is often used to anchor pictures. Captions, for example, often anchor images published in news articles. The words in the caption will often point to a certain path of perception and shape how the reader-viewer navigates through and understands the iconic text (Serafini, 2012). In this interpretive scenario, written language anchors the image, pulling the image into the confines of the narrative detailed by the words.

Rio was one of the first students who asserted the authority of language, stating, "Our stories could be mildly accurate, or words could destruct our stories." Amy wrote that words "will help me improve and find the true meaning/plot of the story rather than just pictures because our thoughts could be placed anywhere, and our imaginations can go wild." Adam, echoing the statement that Mako will later inject into the class discussion, stated that language "will put us all on the same page to get one answer." Jayli explained that the images "made the viewers wonder and speculate. But just being given

the story takes all that away.” In the midst of this discussion, Mako shared the opinion quoted at the opening to this article: that words helped the reader understand the “true story” and that the stories they created yesterday went “past the truth.” Mako’s writing (see Figure 4) testified to this perspective.

Only two students in the class expressed different ideas about how words affected their experience of reading the text. Ryo expressed words as “equal[ing] another point of view.” Alberto also believed that the words would add another point of view and, as points of view, readers should “continue viewing them [words] like opinions.” Although their classmates advocated for the clarity of language and the closure words would provide, these two students maintained that they, as readers, could still determine meanings and offer individual interpretations.

The students’ repeated use of “will” when explaining what words will do showed a collective commitment to the power of written language to represent and communicate the truth, or, as Bailey averred, “the one correct answer.” Hodge and Kress (1988) emphasized that “truth” and “reality” are socially constructed categories expressed by groups of people who are acting and thinking in solidarity with each other. The versions of truth and reality that prevail are secured by dominant social collectives through an affinity that secures their power. In this classroom, the dominant view was



IN MY OPINION, ADDING WORDS WILL DEFINITELY PROVIDE SOME MUCH-NEEDED CLOSURE, BUT MAY END UP BEING DISAPPOINTING DUE TO ALL OUR RESPONSES FROM YESTERDAY. WHATEVER WILL HAPPEN, I WOULD MUCH RATHER BE UNIMPRESSED THAN REMAIN IN THE DARK.

Figure 4. Mako’s written reflection on how words alter the experience of reading Tan’s (2009) narrative

that written language will present the truth. The least powerful viewpoint, shared by only two students, indicated that this viewpoint could be contested. However, their shared affinity was not significant enough to overturn the prevailing ideas on written language and visual images.

What is noticeable about the students’ responses is that closeness to truth is not determined intramodally through an analysis of language-use or visual design features within the individual modes of language and image. Instead, without even viewing the accompanying written language, the students argued for words as the harbinger of truth. Closeness to truth in representational form was determined by students outside of both sign systems. Within this essayist and assessment-oriented educational environment (Gee & Hayes, 2011), written language was designated as the more authoritative and factual mode.

Written language as site of inherent meaning.

With regards to Tan’s (2009) story, the majority of these students believed that language would reveal the “truth,” a “correct answer,” or “the real story,” even in a fictional text where, essentially, truth and reality are purposefully constructed by a human author. It is no surprise, perhaps, that when students read the entire story, they acted disappointed. They expected language to go beyond anchoring the image and seemed to want language to do more than provide direction for their interpretation. Most of the students expected language to offer up a single, stable meaning to the story and provide closure to their reading of the text. This story, though, defied students’ expectations. The words in this story did not provide closure or textual coherence.

After reading the whole text, Ms. Scarlett asked students to share their immediate reactions. Once again, they were given time to reflect in writing before sharing with the class. Moses described the

story as “unsatisfying.” Mako “really wanted it to be a good story” but determined that, in terms of meaning, “There was really nothing there.” Arvind stated that, “It’s not even really a story.” Rya elaborated on their disappointment: “The images and the words shift our attention. The images hide the humans, while the words show them. The words don’t reveal anything about the stick figures.” These reactions to the modal dissonance in the story resulted in a seemingly collective decision that the narrative had no meaning and that the narrative was simply a bad story. Moses argued that the author had no purpose and concluded that, “If an author had no purpose, it’s a pointless text.”

At this point, only Alberto rallied against popular opinion in this classroom. In line with Iser’s (2000) scholarship, Alberto stated that words helped the reader fill in some gaps and that images helped fill in other blanks. Alberto stated that there is one more blank, and that’s for the reader to fill in. Alberto concluded by stating, “Each person fills that blank in differently, which is why we get multiple perspectives.”

Ms. Scarlett quieted the class for a moment and reminded students that Shaun Tan was not available to comment on their precise purpose in composing this story. Ms. Scarlett then discussed students becoming “comfortable with ambiguity and not always having the answers.” She continued: “I don’t teach you so that you can have one answer. I’m here to teach you how to think.” At this point, Arvind lightened the mood in the classroom, saying loudly, “I need a Tylenol. This class has exhausted me.” The class laughed, except for Mako. “Are we all okay?” Ms. Scarlett asked the class. Most nodded. Mako still appeared frustrated and gave voice to the frustration that remained. Mako’s final contribution to the discussion was significant: “We are taught that everything has a definitive answer. Why are we taught that if it’s not true?”

Later, in a voice-recorded reflection, Ms. Scarlett acknowledged the tension that this discussion about Tan’s text had generated. She had planned another task for the end part of the lesson but had felt compelled to continue with the debate over words, images, and meaning. In the voice message, Ms. Scarlett said:

I also decided during seventh-period to scrap the lesson and continue pushing the students with the conversation about the effect of words and images together when it came to *Stick Figures*. Because they were telling me that they wanted specific and clear answers. I think this is a condition of maybe the school I teach at, I think this is a condition of adolescence where we’re told—students are told—that there is one right answer. There’s an opportunity to teach students to be okay with ambiguity and to be okay with no resolution, and to be okay with multiple perspectives when we introduce words and images together.

For Ms. Scarlett, this multimodal short story and the discussion that ensued became an opportunity to discuss truth, reality, and fiction, as social constructs. The class returned to this discussion repeatedly throughout the curriculum unit. This text enabled Ms. Scarlett to help these students delve further into the notion of *representation*. Just as images will never mirror words in perfect symmetry (Lewis, 2001), neither will the signs belonging to any sign system precisely replicate the referent (Peirce, cited in Dressman, 2015), the thing or concept in the world that a sign represents.

Students' Preference for Word-Image Coherence

Despite the antagonism, tension, and debate in the second lesson over word-image relationships and meaning, and despite Ms. Scarlett introducing students to elements of Kress and van Leeuwen's (2006) visual grammar and allowing them to self-select reading material from a range of multimodal texts, the students appeared attached to the notion that images are supplementary to words. Nodelman (1988) talked about how words affect how readers see images in picturebooks but also stated that the reverse was true: that pictures can also affect and change how people read and interpret words. Students seemed primarily attached to Nodelman's former assertion, and most students did not, in these four lessons, entertain the latter assertion.

At the conclusion of the four-lesson sequence, Ms. Scarlett asked students to reflect on the classroom discussion, Tan's (2009) short story, and the images from Smy's (2017) novel. Ms. Scarlett then asked students to represent in writing, images, or a combination of the two their understanding regarding how words and images work together. Because the discussion in the class had become quite heated, I had expected students to problematize word-image relationships and express that there were word-image combinations that disrupted their reading and altered their meaning-making expectations. Instead, students seemed to default to coherence akin to symmetry (Nikolajeva & Scott, 2000) in word-image relationships, seemingly confirming their adherence to the idea of one kind of relationship as opposed to the existence of multiple interpretive possibilities.

This default to word-image coherence and modal unity was discernible across the four days that this

study took place. Predominantly, students' understanding of word-image relationships positioned image as subordinate to words and illustrative of the story the words were telling. Ryo, for example, during the debate in lesson two, stated: "Images should extend the words and be clearly connected to the words. That's what pictures do." Ryo spoke as if there are strict rules and conventions regarding word-image relationships in texts, and these words suggest Ryo's rejection of alternative word-image possibilities. Even though the class had just read a text in which the images are not directly connected to the words, Ryo persisted with the notion that images *should* be connected to the words, their connection unproblematic even if the images provide more information. Ryo spoke as if

their preferred word-image relationship were fact.

"To default to word-image coherence is to limit access to the meaning potential of iconic texts and the use of image in both interpretive and composition activities."

Ryo's assertion about the work that images do when paired with words was also repeated throughout the students' final written reflections in this sequence of lessons. Seven students talked about how images "extend" words. Two

students wrote about how images should "enhance" words. When students drew or developed verbal similes or metaphors to capture their understanding of word-image relationships, several students showed words carrying or supporting images (see Figure 5 for two examples). However, the most frequent comparative statement made was that words and images go together like peanut butter and jelly (see Figure 6 for an example), thus presenting words and images as different but complementary. The ingredients are better together than they are alone, but, together, they work well.

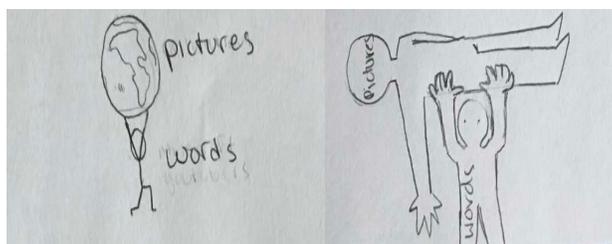


Figure 5. Moses's and Ryo's visual representation of how words and images interact

Only Amu represented image as a significant mode with important meaning-making resources or potential. Amu described how images “can trigger different feelings. Size and proportion can also do this.” Amu's visual depiction of images and words show a lively and animated image being turned into a seated, tired-looking book containing words and lines (see Figure 7). However, Amu's illustration also suggested a preference for the mode of image. The renderings of the two books with images on their pages show the books as expressive and gesturing animatedly. The third book, containing only words, is seated on a stool, its eyes sleepy.

Overall, the students' work suggested that they did not view image as a significant source of meaning potential. Indeed, as Nikolajeva and Scott's (2000) word-image interaction spectrum suggests, there are instances when images do appear to reflect the words and complement or enhance them. However, as Nikolajeva and Scott (2000), as well as other scholars (Lewis, 2001; Serafini, 2011, 2014; Sipe, 2012), have pointed out, this is not the only way that images and words interact. To default to word-

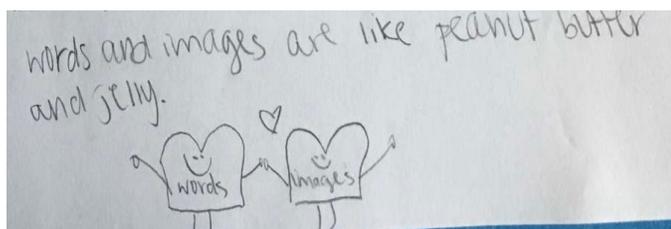


Figure 6. Chloe's visual representation of how words and images interact

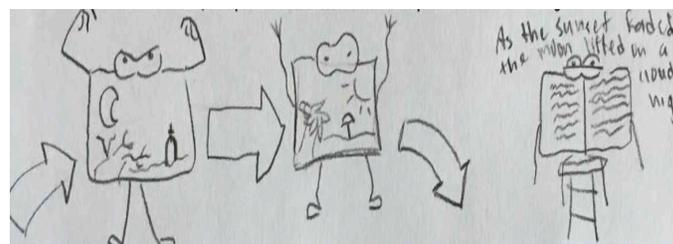


Figure 7. Amu's visual representation of how words and images interact

image coherence is to limit access to the meaning potential of iconic texts and the use of image in both interpretive and composition activities.

Discussion

This study highlights three important issues for literacy education researchers and classroom practitioners to consider. First, the study suggests that these students had been socialized to conceptualize textual interpretation as a search for one inherent and agreed-upon meaning. Second, the study highlighted the need for multimodal literacies instruction at the secondary level that supports students in understanding how different modes of representation and communication work in combination. Third, this study suggests the necessity of critical literacy frameworks designed for use with multimodal texts so that students can challenge and critique representations communicated through any mode.

Students Continue to Search for One Inherent Meaning

One of the key premises of a social semiotics approach to multimodality is that meaning is made anew each time an idea or text is communicated (Kress, 2010). As meaning-makers and text-interpreters themselves, students pay attention to certain criteria when they encounter text while throwing other criterial possibilities into shadow (Bezemer & Kress, 2016). As Hodge and Kress (1988) state, “The interpretation of texts is always a matter of guesses not facts” (p. 168).

My study highlighted how high-achieving students deemed to be proficient or highly proficient (as determined by state and other standardized tests) may not be prepared to interrogate texts or engage in dialogue regarding multiple interpretations. Students became frustrated when a single meaning for the text could not be determined, because they had been taught to search out main ideas. Instead of accepting the existence of multiple interpretations, some students declared that the text's meaning was null and void. Because they could not identify a single meaning or compress the words and images into tight alignment with each other, students dismissed the text as meaningless. This study suggests that these students may have learned unproductive and restrictive versions of reading and may have "suspended their interpretive authority" in favor of "locating a single authoritative meaning" (Davis & Vehabovic, 2018, p. 584).

Davis and Vehabovic (2018) called upon teachers to resist practices that limit the meaning-making abilities of students and the meaning potential of texts, and, throughout the study, Ms. Scarlett sought to answer this call by resisting the test-centric practices to which students in this class had grown accustomed. She talked directly with students about defying the habit of searching for a single meaning. But what happens when Ms. Scarlett's students move into new English Language Arts contexts? Will they encounter a teacher willing to resist test-centric literacy practices? Or a teacher that will re-establish test-oriented literacy practices (Gee & Hayes, 2011). These teacher effects are a matter to which all stakeholders in literacy education should attend. Research in classroom contexts and the communication of that research to multiple stakeholders (including school administrators and policymakers) continues to be essential.

Students Continue to Need Multimodal Literacies Instruction

The need for multimodal literacies instruction is not a new observation. For example, while Siegel (2012) argued that school literacy curricula must change in order to reflect the changing literacy practices embedded within the world beyond school, they also noted that "the expansion of modes for meaning-making" is a storyline colliding with "the confinement of this multiplicity through the practices of accountability" (p. 675). My study indicates the continued need to advocate for multimodal texts and composition learning experiences in schools. Not only were the students in my study searching for one singular meaning, they were also searching for that meaning within one single mode of representation and communication: written language.

Serafini (2015) contends that a multimodal literacies curriculum would enable students to consider the semiotic options made available through single modes of representation and communication, and attend to how meaning can be constructed across combinations of multiple modes. A multimodal literacies curriculum requires teachers to diverge from traditional, canonical literature that privileges written language in order to incorporate multimodal texts and concepts into their instruction. However, although I believe that visual literacy frameworks (Callow, 2008; Cappello, 2017) will help teachers support their students in interpreting images and navigating design features, teachers will also need to learn new metalanguage and analytical perspectives (Callow, 2018; Serafini, 2015) in order to make pedagogical moves that address the intermodal relationships at play in texts.

My study further highlights two additional points related to the ongoing need for multimodal literacies instruction in school. First, the study points to a potential concern. The students in Ms.

Scarlett's classroom were quick to dismiss Tan's story as meaningless when they realized that the words and images did not complement or cohere with each other in an expected way. Instead of viewing this dissonance as evidence of text complexity, most students determined that the dissonance resulted in a text so simple they could not "drain" any meaning from it (Mako's words). If students dismiss dissonant modal combinations as meaningless or a "bad story" (also Mako's words), what will they do when they encounter other texts that lack modal complementarity? Students risk under-developing their interpretive repertoires (Serafini, 2015). I believe that education researchers need to advance, promote, and circulate expanded conceptualizations of text complexity. These ideas go beyond traditional linguistic formulas to address the multiple kinds of intermodal relationships possible within texts (Serafini, 2014).

The second aspect my study revealed is students' capacity to theorize and engage in discussions about how different modes work in isolation and together. Without any instruction in word-image scholarship, students articulated key terms frequently used by leading children's literature scholars. For example, students talked about word-image relationships in terms of symmetry, complementariness, enhancement, extension, or expansion, all of which touch upon ideas found in Nikolajeva's and Scott's (2000) spectrum of word-image relationships. Ms. Scarlett asked students to consider more typical literary qualities such as theme and plot, but she also engaged them in representing and communicating ideas about multimodality.

Although Ms. Scarlett's students needed more exposure to multimodal texts featuring additional kinds of word-image relationships and additional

experience addressing the gaps left when modal components do not align, they engaged in passionate discussion and dialogue about complex theoretical ideas. Theoretical knowledge about how multimodal texts work spans texts and interpretive experiences and is not limited to the reading of particular texts in particular contexts.

Students Need Multimodal and Critical Literacies Instruction

The final aspect of my discussion stems from the majority of students articulating that the words would reveal a single truth. It is interesting that students would seek truth in a fictional story and believe that truth can be located. It is also concerning that the students who sought truth

“Students need to understand that any representational work can be challenged, critiqued, and, ultimately, rejected.”

searched for it only in written language. In a culture replete with visual and graphically-designed texts (Barnard, 2001), students need to be aware that images represent and circulate myths that naturalize and normalize the ideologies

belonging to dominant groups (Aiello, 2006; Barthes, 1977). It is not only the images in more obviously manipulative texts such as print advertisements or television commercials that require analysis, critique, and challenge. Students should also be prepared to analyze, challenge, and critique the images that populate the fictional texts we consume: Disney movies (Ajayi, 2011, 2012), picturebooks (Serafini, 2010), or a fictional multimodal short story about stick figures, for example. To dismiss images is to be immune to the role they play in shaping readers' social settings.

Aiello (2006) described the three kinds of reading that Stuart Hall (1997), a Cultural Studies scholar, believed can take place in response to a multimodal or media message (Aiello, 2006, p. 98). Hall argued that a *dominant reading* involves the reader

accepting the ideological messages imbued in the text by the privileged and powerful. An *oppositional reading* requires the reader to challenge and reject the hegemonic messages. A *negotiated reading* occurs when the reader accepts aspects of the dominant reading but challenges and rejects other aspects. Approaching the reading of multimodal texts from Hall's perspective would provide students with the opportunity to negotiate their readings of messages made material through multiple modes of representation and communication. Students need to understand that any representational work can be challenged, critiqued, and, ultimately, rejected. As Peel (2017) observes, "Truly rigorous reading does not rely on the inherent complexity of text, but on the complicating of the text by the reader. Rigorous reading needs to be more than hard; it needs to be liberatory" (p. 106).

Limitations and Future Research

This study has multiple limitations that signal the need for future research. Primarily, the study spans four days in one English Language Arts classroom and examines students' responses to a single multimodal narrative text. Thus, extending this research to additional classrooms within other schools and districts would enable researchers to construct assertions (Erickson, 1986) that can be drawn across the individual contexts in which readers and multimodal texts are situated. I also believe that more long-term formative and practice-based research (Hinchman & Appleman, 2017) in which teachers, students, and researchers work together to examine and build understandings of

word-image relationships across multiple genres, mediums, and even disciplines would be advantageous to the development of multimodal literacies curricula and practices in K-12 classrooms. Furthermore, although classroom discussions and student interactions with multimodal texts were not audio or video-recorded in this study, future research could employ these data collection strategies to examine such literacy events in more fine-grained detail.

Conclusion

This study provides insight into the frustrations experienced by students when word-image relationships within a multimodal short story did not meet their expectations for how different modes should interact within a narrative text. However, the study also reveals broader issues and concerns regarding students' assumptions about texts, images, and language. A multimodal literacies approach to curriculum and instruction could help students actualize their more complete capacities as meaning-makers and support students in realizing their agency during meaning-making work. In an era where differing accounts of truth and reality coexist, engaging students in analytical and critical work on representations constructed through multiple modes is necessary work. Such work could help students realize their options as readers and consumers of information. Every representation can be accepted, partially challenged, or wholly rejected (Hall, 1997, cited in Aiello, 2006). Representation is always a matter of reality removed. Representation can never be reality itself.

References

- Aiello, G. (2006). Theoretical advances in critical visual analysis: Perception, ideology, mythologies, and social semiotics. *Journal of Visual Literacy*, 26(2), 89–102.
- Ajayi, L. (2011). A multiliteracies pedagogy: Exploring semiotic possibilities of a Disney video in a third-grade diverse classroom. *The Urban Review*, 43(3), 396–413.
- Ajayi, L. (2012). Video “reading” and multimodality: A study of ESL/literacy pupils’ interpretation of Cinderella from their socio-historical perspective. *The Urban Review*, 44(1), 60–89.
- Barnard, M. (2001). *Approaches to understanding visual culture*. New York, NY: Palgrave.
- Barthes, R. (1977). Rhetoric of the image. In S. Heath (Ed.), *Image - music - text* (pp. 32–51). New York, NY: Hill and Wang.
- Bazeley, P. (2013). *Qualitative data analysis: Practical strategies*. Thousand Oaks, CA: SAGE.
- Bezemer, J., & Kress, G. (2016). *Multimodality, learning and communication: A social semiotic frame*. New York, NY: Routledge.
- Calkins, L., Ehrenworth, M., & Lehman, C. (2012). *Pathways to the Common Core: Accelerating achievement*. Portsmouth, NH: Heinemann.
- Callow, J. (2008). Show me: Principles for assessing students’ visual literacy. *The Reading Teacher*, 61(8), 616–626.
- Callow, J. (2018). Classroom assessment and picture books: Strategies for assessing how students interpret multimodal texts. *Australian Journal of Language & Literacy*, 41(1), 5–20.
- Cappello, M. (2017). Considering visual text complexity: A guide for teachers. *The Reading Teacher*, 70(6), 733–739.
- Cowan, K., & Kress, G. (2017). Documenting and transferring meaning in the multimodal world: Reconsidering transcription. In F. Serafini & E. Gee (Eds.), *Remixing multiliteracies: Theory and practice from New London to new times* (pp. 50–62). New York, NY: Teachers College Press.
- Davis, D. S., & Vehabovic, N. (2018). The dangers of test preparation: What students learn (and don’t learn) about reading comprehension from test-centric literacy instruction. *The Reading Teacher*, 71(5), pp. 579–588.
- Davis, D. S., & Willson, A. (2015). Practices and commitments of test-centric literacy instruction: Lessons from a testing transition. *Reading Research Quarterly*, 50(3), 357–379.
- Dressman, M. (2015). Reading as the interpretation of signs. *Reading Research Quarterly*, 51(1), 111–136.
- Emerson, R. M., Fretz, R. I., & Shaw, L. L. (2011). *Writing ethnographic fieldnotes*. Chicago, IL: University of Chicago Press.
- Erickson, F. (1986). Qualitative methods in research on teaching. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (pp. 119–161). New York, NY: Macmillan.
- Fisher, D., & Frey, N. (2014). *Text-dependent questions, grades 6-12: Pathways to close and critical reading*. Thousand Oaks, CA: Corwin Press.

- Garcia, A. (2017). A literature of possibilities: Games, agency, and redefining texts. *Journal of Adolescent & Adult Literacy*, 61(3), 339–341.
- Gee, J. P., & Hayes, E. R. (2011). *Language and learning in the digital age*. New York, NY: Routledge.
- Hall, S. (1997). The work of representation. In S. Hall (Ed.), *Representation: Cultural representations and signifying practices* (pp. 13–74). London, UK: Sage.
- Hinchman, K. A., & Appleman, D. A. (2017). Introduction. In K. A. Hinchman & D. A. Appleman (Eds.), *Adolescent literacies: A handbook of practice-based research* (pp. xiii–xxii). New York, NY: Guilford.
- Hodge, R., & Kress, G. (1988). *Social semiotics*. Ithaca, NY: Cornell University Press.
- Hull, G. A., & Nelson, M. E. (2005). Locating the semiotic power of multimodality. *Written Communication*, 22(2), 224–261.
- Iser, W. (2000). *The range of interpretation*. New York, NY: Columbia University Press.
- Jewitt, C. (2017). An introduction to multimodality. In C. Jewitt (Ed.), *The Routledge handbook of multimodal analysis* (pp. 15–30). New York, NY: Routledge.
- Jewitt, C., Bezemer, J., & O'Halloran, K. (2016). *Introducing multimodality*. New York, NY: Routledge.
- Jiménez, L. M., & Meyer, C. K. (2016). First impressions matter: Navigating graphic novels utilizing linguistic, visual, and spatial resources. *Journal of Literacy Research*, 48(4), 423–447.
- Jiménez, L. M., Roberts, K. L., Brugar, K. A., Meyer, C. K., & Waito, K. (2017). Moving our can(n)ons: Toward an appreciation of multimodal texts in the classroom. *The Reading Teacher*, 71(3), 363–368.
- Khadka, S., & Lee, J. C. (2019). *Bridging the multimodal gap: From theory to practice*. Boulder, CO: University Press of Colorado.
- Kress, G. (2010). *Multimodality: A social semiotic approach to contemporary communication*. New York, NY: Routledge.
- Kress, G., & van Leeuwen, T. (2001). *Multimodal discourse: The modes and media of contemporary communication*. London, UK: Arnold Publishers.
- Kress, G., & van Leeuwen, T. (2006). *Reading images: The grammar of visual design*. New York, NY: Routledge.
- Lemke, J. (1998). Multiplying meaning. In J. R. Martin & R. Veel (eds.), *Reading science: Critical and functional perspectives on discourses of science* (pp. 87–113). New York, NY: Routledge.
- Lewis, D. (2001). *Reading contemporary picturebooks: Picturing text*. New York, NY: Routledge.
- McCloud, S. (1993). *Understanding comics: The invisible art*. New York, NY: HarperCollins.
- Meek, M. (1992). Children reading--now. In M. Styles, E. Bearne, & V. Watson (Eds.), *After Alice: Exploring children's literature* (pp. 172–187). London, UK: Cassell.
- Miles, M. B., Huberman, A. M., & Saldana, J. (2019). *Qualitative data analysis: A methods sourcebook*. Thousand Oaks, CA: Sage.

- Mills, K. A., & Exley, B. (2014). Narrative and multimodality in English language arts curricula: A tale of two nations. *Language Arts*, 92(2), 136–143.
- National Council of Teachers of English. (2005). *Multimodal literacies: NCTE position statement*. Urbana, IL: Author. Retrieved from <http://www2.ncte.org/statement/multimodalliteracies/>
- New London Group. (1996). A pedagogy of multiliteracies: Designing social futures. *Harvard Educational Review*, 66(1), 60–93.
- Nikolajeva, M., & Scott, C. (2000). The dynamics of picturebook communication. *Children's Literature in Education*, 31(4), 225–239.
- Nodelman, P. (1988). *Words about pictures: The narrative art of children's picture books*. Athens, GA: University of Georgia Press.
- O'Brien, D. (2012). "Struggling" adolescents' engagement in multimediating: Countering the institutional construction of incompetence. In D. E. Alvermann & K. A. Hinchman (Eds.), *Reconceptualizing the literacies in adolescents' lives: Bridging the everyday/academic divide* (pp. 71–92). New York, NY: Routledge.
- Olson, K. (2011). *Essentials of qualitative interviewing*. Walnut Creek, CA: Left Coast Press.
- Overstreet, R. R. (2018). When stories transcend books: Ideas for teaching transmedia. In J. S. Dail, S. Witte, & S. T. Bickmore (Eds.), *Young adult literature and the digital world: Textual engagement through visual literacy* (pp. 17–27). Lanham, MD: Rowman & Littlefield.
- Pantaleo, S. (2012). Exploring grade 7 students' responses to Shaun Tan's *The Red Tree*. *Children's Literature in Education*, 43(1), 51–71.
- Pantaleo, S. (2019). Metalepsis in elementary students' multimodal narrative representations. *Research in the Teaching of English*, 54(1), 8–30.
- Peel, A. (2017). Complicating canons: A critical literacy challenge to common core assessment. *Literacy Discussion*, 51(2), 104–110.
- Renaissance Learning. (2019). *What kids are reading: World's largest annual study of K–12 reading habits: 2019 edition.*: Wisconsin Rapids, WI: Author.
- Saldaña, J. (2016). *The coding manual for qualitative researchers*. Thousand Oaks, CA: Sage.
- Seglem, R., & Witte, S. (2009). You gotta see it to believe it: Teaching visual literacy in the English classroom. *Journal of Adolescent & Adult Literacy*, 53(3), 216–226.
- Serafini, F. (2010). Reading multimodal texts: Perceptual, structural and ideological Perspectives. *Children's Literature in Education*, 41(2), 85–104.
- Serafini, F. (2011). Expanding perspectives for comprehending visual images in multimodal texts. *Journal of Adolescent & Adult Literacy*, 54(5), 342–350.
- Serafini, F. (2012). Expanding the four resources model: Reading visual and multi-modal texts. *Pedagogies: An International Journal*, 7(2), 150–164.

- Serafini, F. (2014). *Reading the visual: An introduction to teaching multimodal literacy*. New York, NY: Teachers College Press.
- Serafini, F. (2015). Multimodal literacy: From theories to practices. *Journal of Children's Literature*, 92(6), 412–423.
- Siegel, M. (2012). New times for multimodality? Confronting the accountability culture. *Journal of Adolescent & Adult Literacy*, 55(8), 671–681.
- Sipe, L. R. (2012). Revisiting the relationships between text and pictures. *Children's Literature in Education*, 43(1), 4–21.
- Slota, S. T., Young, M. F., O'Byrne, W. I., & Ballestrini, K. (2016). A new hope: Negotiating the integration of transmedia storytelling and literacy instruction. *Journal of Adolescent & Adult Literacy*, 59(6), 642–646.
- Smith, B. E. (2017). Composing across modes: A comparative analysis of adolescents' multimodal composing processes. *Learning, Media and Technology*, 42(3), 259–278.
- Smy, P. (2017). *Thornhill*. New York, NY: Roaring Brook Press.
- Stotsky, S. (2010). Literary study in grades 9, 10, and 11: A national survey. *Forum: A Publication of the Association of Literary Scholars, Critics, and Writers*, 4, 1–75.
- Tan, S. (2009). Stick figures. In S. Tan (Author & Illustrator), *Tales from Outer Suburbia* (pp. 64-69). New York, NY: Arthur A. Levine.
- Thiel, J. J., & Kuby, C. R. (2018). Careful! There are monsters in this chapter. In J. J. Thiel, K. Spector, & C. R. Kuby (Eds.), *Posthumanism and literacy education: Knowing/becoming/doing literacies*. New York, NY: Routledge.
- Unsworth, L., & Cléirigh, C. (2017). Multimodality and reading: The construction of meaning through image-text interaction. In C. Jewitt (Ed.), *The Routledge handbook of multimodal analysis* (pp. 176–188). New York, NY: Routledge.
- Wohlwend, K. (2017). The expression of multiliteracies and multimodalities in play. In F. Serafini & E. Gee (Eds.), *Remixing multiliteracies: Theory and practice from New London to new times* (pp. 162–175). New York, NY: Teachers College Press.