

## Reading Multimodal Texts: Perceptual, Structural and Ideological Perspectives

Frank Serafini

Published online: 2 April 2010  
© Springer Science+Business Media, LLC 2010

**Abstract** This article presents a tripartite framework for analyzing multimodal texts. The three analytical perspectives presented include: (1) perceptual, (2) structural, and (3) ideological analytical processes. Using Anthony Browne's picturebook *Piggybook* as an example, assertions are made regarding what each analytical perspective brings to the interpretation of multimodal texts and how these perspectives expand readers' interpretive repertoires. Drawing on diverse fields of inquiry, including semiotics, art theory, visual grammar, communication studies, media literacy, visual literacy and literary theory, the article suggests an expansion of the strategies and analytical perspectives readers bring to multimodal texts and visual images. Each perspective is presented as necessary but insufficient in and of itself to provide the necessary foundation for comprehending texts. It is through an expansion of the interpretive strategies and perspectives that readers bring to a multimodal text, focusing on visual, textual, and design elements that readers will become more proficient in their interpretive processes.

**Keywords** Multimodal texts · Picturebooks · Interpretation · Comprehension · Visual grammar

Multimodal texts, in particular contemporary picturebooks, are used extensively in many elementary reading programs, conveying meanings through the use of two sign systems; written language and visual image. However, the primary focus in

---

Dr. Frank Serafini is an author, illustrator, photographer, educator, musician, and an Associate Professor of Literacy Education and Children's Literature at Arizona State University.

---

F. Serafini (✉)  
College of Teacher Education and Leadership, Arizona State University,  
8084 E. High Point Dr., Scottsdale, AZ 85257, USA  
e-mail: serafini@asu.edu  
URL: [www.frankserafini.com](http://www.frankserafini.com); [fserafini@mac.com](mailto:fserafini@mac.com)

elementary reading education has been on the strategies and skills necessary for understanding written language. This lack of pedagogical attention to visual images and visual systems of meaning presents serious challenges to teachers at a time when image has begun to dominate the lives of their students (Fleckenstein, 2002; Kress, 2003). While numerous resources are readily available that focus on strategies for reading and comprehending written text (Harvey and Goudvis, 2000; Owocki, 2003; Snow and Sweet, 2003), pedagogical approaches addressing various strategies for comprehending visual images, in particular those included in contemporary picturebooks, have only recently emerged in elementary pedagogical discussions (Anstey and Bull, 2006; Albers, 2008).

Visual images are drawn upon with increased frequency to make sense of one's world, often overshadowing the once dominant mode of written language. Anne Seward Barry (1997) recognized, "that the visual mode of communication now dominates verbal communication within the American culture cannot be doubted, and it is past time to realize that traditional methods of analysis must be expanded to include perceptual principles and visual mosaic logic as well" (p. 7). Nicholas Mirzeoff (1998) suggests that contemporary society is *ocularcentric*, not because visual images are more common, but because our experiences are better understood as totally constructed visual experiences.

As readers interact with images on the computer screen, as well as in printed materials, we must remember the world shown is different from the world told (Kress, 2003). The semiotic resources used to create multimodal texts are different from the ones drawn upon to create printed texts, and bring with them different potentials for making meaning. This shift from a linguistic focus to a multimodal one requires readers to navigate, design, interpret and analyze texts in new and more interactive ways (Unsworth, 2002; Anstey and Bull, 2006).

A shift from the single mode of written language to multimodal texts that include extensive design and visual elements requires a parallel shift in the strategies and skills required by readers. Len Unsworth and Janet Wheeler (2002) suggest, "if children are to learn how to analyze the ways images make meanings, they need to gain knowledge of the visual meaning-making systems deployed in images" (p. 69). In order to successfully interact with and interpret the meanings of the visual images and graphic designs included in multimodal texts, readers must employ a new set of strategies that extend beyond the various cognitive-based reading comprehension strategies used to understand written text.

To expand one's interpretive repertoires to include approaches to visual analysis, readers will be required to synthesize perceptual abilities with structural approaches and political, historical, and cultural understandings. Giorgia Aiello (2006) states, "In analyzing images, then, it is necessary to account not only for their cultural norms, but also for their perceptual qualities" (pp. 89–90). A shift in attention from the basic design elements, objects and semiotic resources used in creating visual images and multimodal texts to the socio-cultural contexts of production and reception is necessary, but should not abandon the analytical approaches put forth by perceptual psychology (Arnheim, 1986; Seward Barry, 1997), visual grammar and semiotics (Kress, 2003; Kress and van Leeuwen, 1996; van Leeuwen and Jewitt, 2001), visual communications (Leeds-Hurwitz, 1993; Smith-Shank, 2004), and

visual literacy (Elkins, 2008; Duncum, 2004). The images contained in multimodal texts encountered in complex social contexts are created with particular semiotic resources, basic design elements, and visual structures. To ignore the perceptual and structural aspects of visual images and multimodal texts in favor of a socio-cultural perspective would limit readers' interpretive repertoire and forego relevant perspectives for making sense of images and multimodal texts.

## Multimodal Texts

Multimodal texts present information across a variety of modes including visual images, design elements, written language, and other semiotic resources (Jewitt and Kress, 2003). The mode of written language and that of visual image are governed by distinct logics; written text is governed by the logic of time or temporal sequence, whereas, visual image is governed by the logic of spatiality, organized arrangements, and simultaneity (Kress, 2003). In written text, meaning is derived from position in the temporal sequence, whereas meaning is derived in visual images from spatial relations or visual grammar (Kress and van Leeuwen, 1996).

Readers in schools today interact with traditional texts that contain multimodal elements, for example picturebooks, informational texts, magazines and newspapers, as well as contemporary texts that contain hypertext, videos, music, and graphic designs. In fact, most contemporary written texts are accompanied by visual images. Paul Duncum (2004) states, "... there is no avoiding the multimodal nature of dominant and emerging cultural sites" (p. 259). Images and texts are being combined in unique ways, and readers in today's world need new skills and strategies for constructing meaning in transaction with these multimodal texts as they are encountered during the social practices of interpretation and representation (Serafini, 2009).

## Visual Literacy and Perception

Visual literacy, defined as the ability to access, analyze, evaluate, and communicate information in any variety of form that engages the cognitive processing of a visual image (Chauvin, 2003), combines psychological theories of perception with the socio-cultural and critical aspects of visual design, social semiotics, and media studies. Merging physiological and cognitive theories that focus on vision and perception with the social and cultural aspects of literacy and interpretation allows for a more inclusive and expansive analytical approach to multimodal texts.

Visual perception is conceptualized as more than the passive impression of light on the retina. It is viewed as a dynamic process in which the brain, largely automatically, filters, discards, and selects information, and compares it to an individual's stored record (Stafford, 2008). Neuroscientists have argued, "... the brain is only interested in obtaining knowledge about those permanent, essential, or characteristic properties of objects and surfaces that allow it to categorize them" (Zeki, 1999, p. 77). In other words, before images are interpreted in the social

contexts of their production, reception and dissemination, qualities of the image must be perceived, processed, and categorized. These perceptual processes, as well as the socio-cultural contexts in which an image or multimodal text are viewed, are important considerations for understanding the meanings readers and viewers construct during transactions with multimodal texts.

In order to be successful understanding multimodal texts, readers will need to draw upon a variety of converging and interconnected perspectives. To fully participate in today's cultural and political contexts one must become competent in the design, production, and dissemination of representations as messages (Kress, 2010). Each of the interconnected perspectives presented in this article provides a different focus, a different set of interpretive strategies and analytical tools, calling readers' attention to different aspects of the construction, design and analysis of multimodal texts.

### Three Analytical Perspectives

In this article, I propose an analytical framework for investigating multimodal texts, in particular reading, viewing, and responding to contemporary children's picturebooks that broadens the perspectives we can and should bring to these texts. The framework addresses three interconnected analytical perspectives, namely: (1) perceptual, (2) structural, and (3) ideological perspectives. These analytical perspectives do not exist in isolation. In fact, Roland Barthes (1977) suggests that the viewer of an image receives the perspective (denotative) message and the cultural (connotative) message simultaneously and that the denotative message is constituted by what remains in the image when one removes mentally the connotative sign. The blurry lines that exists between denotative and connotative messages will not be addressed in this discussion, however the possibility of these perspectives being discussed separately will be advocated.

Robert Scholes (1985) draws a similar distinction, suggesting his three dimensions of literary competence, reading, interpretation, and criticism, "cannot be divided into separate bits, but are sufficiently distinguishable for us to understand them ourselves, *and* (italics in original) for us to present them to our students, as discrete enterprises that can be practiced separately (p. 21). In similar fashion, the three analytical perspectives proposed in this article are sufficiently distinguishable from one another, and are therefore described and presented separately.

The three analytical perspectives may be conceptualized as a set of three concentric circles, with each perspective (perceptual, structural, and ideological) nested within subsequent perspectives rather than seen as a leveled hierarchy. The perceptual analytical perspective is located in the center of the three concentric circles, nested within the structural and ideological perspectives. E. H. Gombrich (1961) stated, "the innocent eye is a myth" referring to the idea that the unknowing mind sees nothing and therefore what we perceive is always conditioned by what we know (p. 278). In similar fashion, one's perceptions of multimodal texts are colored by prior knowledge, personal experiences, and socio-cultural and historical contexts of reception and production (Berger, 1972). Judith Graham (1990) states, "readers'

heads are not empty of images when they open a picture-book any more than they are empty of ideas when they open a text” (p. 18). Images and texts mean things because readers bring experiences and understandings of images, language and the world to them when reading.

The structural perspective, positioned in the middle ring, is nested within the ideological perspective and affected by the socio-cultural and historical contexts in which we interpret and compose multimodal texts. Kress and van Leeuwen (1996), drawing upon semiotic theory (Chandler, 2007) and systemic functional linguistics (Halliday, 1975, 1978), provided, “inventories of the major compositional structures which have become established as conventions in the course of the history of visual semiotics, and to analyze how they are used to produce meaning by contemporary image makers” (p. 1). These structures do not exist in isolation from social contexts and practices, rather are nested within the ideological or social semiotic milieu in which they are utilized.

Finally, the ideological perspective, conceptualized as the outer ring, is informed by our perceptions of the images and texts we encounter, the basic elements of design (Dondis, 1973), and the structures of visual images (Kress and van Leeuwen, 1996). Visual texts are *motivated* or developed to perform specific social actions or tasks, and various semiotic resources are utilized to create a field of possible meanings, which need to be activated by the producers and viewers of images (Jewitt and Oyama, 2001). The structures and visual resources of multimodal texts are always interpreted within a social context and through particular social practices, and the social practices draw upon the semiotic resources, basic design elements and perceptual qualities of semiotic materials employed in their conception.

Each of the three perspectives in the analytic framework draws on particular semiotic, art, visual design and socio-cultural theories. Similar to the four resources model (Freebody, 1992; Freebody and Luke, 1990), rather than privileging any single perspective over the others, each perspective should be considered necessary but insufficient in and of itself to render a viable interpretation or analysis of multimodal texts. Since no single perspective can provide a value-free, universal depiction of reality, each analytic perspective offers a different lens for investigating multimodal texts.

Additionally, the three analytical perspectives described in this article differ in epistemological assumptions, modes of analysis, and warrants used as evidence of interpretive viability. Trying to work across such contested terrain may seem a fool's errand, or worse a doomed project trying to solve incommensurable differences among theoretical paradigms (Moss et al., 2009). A disinterested and sympathetic attention to and contemplation of a literary text or work of art has been challenged by art and literary theorists (Goodman, 1976; Gombrich, 1961; Culler, 1997; Fish, 1980). Gombrich (1961) insists, “the eye comes always ancient to its work... not only how, but what it sees is regulated by need and prejudice (p. 279). Reception and interpretation are not separate mental operations, rather they are thoroughly interconnected processes, and any approach to understanding visual images or multimodal texts must acknowledge this interconnection. However, bypassing the forms, visual structures, design elements, and objects rendered in an

image or multimodal text to consider the socio-cultural influences and contexts of production and reception may mistakenly overlook the interpretive possibilities other analytical tools and approaches make available.

It is not my goal to theoretically dissect perception from interpretation, draw definitive or final distinctions among the three perspectives, nor attempt to cross the epistemological chasm that divides objectivism from postmodern or post-structural thought. Rather, my intention is to suggest that each perspective affords the reader an alternative viewpoint and set of analytical tools from which to consider multimodal texts and the visual images contained therein. Peggy Albers (2008) suggests, “scholarship must also turn to perspectives that are often unassociated with the field of literacy, which provide additional insight into how readers make sense of their worlds” (p. 167). Acknowledging that each perspective draws upon diverse fields of inquiry, provides alternative lenses and understandings, focuses one’s attention on different aspects of multimodal texts, and expand one’s interpretive repertoire for comprehension is paramount.

To clarify the distinctions made among the three analytical perspectives presented, I will conduct an abbreviated analysis of the contemporary picturebook *Piggybook* by Anthony Browne (1986) from each of the three analytical perspectives. *Piggybook* is the story of a mother, father, and two sons that focuses on gender roles and expectations and how members of the family believe they should be treated. When the father and two boys disregard the mother and the work she does around their house, the mother leaves unannounced. Subsequently, the house falls into disarray due to neglect and lack of domestic skills of the male members of the family. In the visual images provided by Browne, we see the men turn into pigs both literally and symbolically as they fail to maintain domestic routines.

Browne’s book provides notable opportunities for analyzing images, text, and design elements across the perceptual, structural and ideological perspectives presented in this article. The front cover of the book and four double-page spreads—or openings—have been purposefully selected to provide examples of each analytical approach.

## Theoretical Foundations

Literary, visual communication, and art-based researchers, educators and theorists have proposed analytical perspectives similar to the tripartite framework described in this article (Scholes, 1985; Berger, 1972; Panofsky, 1955). In general, these analytical frameworks begin with a literal or denotative level of meaning based on perception of visual and textual elements (attention to the physical effects of visual stimuli on the retina), and progress through various interpretive perspectives, where viewers construct meaning in transaction with visual texts and images, culminating in a critical or socio-cultural frame of analysis. Although this is certainly an oversimplification of some rather diverse and complex theories, differences among each of these perspectives are worthy of consideration before proceeding.

From a literary theory perspective, Scholes (1985) constructed an analytical framework across three interrelated dimensions: (1) reading, (2) interpretation, and (3) criticism. The first dimension, reading, focused on knowing the codes that operate in any given text, and the processing of text without confusion or delay. The second dimension, interpretation, suggested, a move from a summary of the events to a discussion of the meaning or theme of a work of fiction (Scholes, 1985). Criticism, the third dimension, involved a critique of themes in a given work of fiction or the codes from which a text has been constructed. Each of these three dimensions evokes a different literary practice. Scholes (1985) suggested the teacher's role was to demonstrate how these codes and processes are used in reading literary texts and to encourage students' own textual practices.

Erwin Panofsky (1955), an art historian and critic, introduced iconological methods in artistic interpretation and distinguished three strata for interpreting Renaissance art. The first strata, which Panofsky (1955) entitled, primary or natural meaning, consisted of the identification of "pure forms, that is: certain configurations of line and color ... as representations of natural objects" (p. 28). This strata focused on identifying visual data with objects known from experience (Hassenmuller, 1987). Panofsky described a second strata entitled, secondary or conventional meaning, which required viewers to move beyond the literal image to consider their experiences and social contexts. This second strata was concerned with connecting artistic motifs with themes or concepts. The third strata included intrinsic meanings or content. Panofsky (1955) suggested, "it is apprehended by ascertaining those underlying principles which reveal the basic attitudes of a nation, a period, a class, a religious or philosophical persuasion" (p. 30). These *three analogous levels of meaning* in art interpretation became the object of pre-iconographic, iconographic and iconological interpretative processes (Hassenmueller, 1987).

Panofsky's model of interpretation began with pre-iconographic descriptions, the recognition of pure forms or direct association of a new visual experience with memory (Hasenmueller, 1987). Iconography was considered an intellectual interpretation of secondary or conventional subject matter, themes and concepts. The identification of such images, themes and concepts, in contrast to identifying forms, distinguished the iconographic from the pre-iconographic. Taking the process a step further, iconological processes focused on ascertaining the underlying principles of a nation, culture, or period that was rendered through a work of art. Iconology can be viewed as an ideological interpretation of an image or work of art. Panofsky "... knew the difficulties of verifying the conclusions of investigation that transcended empirical data," but, progressing through the identification of forms, the construction of images and themes, to the interpretation of these images in light of the socio-cultural contexts of these images was the basis of his interpretive framework (Hasenmeuller, 1987, p. 291).

Gillian Rose (2001) distinguished three modalities that can contribute to a critical understanding of images, namely, (1) technological—any form or apparatus designed to be looked at, (2) compositional—formal structures or strategies, for example color, spatial organization or content, and (3) social—the range of economic, social and political relations, institutions and practices that surround an

image (pp. 16–17). Rose (2001) noted how some methodologies focused on the image in and of itself. This focus on composition, color, spatial organization, and light suggested an image's *expressive content* or the combined effect of visual form and subject matter. Although she acknowledged, “visual images do not exist in a vacuum, and looking at them for ‘what they are’ neglects the ways in which they are produced and interpreted through particular social practices” (Rose, 2001), viewers still needed to consider what was presented or rendered in an image and the structures and modalities that were used in creating images before proceeding to interpretation and ideological analysis.

In summary, these theorists draw from different fields of inquiry and offer a variety of analytical perspectives for looking at visual images, art, and multimodal texts. Each comes to the act of interpretation with different theoretical assumptions and analytical tools. Each theorist acknowledges the need to perceive what is rendered or presented in an image or text before interpreting what is depicted, and further consider the social, cultural, political and historical contexts of the reception and production of these texts and images. In concordance, Barbara Stafford (2008) distinguishes between *visual competence*, a baseline skill, like the ability to decode, “that is a necessary, but far from sufficient, condition for the more advanced and specialized skills” she defines as *visual literacy* (p. 13).

An important aspect of these various approaches, is the fact that each theorist tried to differentiate between what John Berger (1972) defined as *looking*, the physical act of light falling on the retina, and *seeing*, one's ability to transact with an image or text, to construct meaning and to situate ourselves within an image. Berger suggests that looking is a physiological or perceptual act, while seeing is an interpretive act, based on socio-cultural considerations and contexts. Rose (2001) distinguishes between *vision*, what the human eye is physiologically capable of seeing, and *visuality*, how vision is constructed in various ways. In order to provide a variety of analytical perspectives for interpreting and comprehending multimodal texts, I have drawn from these theorists to create a composite framework addressing three distinguishable, if not theoretically distinct, analytical perspectives for interpreting multimodal texts.

### Perceptual Analytical Perspectives

The perceptual analytical perspective focuses on the literal or denotative contents of an image or series of images in a multimodal text, the elements of design, for example borders and font, and other visual and textual elements of these texts. Focusing on visual images, Monroe Beardsley (1981) suggested, “a picture is two things at once: it is a design, and it is a picture *of* (italics in original) something. In other words, it presents something to the eye for direct inspection, and it represents something that exists, or might exist outside the picture frame” (p. 267).

What is presented to the eye for “direct inspection,” or close attention to the literal aspects of an image, naming the visual elements of a multimodal text, and taking an inventory of its contents is the focus of the perceptual analytical perspective. Readers cannot interpret what has not been noticed or recognized. The



first step in expanding readers' interpretive repertoires is by calling attention to the elements and designs of multimodal texts that may at times be overlooked. What is being suggested here is the possibility of a denotative or pre-iconographical enumeration of the content of an image (van Straten, 1994). In other words, readers create an inventory of the literal elements of an image or series of images in a picturebook and use this inventory as the starting point for one's interpretive processes. As Scholes (1985) suggested, noticing visual elements and interpreting what is noticed are sufficiently distinguishable processes for us to discuss and teach them separately. In Stafford's (2008) terms, readers of multimodal texts need to increase their *visual competence* before they are able to become *visually literate*.

Jerome Stolnitz (1960) defined aesthetics as a particular sort of perception, one that values the experience of viewing in and of itself. He separated perception from interpretation in the sense that attention is selective, dictated by purpose and prior experiences. Perception is seen as a "goal-directed behavior" that *prepares* (italics in original) the viewer to respond and interpret (Stolnitz, 1960, p. 30). It is not interpretation in and of itself. "The visual field is created by light falling on our eyes; the visual world, however, interprets these patterns of light as reality. The visual world, then, is an interpretation of reality but not reality itself" (Seward Barry, 1997, p. 15).

Our perceptual system simultaneously *limits* and *calls attention* to what we are able to perceive and understand. Elements are rendered in visual fields by artists and graphic designers and light is registered on the eyes (retinas) of the reader-viewer. However, perception, unlike the registering of light or the apprehension of sense data, is not innocent or naïve. It is guided by the experiences and knowledge of the individual receiving the information. We attend to what we notice, and what we notice depends on what we understand. Readers cannot interpret that which is not perceived, and what is perceived can change based on what is understood.

The construction of meaning or interpretation of multimodal texts is closely connected to the perception of the visual and textual elements inherent in these texts. Because of this, it is important to help young readers notice, attend to, and describe as many of the design elements and images included in multimodal texts as possible. Developing an expanded vocabulary for discussing paratextual elements (Genette, 1997), basic elements of design (Dondis, 1973), visual displays (Moss, 2003), and fonts and other graphic design elements (Samara, 2007) provides the foundation for developing more sophisticated interpretations.

### *Piggybook* from a Perceptual Analytical Perspective

Since *Piggybook* is not paginated, the book will be described as a series of *openings*, meaning a two-page display in the story sequence. Approaching the front cover of *Piggybook*, one is struck by the unusual title of the book and the image of what seems to be a traditional, nuclear family. The borders of the image are broken by the heads of the two boys sticking out the top of the image. The orientation of the book is considered vertical or portrait, which aligns with the family portrait displayed on the cover. The colors of the image are bright, especially the color of the red jackets the boys are wearing which seem to stand out. Two boys and an older man, we

assume the father, are being carried piggyback by a woman, we assume the mother. The two boys and the man are smiling, while the woman is not. The border of the cover image is green and black, and the color of the book is a light pink.

The second opening of the book presents the two boys and their father calling for their breakfast from a rectangular wooden dining table. The father is hidden behind a newspaper, while the boys' mouths are opened wide as they call to their mom for their breakfast. The circle of their open mouths is repeated in the round dishes and cups on the table, and in numerous images contained in the newspaper the father is holding. The father and the two boys are asking the mother to "hurry up" so they won't be late for their important jobs and school, respectively.

The third opening of the book presents four sepia-toned images of a faceless mother doing the dishes, vacuuming, making the beds and leaving for work, arranged in four symmetrical quadrants. The mother is presented in a monochromatic color scheme in comparison to the brightly colored clothes worn by her sons and husband on other pages of the book. We do not see her face as she goes about her work. Three of the images are set inside the house, with the fourth outside on a sidewalk. In the fourth image, a small pig face is presented as graffiti on the brick wall behind the mother. She is dressed in a coat, standing next to a flag pole, as she looks in her purse for something.

The next section discussed is the eighth opening of the book. On the left side or recto, we see a mantle and fireplace with a copy of the painting *Mr. and Mrs. Andrews* by Thomas Gainsborough above the fireplace surround. The female character is missing from the painting, with a white outline left in her place. Pigs are depicted in the mosaic tiles in the fireplace surround, fireplace tools, baseboard, and a vase and photograph on the mantle. The right side or verso of the opening reveals a hoof-like hand in a suit jacket holding an unsigned, handwritten note against a pig-patterned wallpaper stating, "You are pigs." Since the text states Mr. Piggott opens the letter, we can only assume the note is from the missing mother and the hoof is the father's transformed hand.

In the eleventh opening, the sons and father have completed their metamorphoses into human-like figures with pig heads. The wallpaper, furniture covers, cans of food and newspaper all resemble or contain pig faces. The painting of the *Laughing Cavalier* by Frans Hals has been transformed into a pig-headed cavalier. The male characters are shown "rooting around for scraps" and are presented crouched on all fours roaming the floor. The text explains that one evening, as they were scrounging for something to eat behind the chairs in the living room, the mother returned. In the verso of the opening, the mother's shadow is projected casting a blue shadow on the wall, framed by the doorway in which she stands.

This taxonomy, or inventory of the objects contained in the picturebook *Piggybook*, is used as the foundation for the analysis and interpretations to come. The classification of the paintings included in some of the images required research beyond the covers of the text, while the rest of the descriptions focused on literal presentations and depictions of the visual elements of the picturebook. The cover of the book and the four openings described will be analyzed further in the next two sections. A complete analysis of every image in the picturebook would require more space than publication limits allow.

## Structural Analytical Perspectives

Structural approaches to interpreting visual images come from a wide variety of disciplines and fields of inquiry. In the field of visual communication, Radan Martineo and Andrew Salway (2005) offer a generalized system of image-text relations which applies to different genres of multimodal discourse. From the perspective of literacy education, attention is being paid to the concept of multimodality and image-text relations, suggesting a “semiotic turn” in approaches to using multiple sign systems and transmediation in the elementary classroom (Siegel, 2006). Neil Cohn (2007) offers a visual lexicon for reading and interpreting graphic novels and comics, addressing various levels of representation in visual language compared with the structural makeup of verbal language.

William Moebius (1986) and Perry Nodelman (1984) detail the various codes and conventions of contemporary and classic picturebooks. Moebius describes the various pictorial and written elements depicted, arguing that picturebooks obey certain conventions of *recognizability* and *continuity*. Maria Nikolajeva and Carole Scott (2000) present a taxonomy outlining the various word-image interactions in picturebooks. Lewis (2001), Unsworth and Wheeler (2002), and Serafini (2009) draw upon Kress and van Leeuwen's (1996) framework of visual grammar to interrogate the images and elements of visual design contained in contemporary picturebooks and multimodal texts. In addition, Golden and Gerber (1990), Patricia Crawford (2000) and Lawrence Sipe (1998) offer approaches to understanding picturebooks from a semiotic perspective.

Drawing upon Halliday's (1975, 1978) notion of metafunctions in systemic functional linguistics, Kress and van Leeuwen set forth a grammar of visual design to be used to understand how visual images are produced, depicted, and interpreted. Halliday proposed three “metafunctions” for linguistic systems, namely: (1) ideational, (2) interpersonal, and (3) textual. The ideational metafunction focuses on the content or knowledge of the world represented in language. The interpersonal metafunction focuses on the relationships constructed by participants through language. Finally, the textual metafunction deals with the ways texts are structured and composed, in particular issues of grammar.

Kress and van Leeuwen (1996) proffered the terms *representational*, *interactive* and *compositional* to refer to Halliday's three metafunctions for visual images and multimodal texts, respectively. Using these metafunctions as an organizing device, Kress and van Leeuwen (1996) describe a taxonomy of structures within each metafunction that can be employed to understand visual images. Narrative representations, framing, information zones, composition, color and position are just a few of the structures that can be used to interpret visual images and multimodal texts.

The literal “naming” (Serafini and Ladd, 2008) of components of a visual image is a basic type of interpretive move. To simply name or give language to an aspect of a visual image is an act of transduction—a shift from one mode (image) to another (language). Every act of transduction is an act of interpretation, no matter how literal the emphasis. However, the potential of structural and semiotic resources to serve as prompts goes beyond naming or literal meaning. The visual cues

presented in multimodal texts can be interpreted in numerous ways by drawing on the various grammars inherent in images and design elements. Images and design elements can be interpreted because meaning is intended by its creator. Various semiotic resources are selected to serve the interests of the designer or rhetor (Kress, 2010) whether deliberately intended or not.

Semiotic resources, for example color, photography, and painting not only mean different things, they mean things differently (Kress, 2003). Through socio-cultural uses and associations, metaphorical connections and synecdochal relationships these resources provide prompts from which viewers can construct meanings and interpretations. Each semiotic resource brings meaning to light through its materiality. Meaning is made and messages are conveyed through the particular affordances and limitations each resource sets forth. Visual images rely on spatial and compositional layouts, where written and oral language and music are displayed in temporal sequence.

There are numerous taxonomies and structural inventories one can draw upon when interpreting visual images and multimodal texts. Understanding the relationships among various visual structures or grammars and the meanings associated with them in a given culture is an important aspect of the structural analytical perspective. Readers of multimodal texts need to develop a metalanguage for noticing, discussing and interpreting visual images if they are going to move beyond the literal perception of images and multimodal texts.

### *Piggybook* from a Structural Analytical Perspective

Approaching the cover of *Piggybook* from a structural perspective, it would be important to address each metafunction; ideational, interpersonal, and compositional. From the perspective of the ideational metafunction, one would consider the way that characters are represented and the interactions among them. On the cover, the mother is carrying the father and two boys. The arms of the father and two boys encircle one another, while the mother's arms support the father and boys by their legs.

From the perspective of the interpersonal metafunction, the characters are positioned in a middle range of social distance, not too close and not too far from the viewer. It is like a full body portrait, where the viewer is positioned at eye-level. The characters are all looking directly at the viewer, in a frontal orientation, demanding that we interact with them. The modality of the artwork is realistic, yet not as realistic as a photograph.

From the perspective of the compositional metafunction, the characters are framed by a green and black border, but the boys heads are breaking the border, suggesting a more intimate connection between characters and the viewer. The characters are positioned in the center of the image, increasing their importance or salience. The characters are moving from left to right, suggesting a move from where they have traditionally been to where they are heading in the future.

Looking at the second opening, the boys are positioned on the left and right sides of the image, with the father positioned in the center attracting more attention even though he is hidden behind the newspaper. The boys faces are looking up and off

page suggesting they are talking to the mother in another part of the house. The viewer is brought in closer to the participants in this image. White space frames the entire image with no background provided.

In the third opening, the mother's face is drawn from the side and from behind. We are not given access to her face, suggesting she is less important or an anonymous member of the family. The sepia toned images suggest a traditional orientation, a connection to an earlier part of the twentieth century when women had less social standing. The mother is clothed in monotone colors and from a greater social distance than the father and boys in the previous image.

The female character in the painting by Thomas Gainsborough entitled *Mr. and Mrs. Andrews* on the recto (left side) of the eighth opening is missing. The male character, Mr. Andrews, is presented with a pig head, and is staring wide-eyed at the viewer, suggesting he is surprised by his wife's disappearance. The wallpaper and fireplace surround contain images of pigs and pig faces suggesting a metamorphosis is taking place. On the verso (right side) we see a pig hoof in a suit coat holding a note that reads, "You are pigs." The pig faces that adorn the wallpaper are staring at the viewer with a circle shape to the mouth suggesting surprise or alarm.

In the eleventh opening, the male figures, having completed their metamorphoses into pigs and are positioned on the floor on all fours drawn from behind with their rears pointing at the viewer. The mother enters the room and is framed by the doorway in which she stands. We see her shadow from behind as she enters the room. The viewer sees what she sees. The shadow she casts across the door is "Madonna-like" suggesting her return to save her family. The pig faces on the wall and in the wallpaper have down-turned mouths suggesting concern and disappointment.

Each of the interpretations drawn from a structural analysis moves from the literal naming of the perceptual perspective to a proposition of meaning constructed by the viewer. The images serve as prompts containing socially recognized conventions and schema (Albers, 2007; Kress, 2010) for the viewer to draw upon when interpreting the meaning of the images and design of the book. There is no single, direct connection between these images and their meaning. The reader-viewer generates meanings based on her or his previous experiences, culture, and knowledge of social and image conventions. Color, composition, the use of borders, book orientation, negative or white space, salience and modality all bring different meaning potentials that can be drawn upon when interpreting multimodal texts.

### **Ideological Analytical Perspectives**

In similar fashion to structural analytical perspectives, ideological or critical approaches to visual images and multimodal texts are constructed from a variety of disciplines and fields of inquiry, including visual discourse analysis (Albers, 2007), critical content analysis (Beach et al., 2009), critical media studies (Semali, 2003), visual communication (Messaris, 2003), advertising (Williamson, 1978) and cultural studies (Lister and Wells, 2001). Each of these approaches focuses on the socio-cultural, historical, and political contexts of the production, and dissemination of visual images and multimodal texts.

“To explore the meaning of images is to recognize that they are produced within dynamics of social power and ideology. Ideologies are systems of belief that exist within all cultures” (Sturken and Cartwright, 2001, p. 21). They continue, “images are an important means through which ideologies are produced and onto which ideologies are projected” (p. 21). Rose (2001) argues, “one of the central aims of ‘the cultural turn’ in the social sciences is to argue that social categories are not natural but instead are constructed. These constructions can take visual form” (pp. 10–11).

To understand the images and design elements presented in multimodal texts requires readers to consider aspects of production and reception, in addition to the aspects of the image and text itself. The capacity of images to affect us as viewers is dependent on the larger cultural meanings they evoke and the social, political and cultural contexts in which they are viewed (Sturken and Cartwright, 2001). Anne Wolcott (1996) argues that readers must look not only at the relationships within a work of art but beyond the work itself to the historical, cultural and social contexts in order to comprehend its meaning.

Albers (2007) views a visual text as, “a structure of messages within which are embedded social conventions and/or perceptions, and which also present the discourse communities to which visual text maker identifies” (p. 84). Albers (2007) approach, known as “visual discourse analysis” goes beyond the perceptual and structural perspectives, and is an attempt to situate the interpretation of visual texts in a socio-cultural and critical framework. “Visual discourse analysis is a general term for an approach to analyzing art as a language and its use. It is concerned with a theory and method of studying the structures and conventions within visual texts, and identifying how certain social activities and social identities get played out in their production” (Albers, 2007, p. 83).

Rebecca Rogers draws upon Norman Fairclough's (2003) concepts of genre, discourse, and style to interpret a contemporary picturebook from a critical perspective, arguing, “the theories and methods of critical discourse analysis can provide insight into not just *what* (italics in original) is written and illustrated but *how* they are written and illustrated” (Beach et al., 2009). Aiello (2006) argues that a structuralist perspective focuses on deconstructing multimodal texts to identify codes and conventions in order to attach the same meanings to the same signs. However, approaching visual images and multimodal texts from an ideological perspective “provides the possibility for renegotiating the meanings inherent in such constructs rather than seeing these as fixed, irrevocable and natural” (Idema, 2001).

The three metafunctions put forth by Kress and van Leeuwen (1996) are not direct relationships between semiotic resources and meaning. Concepts like power, detachment, involvement and interaction are not meanings *hidden* in the images themselves, they are meaning potentials, a field of possible meanings which are activated by the producers and viewers of images (Jewitt and Oyama, 2001). For instance, signs included in advertisements draw upon shared meanings and cultural codes (Williamson, 1978). Understanding these codes requires knowledge of the structural and semiotic resources used to create advertisements and other multimodal texts. Dependent upon readers ability to make references to the products endorsed through advertisements, the messages presented in

advertisements help invest commodities with value, and these values are attached to the images contained in the advertisements.

The shift from a structural perspective on semiotics to social semiotics represents a shift from interpreting images through the recognition of the codes and conventions used in the creation of visual images, and to considering the socio-cultural contexts of production and reception, in addition to the image itself. Linda Scott (1994) argues, “pictures are not merely analogues to visual perception but symbolic artifacts constructed from the conventions of a particular culture” (p. 252). Going beyond the perceptual and structural analytical perspectives to consider the cultural, historical, and political or ideological ramifications of the production and reception of visual images and multimodal texts is an important consideration in understanding multimodal texts in contemporary society.

### *Piggybook* from a Ideological Analytical Perspective

It should first be noted that *Piggybook* is a commercial product intended for use by children, parents, and teachers in school settings and the home for pleasure reading. Browne brings to his picturebook his experiences, perspectives, and intentions in his production of the book. The publisher (Penguin, UK) selects particular manuscripts to produce, and distributes their products through various commercial channels. The context of the production of the images and text, its distribution and how and where it is read or received, are important considerations in addition to the book itself.

Approaching *Piggybook* in and of itself from an ideological analytical perspective, the image on the front cover suggests a feminist perspective might be useful as an analytical frame. Questions, such as, “Why is the mother carrying a grown man and two boys? Is the image some playful reference to the title? What relationships in the family are working here?,” can all be useful for interrogating the status and perceptions of the family members.

In the second opening, the mother is not shown in the image. Are we to assume that she is too busy cooking to be eating? The traditional roles of domestic housewife and working father need to be contested and brought forth for discussion. The family is portrayed as a traditional “nuclear” family, yet this type of family arrangement makes up less and less of the total population of most countries. Is this family arrangement being portrayed as “normal”?

The images in the newspaper are reminiscent of the iconic, expressivist painting by Edvard Munch known as *The Scream*. The characters mouths also seem related to young birds’ mouths opening wide to be fed by their mothers. From a traditional family dynamic, the male characters sit at the table waiting to be served by the female character. Is this relationship being contested?

In the eight opening, the male character in the painting by Thomas Gainsborough has a look of surprise on his pig-like face. The hoof-like hand portrayed on the verso suggests the men have turned to pigs. This is a visual metaphor or direct reference to them being “male chauvinist pigs.” The fact that the mother has left home, abandoning her family and leaving them to fend for themselves, and that the male characters soon slide into domestic disarray, speaks to the roles that they were required to fill by societal norms. The neatness of the house at this point due to the

mother's work habits and domestic routines will soon be a thing of the past as the men try to take care of themselves. Why are men portrayed as "domestically challenged" when in real life many men are fully capable of taking care of a house, cooking and cleaning for themselves and their families? This stereotypical portrayal of gender roles and norms should be interrogated and connected to the lives of the reader.

In one of the most revealing scenes in the book, the eleventh opening shows the mother returning to find her husband and children "rooting" around on the floor, looking for scraps to eat. She is portrayed as a blue, "Madonna-like" shadow, framed by the doorway in which she stands. The father and two boys are positioned below her, squatting on all fours looking for food. The house is completely in shambles, representing the male characters' transformation from human to animal, in both their physical features as well as their actions and speech. Is there a connection between the outline of the mother and numerous Renaissance images of the Virgin Mary? What connections can be generated from this similarity? Has she returned to save her family? Or has she been lurking outside the home to simply watch her husband and sons learn to appreciate what she does for them?

The roles traditionally associated with men and women are brought to light in this book. However, they are neither contested nor completely explained. The men find it useful to help out with various domestic chores, but only because they were forced into doing so by the mother's abandonment. The images of the mother change power positions with the father, at times being placed below, and then above him in various openings. At the end of the story, the wife goes out and fixes the car, a traditionally male role. From an ideological perspective, gender expectations, norms and family roles needs to be an important aspect of any discussions of this picturebook.

### **Pedagogical Implications**

Theories of visual grammar (Kress and van Leeuwen 1996), semiotic and iconographical analysis (Sipe, 1998; Panofsky, 1955), and visual literacy (Elkins, 2008; Messaris, 1994), though replete in the research and theoretical literature, need to be translated into instructional frameworks that address the pedagogical needs of classroom teachers. Making these theories practical is an important step in bringing instruction in visual literacy into elementary, middle and high school classrooms. Hassett and Schieble (2007) argue, "finding space and time for the visual in K-12 literacy instruction is not only possible when new literacies and new texts can be used in the classroom without sacrificing curricular goals, it is also necessary in a world influenced by changing forms of communication, information and mass media" (p. 67).

Moving beyond an enumeration of the literal contents of an image is an important consideration for comprehending multimodal texts (Serafini and Ladd, 2008). In addition, drawing inferences from the images and text, and interpreting multimodal texts in light of one's experiences and socio-cultural contexts is a challenge for many readers. In order to do so, teachers need to expand their theoretical



foundations by reading outside the traditional disciplinary boundaries of literacy education to consider the impact studies in art theory, visual communication, semiotics and multimodality play on the interpretation and comprehension of contemporary texts. In addition, students need to develop a “metalanguage” to describe and interrogate multimodal texts in purposeful ways to understand the structures and elements used to convey meanings (Zammit, 2007).

Expanding readers' interpretive repertoires to consider the perceptual, structural and ideological perspectives may help readers deal with the complexity inherent in multimodal texts. Rather than assume a single aesthetic or reader-response perspective, the three analytical perspectives addressed in this article force readers to consider alternative interpretive moves that involve the depictions of the world presented, the structures and design elements used, the semiotic resources made available in socio-cultural contexts, and the ideological influences that alter one's perceptions and interpretations.

The challenges facing readers and teachers alike demand that we expand the analytical tools we bring to bear in the process of interpreting visual images and multimodal texts. Simply teaching readers to visualize, summarize and predict will force them to only consider the written textual elements of the texts they encounter in today's society. As we move from a typographic era dominated by the printed word to a post-typographic epoch dominated by the visual image and multimodal texts, the analytical tools and interpretive repertoires we draw upon need to expand to support readers in new times.

## Concluding Remarks

“The unsettled status of the field appears to be a productive moment of experimentation, invention, and problem-posing as researchers design analytic approaches that draw on a range of theoretical frameworks relevant to their research interests, purposes and questions” (Siegel and Panofsky, 2009).

Widening the range of theoretical frameworks to draw upon during the reading and interpreting of multimodal texts enriches the literary and visual experience and allows readers to bring multiple perspectives to their interpretive repertoires. Rogers (2009) suggests, “rather than searching for one particular set of methods, our challenge is to continue exploring the range of methodological tools that can be used alongside different critical theories to analyze children's literature” (p. 142). Broadening the methods and perspectives brought to bear on multimodal texts expands the interpretations generated, allowing readers to challenge the messages presented in visual images and multimodal texts.

## References

- Aiello, Giorgia. (2006). Theoretical Advances in Critical Visual Analysis: Perception, Ideology, Mythologies, and Social Semiotics. *Journal of Visual Literacy*, 26(2), 89–102.

- Albers, Peggy. (2007). Visual Discourse Analysis: An Introduction to the Analysis of School-Generated Visual Texts. In D. W. Rowe, R. T. Jiminez, D. L. Compton, D. K. Dickinson, Y. Kim, K. M. Leander & V. J. Risko (Eds.), *56th Yearbook of the National Reading Conference* (pp. 81–95). Oak Creek, WI: NRC.
- Albers, Peggy. (2008). Theorizing Visual Representation in Children's Literature. *Journal of Literacy Research*, 40(2), 163–200.
- Anstey, Michele, & Bull, Geoff. (2006). *Teaching and Learning Multiliteracies: Changing Times, Changing Literacies*. Newark, DE: International Reading Association.
- Arnheim, Rudolf. (1986). *Art and Visual Perception: A Psychology of the Creative Eye*. Berkeley, CA: University of California Press.
- Barthes, Roland. (1977). The Rhetoric of the Image. In R. Barthes (Ed.), *Image, Music, Text* (pp. 32–51). New York: Hill and Wang.
- Beach, Richard, Encisco, Patricia, Harste, Jerome, Jenkins, Christine, Raina, Seemi Aziz, Rogers, Rebecca, Short, Kathy G., Sung, Yoo Kyung, Wilson, Melissa, & Yenika-Agbaw, Vivian. (2009). Exploring the "Critical" in Critical Content Analysis of Children's Literature. In K. M. Leander, D. W. Rowe, D. K. Dickinson, M. K. Hundley, R. T. Jiminez & V. J. Risko (Eds.), *58th Yearbook of the National Reading Conference* (pp. 129–143). Oak Creek, WI: NRC.
- Beardsley, Monroe C. (1981). *Aesthetics: Problems in the Philosophy of Criticism*. Indianapolis, IN: Hackett Publishing Company.
- Berger, John. (1972). *Ways of Seeing*. London: Penguin.
- Browne, Anthony. (1986). *Piggybook*. New York: Alfred A. Knopf.
- Chandler, David. (2007). *Semiotics: The Basics*, 2nd ed. London: Routledge.
- Chauvin, B. A. (2003). Visual or Media Literacy? *Journal of Visual Literacy*, 23(2), 119–128.
- Cohn, Neil. (2007). A Visual Lexicon. *The Public Journal of Semiotics*, 1(1), 35–56.
- Crawford, Patricia A. (2000). Inside the Picture, Outside the Frame: Semiotics and the Reading of Wordless Picture Books. *Journal of Research in Childhood Education*, 15(1), 66–80.
- Culler, Jonathan. (1997). *Literary Theory: A Very Short Introduction*. Oxford: Oxford University Press.
- Dondis, Donis A. (1973). *A Primer of Visual Literacy*. Cambridge, MA: MIT Press.
- Duncum, Paul. (2004). Visual Culture isn't Just Visual: Multiliteracy, Multimodality, and Meaning. *Studies in Art Education*, 45(3), 252–264.
- Elkins, John. (2008). *Visual Literacy*. New York: Routledge.
- Fairclough, Norman. (2003). *Analysing Discourse: Textual Analysis for Social Research*. London: Routledge.
- Fish, Stanley. (1980). *Is There a Text in This Class? The Authority of Interpretive Communities*. Cambridge, MA: Harvard University Press.
- Fleckenstein, Kristie S. (2002). Inviting Imagery into Our Classrooms. In K. S. Fleckenstein, L. T. Calendrillo & D. A. Worley (Eds.), *Language and Images in the Reading-Writing Classroom* (pp. 3–25). Mahwah, NJ: Erlbaum.
- Freebody, Peter. (1992). A Socio-Cultural Approach: Resourcing Four Roles as a Literacy Learner. In A. Watson and A. Beadenhop (Eds.), *Prevention of Reading Failure* (pp. 26–41). Sydney: Ashton-Scholastic.
- Freebody, Peter, & Luke, Allan. (1990). Literacies Programs: Debates and Demands in Cultural Context. *Prospect: Australian Journal of TESOL*, 5(7), 7–16.
- Genette, Gerard. (1997). *Paratexts: Thresholds of Interpretation*. Translated by J. E. Lewin. Cambridge: Cambridge University Press.
- Golden, Joanne M., & Gerber, Annyce. (1990). A Semiotic Perspective of Text: The Picture Story Book Event. *Journal of Reading Behavior*, 22(3), 203–219.
- Gombrich, E. H. (1961). *Art and Illusion: A Study in the Psychology of Pictorial Representation*, 2nd ed. Princeton, NJ: Princeton University Press.
- Goodman, Nelson. (1976). *Languages of Art*. Indianapolis, IN: Hackett Publishing Company.
- Graham, Judith. (1990). *Pictures on the Page*. Exeter: Short Run Press.
- Halliday, Michael A. K. (1975). *Learning How to Mean: Explorations in the Development of Language*. London: Elsevier.
- Halliday, Michael A. K. (1978). *Language as Social Semiotic: The Social Interpretation of Language and Meaning*. London: Edward Arnold.
- Harvey, Stephanie, & Goudvis, Anne. (2000). *Strategies that Work: Teaching Comprehension to Enhance Understanding*. Portland, ME: Stenhouse.

- Hassenmuller, Christine. (1987). Panofsky, Iconography, and Semiotics. *Journal of Aesthetics and Art Criticism*, 35, 289–301.
- Hassett, Dawnene D., & Schieble, Melissa B. (2007). Finding Space and Time for the Visual in K-12 Literacy Instruction. *English Journal*, 97(1), 62–68.
- Idema, Rick. (2001). Analysing Film and Television: A Social Semiotic Account of Hospital: An Unhealthy Business. In T. van Leeuwen & C. Jewitt (Eds.), *Handbook of Visual Analysis* (pp. 183–204). London: Sage.
- Jewitt, Carey, & Kress, Gunther (Eds.). (2003). *Multimodal Literacy*. New York: Peter Lang.
- Jewitt, Carey, & Oyama, Rumiko. (2001). Visual Meaning: A Social Semiotic Approach. In T. van Leeuwen & C. Jewitt (Eds.), *Handbook of Visual Analysis* (pp. 134–156). London: Sage.
- Kress, Gunther. (2003). *Literacy in the New Media Age*. London: Routledge.
- Kress, Gunther. (2010). *Multimodality: A Social Semiotic Approach to Contemporary Communication*. London: Routledge.
- Kress, Gunther, & van Leeuwen, Theo. (1996). *Reading Images: The Grammar of Visual Design*. London: Routledge Falmer.
- Leeds-Hurwitz, Wendy. (1993). *Semiotics and Communication: Signs, Codes, Cultures*. Hillsdale, NJ: Laurence Erlbaum Associates.
- Lewis, David. (2001). *Reading Contemporary Picturebooks: Picturing Text*. London: Routledge Falmer.
- Lister, Martin, & Wells, Liz. (2001). Seeing Beyond Belief: Cultural Studies as an Approach to Analyzing the Visual. In T. van Leeuwen & C. Jewitt (Eds.), *Handbook of Visual Analysis* (pp. 61–91). London: Sage.
- Martineo, Radan, & Salway, Andrew. (2005). A System for Image-Text Relations in New (and Old) Media. *Visual Communication*, 4(3), 337–371.
- Messaris, Paul. (1994). *Visual Literacy: Image, Mind, & Reality*. Boulder, CO: Westview Press.
- Messaris, Paul. (2003). Visual Communication: Theory and Research. *Journal of Communication*, 53(3), 551–556.
- Mirzeoff, Nicholas. (1998). What is Visual Culture? In N. Mirzeoff (Ed.), *The Visual Culture Reader*. London: Routledge.
- Moebuis, William. (1986). Introduction to Picturebook Codes. *Word & Image*, 2(2), 141–158.
- Moss, Barbara. (2003). *Exploring the Literature of Fact: Children's Nonfiction Trade Books in the Elementary Classroom*. New York: Guilford Press.
- Moss, Pamela A., Phillips, Denis C., Erickson, Frederick, Floden, Robert E., Lather, Patti A., & Barbara Schneider, L. (2009). Learning from Our Differences: A Dialogue Across Perspectives on Quality in Educational Research. *Educational Researcher*, 38(7), 501–517.
- Nikolajeva, Maria, & Scott, Carole. (2000). The Dynamics of Picturebook Communication. *Children's Literature in Education*, 31(4), 225–239.
- Nodelman, Perry. (1984). How Picture Books Work. In H. Darling & P. Neumeier (Eds.), *Image and Maker: An Annual Dedication to the Consideration of Book Illustration* (pp. 1–13). La Jolla, CA: Green Tiger Press.
- Owocki, Gretchen. (2003). *Comprehension: Strategic Instruction for K-3 Students*. Portsmouth, NH: Heinemann.
- Panofsky, Erwin. (1955). *Meaning in the Visual Arts*. Chicago: University of Chicago Press.
- Rose, Gillian. (2001). *Visual Methodologies*. London: Sage.
- Samara, Timothy. (2007). *Design Elements: A Graphic Style Manual*. Beverly, MA: Rockport.
- Scholes, Robert E. (1985). *Textual Power: Literary Theory and the Teaching of English*. New Haven: Yale University Press.
- Scott, Linda M. (1994). Images in Advertising: The Need for a Theory of Visual Rhetoric. *Journal of Consumer Research*, 21, 252–273.
- Semali, Ladislaus. (2003). Ways with Visual Languages: Making the Case for Critical Media Literacy. *The Clearing House*, 39(3), 271–277.
- Serafini, Frank. (2009). Understanding Visual Images in Picturebooks. In J. Evans (Ed.), *Talking Beyond the Page: Reading and Responding to Contemporary Picturebooks* (pp. 10–25). London: Routledge.
- Serafini, Frank, & Ladd, Sophie M. (2008). The Challenge of Moving Beyond the Literal in Literature Discussions. *Journal of Language and Literacy Education*, 4(2), 6–20.
- Barry, Seward, & Marie, Anne. (1997). *Visual Intelligence: Perception, Image, and Manipulation in Visual Communication*. Albany, NY: State University of New York Press.
- Siegel, Marjorie. (2006). Rereading the Signs: Multimodal Transformations in the Field of Literacy Education. *Language Arts*, 84(1), 65–77.

- Siegel, Marjorie, & Panofsky, Carolyn P. (2009). Designs for Multimodality in Literacy Studies: Explorations in Analysis. In K. M. Leander, D. W. Rowe, D. K. Dickinson, M. K. Hundley, R. T. Jimenez & V. J. Risko (Eds.), *58th Yearbook of the National Reading Conference* (pp. 99–111). Oak Creek, WI: NRC.
- Sipe, Lawrence R. (1998). How Picture Books Work: A Semiotically Framed Theory of Text-Picture Relationships. *Children's Literature in Education*, 29(2), 97–108.
- Smith-Shank, Deborah Lee. (2004). Semiotics and Visual Culture: Sights, Signs, and Significance. In Deborah L. Smith-Shank (Ed.), Reston, VA: National Art Education Association.
- Snow, Catherine and Sweet, Anne. (2003). *Rethinking Reading Comprehension*. New York: Guilford.
- Stafford, Barbara Marie. (2008). The Remaining 10 Percent: The Role of Sensory Knowledge in the Age of the Self-Organizing Brain. In J. Elkins (Ed.), *Visual Literacy* (pp. 31–58). New York: Routledge.
- Stolnitz, Jerome. (1960). *Aesthetics and the Philosophy of Art Criticism*. Boston: Houghton Mifflin Company.
- Sturken, Marita, & Cartwright, Lisa. (2001). *Practices of Looking: An Introduction to Visual Culture*. Oxford: Oxford University Press.
- Unsworth, Lee. (2002). Changing Dimensions of School Literacies. *Australian Journal of Language and Literacy*, 25(1), 62–77.
- Unsworth, Len, & Wheeler, Janet. (2002). Re-Valuing the Role of Images in Reviewing Picture Books. *Reading*, 36(2), 68–74.
- van Leeuwen, Theo, & Carey, Jewitt. (2001). *Handbook of Visual Analysis*. London: Sage.
- van Straten, Roelof. (1994). *An Introduction to Iconography*. London: Taylor & Francis.
- Williamson, Judith. (1978). *Decoding Advertisements: Ideology and Meaning in Advertising*. London: Marion Boyars.
- Wolcott, Anne. (1996). Is What You See What You Get? A Postmodern Approach to Understanding Works of Art. *Studies in Art Education*, 37(2), 69–79.
- Zammit, Katina. (2007). Popular Culture in the Classroom: Interpreting and Creating Multimodal Texts. In A. McCabe, M. O'Donnell & R. Whittaker (Eds.), *Advances in Language and Education* (pp. 60–76). New York: Continuum.
- Zeki, Semir. (1999). Art and the Brain. *Journal of Consciousness Studies*, 6(6–7), 76–96.