

## SUMMARY

- ◆ Presents a framework for analyzing imagery in multi-modal print documents and Web sites
- ◆ Demonstrates how images and text work together to make meaning for readers/users
- ◆ Provides analytical tools and tips to help choose still images to enhance textual messages

# Visual Social Semiotics: Understanding How Still Images Make Meaning

CLAIRE HARRISON

**P**rofessional communicators are increasingly involved in developing documents, both print and online, that include still images such as photographs, illustrations, and diagrams. During my 20-year career as a writing consultant to government, I have seen the ratio of visuals to text per page in public documents such as brochures, pamphlets, and reports increase steadily. In the case of online documents, rapid advances in browsers and other technologies have transformed the Internet, formerly a text-only medium, into the World Wide Web, a highly visual publishing environment in which Tyner finds an historic echo: "The way that pictures and texts work together in multimedia interfaces is reminiscent of the visually stunning illustrations of Biblical texts seen in the illuminated manuscripts of medieval times" (1998 p. 40).

It is well beyond the scope of this article to explore the historical, social, political, and technological reasons behind the re-emergence of the visual as an important mode of communication within written documents. The important fact for professional communicators is that readers/users no longer rely solely on written text for comprehension; they absorb and process all that they see within a document to create meaning for themselves. Horn calls this multi-modal mix *visual language*:

*... the tight coupling of words, images, and shapes into a unified communication unit. "Tight coupling" means that you cannot remove the words or the images or the shapes from a piece of visual language without destroying or radically diminishing the meaning a reader can obtain from it. (1999, p. 27)*

However, as writers, we are trained and practised in the use of words. We create documents—reports, online

help, promotional materials—that are designed to be rhetorical in the classical Aristotelian sense, that is "the faculty of observing in any given case the available means of persuasion" (1954/2001). In other words, we manipulate words (our available means) to persuade readers/users that our particular message is credible, meaningful, and useful in a world flooded by a torrent of messages. Many of us, however, are not trained and practised in the use of images for rhetorical purposes.

Do we understand sufficiently how visuals persuade readers/users about messages? Unless we are also trained as artists or art historians, how can we acquire this knowledge? The world of still images and their analysis is vast. As O'Toole notes,

*Books are written about individual works, groups of works, the artist's whole oeuvre, schools, movements, centuries of art. A great number of journals are devoted to nothing else. . . . Television films, lectures and tape-slide shows attempt to enrich our knowledge and perceptions. Art is taught in school, in art college, in university. Art is discussed in committees and boardrooms. All these involve verbal discourses about art and about individual works, so people are finding words to talk about art. However, these are diverse and competing discourses, with their own historical, biographical, economic or technical preoccupations and they don't, for the most part, help us as we stand in front of the art work, lost for words. (1994, p. 4)*

Yet, professional communicators cannot afford to be "lost for words," because many of us work on production teams

Manuscript received 22 July 2002; revised 14 September 2002; accepted 26 September 2002.

along with managers, designers, programmers, illustrators, photographers, and so on.

I have, for example, often sat in meetings (as the writer) where the designer passed around stock photos that could potentially be used on a Web site. As a team, we had to decide which image(s) would best represent the intent of the Web site and contribute rhetorically to the text. Like my colleagues, I generally relied on my “gut” feeling about an image to make a judgment and struggled, as they did, to explain my likes and dislikes to help the team reach a consensus.

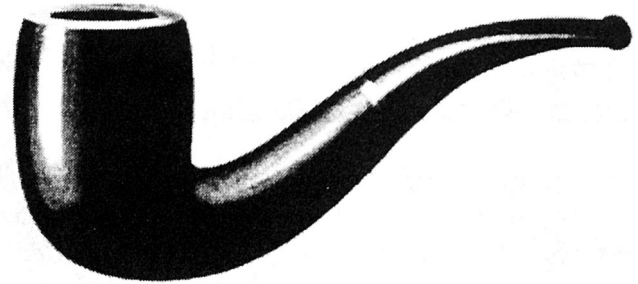
What I needed was an informed vocabulary that would enable me to articulate my reactions when visuals were being selected, tested, and evaluated during print production phases such as design and layout, and Web production phases such as development of the graphical user interface (GUI), asset creation, and prototype/beta testing and evaluation.

It became clear to me that those of us who communicate primarily as writers face three significant challenges in this new multi-modal communication environment. To ensure that our documents are most effective for readers/users, we must

1. Understand how text and still images work together to make meaning together for readers/users.
2. Know when still images enhance or detract from text, and vice versa.
3. Be able to effectively discuss the issues of multi-modal communications with other members of the document's production team.

The purpose of this article, then, is to present a framework—visual social semiotics—that can help professional communicators who need practical tools for image analysis and who may not have the time or inclination to immerse themselves in a new field of study. Although visual social semiotics is not the only theoretical framework for examining how images convey meaning (others include Gestalt theory, art history, psychoanalytical image analysis, and iconography, to name a few), it is unique in stressing that an image is not the result of a singular, isolated, creative activity, but is itself a social process. As such, its meaning is a negotiation between the producer and the viewer, reflecting their individual social/cultural/political beliefs, values, and attitudes.

Because professional communicators must persuade readers/users of one particular viewpoint among many competing and conflicting perspectives, I believe they will find that visual social semiotics can help them better understand the rhetorical, meaning-making potential of still images in relationship to text, provide them with techniques to analyze such images, and contribute to their ability to effectively discuss imagery within a team setting.



*Ceci n'est pas une pipe.*

**Figure 1.** René Magritte's painting makes a striking and explicit statement about the relationship between content (the *signified*) and its visual representation (the *sign*).

## WHAT IS VISUAL SOCIAL SEMIOTICS?

### Semiotics

Semiotics is generally described as the “study of signs.” For a sign to exist, there must be meaning or content (the *signified*) manifested through some form of expression or representation (the *sign*). Figure 1 is a well-known painting by René Magritte that demonstrates this relationship in a striking and explicit manner. By putting the sentence *This is not a pipe* below a highly realistic representation of a pipe, Magritte reminds viewers that the image is not reality but artifice—in other words, a representation or sign.

A thoughtful viewer might note that the word “pipe” itself is an arbitrary combination of four letters that conveys the concept of “pipe” through the form of written expression—once again, a sign.

Signs exist within semiotic systems. For example, the green light in a traffic signal is a sign meaning “go” within the semiotic system of traffic control; words are signs in the semiotic system of language; gestures are signs within the semiotic system of non-verbal communication; and so on. Because semiotic systems encompass the entire range of human practices,

*Semiotics provides us with a potentially unifying conceptual framework and a set of methods and terms for use across the full range of signifying practices, which include gesture, posture, dress, writing, speech, photography, film, television, and radio. . . . As David Sless notes, “we consult linguists to find out about language, art historians or critics to find out about paintings, and anthropologists to find out how people in different societies signal to each other through gesture, dress or decoration. But if we want to know what all these different things have in common then we need to find someone*

*with a semiotic point of view, a vantage point from which to survey our world.” (Chandler 2001)*

It is this cross-cutting vantage point that allows professional communicators to compare and contrast objects from two different semiotic systems—language and imagery—and make a valid, useful analysis.

### Social semiotics

Social semiotics is a branch of the field of semiotics. Lemke notes that

*social semiotics is a synthesis of several modern approaches to the study of social meaning and social action. One of them, obviously is semiotics itself: the study of our social resources for communicating meanings. . . . Formal semiotics is mainly interested in the systematic study of the systems of signs themselves. Social semiotics includes formal semiotics and goes on to ask how people use signs to construct the life of a community. (1990, p. 183)*

Because every community is different, the signs used by one community may be different from those used by another. For example, the color red indicates mourning for people in Ivory Coast, whereas, in contrast, it represents procreation and life for people in India.

Social semioticians apply three important principles when analyzing a semiotic system such as language or imagery—principles that have significance for professional communicators.

**1. Semioticians believe all people see the world through signs.** As Chandler explains,

*Although things may exist independently of signs we know them only through the mediation of signs. We see only what our sign systems allow us to see. . . . Semioticians argue that signs are related to the signifieds by social conventions which we learn. We become so used to such conventions in our use of various media that they seem “natural,” and it can be difficult for us to realize the conventional nature of such relationships. (2001)*

Schriver suggests that successful professional communicators use intuition to “imagine the audience and draw on their internal representation of the audience as a guide to writing . . .” (1997, p. 156). I would add that this intuitive “internal representation” includes a highly sensitized understanding of the sign conventions in a communicator’s particular language semiotic system. This sensitivity contributes to the skills that enable writers to replicate their communities’ discourse in ways that attract interest or please readers.

**2. The meaning of signs is created by people and does not exist separately from them and the life of their social/cultural community.** Therefore, signs have different meanings in different social and cultural contexts—meanings that can range from very different (for example, different languages) to subtle and nuanced (for example, spoken English in U.S. versus spoken English in India). This principle has profound implications for professional communicators who must write for international audiences. The growing number of books and articles on this subject attests to the difficulties writers face when trying to create messages for people whose semiotic systems are different from theirs.

**3. Semiotic systems provide people with a variety of resources for making meaning.** Therefore, when they make a choice to use one sign, they are not using another. As Lemke adds,

*These are the contexts of “what might have been” . . . In the same sentence, what other words could have been used? At the same point in the game, what other plays might have been made? For the same detail in the painting, what other colors could have been used? (1990, p. 188)*

The ability to choose gives communicators a certain amount of power to use signs in unconventional ways and, therefore, affect and even alter meanings.

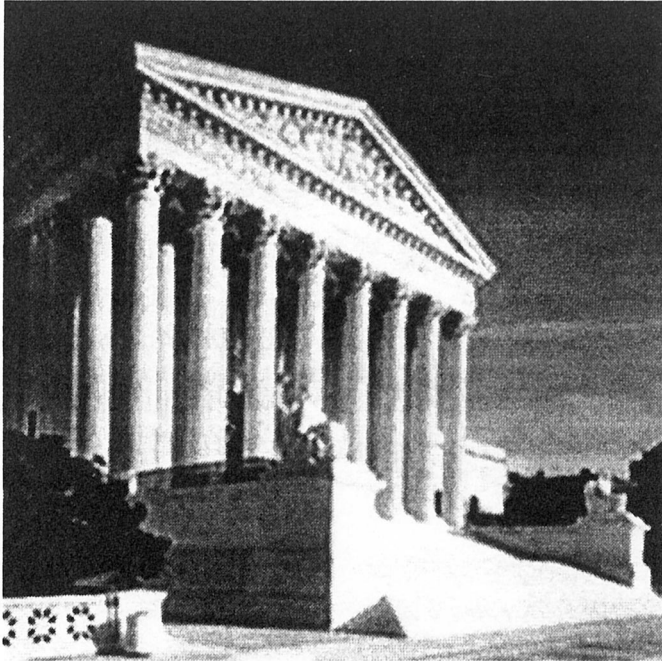
### Visual social semiotics

Visual social semiotics is a new field of study (originating in the 1990s) and has been defined by Jewitt and Oyama as involving “the description of semiotic resources, what can be said and done with images (and other visual means of communication) and how the things people say and do with images can be interpreted” (2001, p. 136).

Here is an example of how visual social semiotics can be used as a tool in analyzing an image on a Web site to see if it enhances, or detracts from, the text. Figure 2 is a photograph from the Web site home page for the Supreme Court of the U.S.

The visual social semiotician would note a significant aspect of this photograph—its point of view. The photograph is taken from an ant’s-eye perspective, placing the Court building at a high vertical angle from the viewer. This angle allows the photographer to glorify the Court by emphasizing the grandeur of its architecture and its classical elegance. The perspective elongates the columns and makes the portico more imposing. Moreover, the high vertical angle compels the viewer to look up at the building—a statement about the pre-eminent power of the Court.





**Figure 2.** Glorifying the Supreme Court through a high vertical angle (<http://www.supremecourtus.gov>).

It's important to remember that the producer of this Web page had other perspective choices that could have altered this "power statement" about the Court. For example, we could have seen the building from a bird's-eye viewpoint, making it smaller, putting it within the context of its surroundings, and reminding us that the U.S. government is "of the people, by the people, and for the people." Or we could have had a photograph of the building taken on its steps so that we would be looking horizontally toward the interior—a perspective that would imply equality between citizens and the judicial system.

Linguistic social semioticians would look at the text that, in addition to the title "Supreme Court of the U.S.," includes a list of links such as Oral Arguments, Case Handling Guides, Court Rules, and Opinions. They would note that the document really contains ellipsis—a form of cohesive textual tie that acts by assuming that the reader is able to presuppose meaning despite omitted words—in this case, verbs and adjectives. The full thought behind the title and links is: "The Supreme Court of the U.S. *bears legal* Oral Arguments, *sets court* Case Handling Guides, *establishes* Court Rules, and *gives legal* Opinions." The linguistic social semiotician would note that this full thought makes heavy use of nominalizations—*arguments*, *guides*, *rules*, and *opinions*. Nominalization, the process of creating a noun from a verb,

- ◆ Deletes the people who do things and the people to whom things are done
- ◆ Eliminates tense (that is, past, present, and future)
- ◆ Omits modal verbs (for instance, *can*, *might*, *should*)

When these elements are cut out from a text, the process (for example, *to judge*) becomes depersonalized and is rendered as an object (for example, *a judgment*). Such textual entities are not only impersonal, they also take on the quality of being timeless and fixed, even though arguments can change and judgments have been overturned. On this Web page, the nominalizations contribute substantial weight to the importance and influence of what nine people do for a living.

Let us imagine the team putting this Web page together. They have already created the text for the title and links and are now choosing a photograph. Say they have a choice of three photographs: the ant's-eye view, a bird's-eye view, and a "level" view. Which would provide them with the best "tight coupling" to make the page's visual language most effective? For example, how "supreme" would the Court appear if the bird's-eye view photograph was used? Or how timeless and fixed would the Arguments, Guides, Rules, and Opinions seem if the perspective of equality were used on the page? Clearly, the team would choose the ant's-eye perspective to enhance the importance of the highest court of the U.S. judicial system, as already implied by the text.

Note that I have not discussed the size or placement of the photograph on the Web page. Although such features—as well as the relationship of the image to other elements of the page, both printed and online—can have a significant impact on visual language, they are part of page design, an aspect of visual communication not covered in this article. To help readers focus on the content of images analyzed here, most of the images I use in this article are excerpted from Web pages.

Professional communicators should be aware that visual social semiotics may not be able to answer all the issues that an image may raise. For example, 10 different creators could be given the same image to produce and would do so in 10 different ways, creating different effects through aesthetics and style—effects that are not addressed through this particular framework. Nor can visual social semiotics help in understanding the rhetorical role that text itself plays as a graphical image through color, typography, and placement. (See Schriver 1997 for a discussion of how people "see" text.) However, as I hope the discussion of the photograph of the U.S. Supreme Court demonstrates, visual social semiotics can be an extremely useful tool for analyzing images and their relationship to text.

### THE THREE CATEGORIES OF IMAGES

The first step in understanding visual social semiotics is knowledge of the three categories of images: the *icon*,



**Figure 3.** When a house represents the home page, it is an iconic image.

*index*, and *symbol*. (The following explanation is adapted from Hammerich and Harrison 2002, pp. 140–142.)

- ◆ **Icon** An image is iconic if it bears a similarity or resemblance to what we already know or conceive about an object or person. Icons include paintings, maps, and photographs and can range from very realistic to very simplistic. On the Web, a common icon is the house used to represent the home page, as shown in Figure 3.
- ◆ **Index** An image is an index if it is recognizable, not because of any similarity to an object or person, but because we understand the relationship between the image and the concept that it stands for. A weather-vane, for example, has no resemblance to any aspect of weather, yet it stands for the concept of wind. A common Web index is the upward-pointing arrow to indicate the top of a Web page as shown in Figure 4. Indexical signs can be confusing and, therefore, often require text to accompany them. This is why the upward-pointing arrow on a scrollable page is often also labeled, “Top of Page.”
- ◆ **Symbol** An image is a symbol when it has no visual or conceptual connection to an object or person. We know the meaning of the image only because of convention; that is, it’s something we’ve learned. A word, for example, is a symbol because it does not resemble what it stands for, nor does it have any indexical relationship to what it signifies. Take the word *rose*. It doesn’t look like a rose or bear any relationship to the concept of a rose. A common Web symbol is the line beneath a word or phrase used to indicate a link, as shown in Figure 5.

Although indexical and symbolic images play important roles in a Web site’s visual language, I will focus primarily on icons in this paper because professional communicators are likely to find that these images are the most contentious during team discussions.

### ANALYZING ICONIC IMAGES

In their seminal work, *Reading images: The grammar of visual design*, Kress and van Leeuwen (1996) set out the first social semiotic framework for analyzing images, noting that “We intend to provide inventories of the major compositional structures which have become established as conventions in the course of the history of visual semiotics,



**Figure 4.** When the upward-point arrow represents the top of the Web page, it is an indexical image.

**This is a link.**

**Figure 5.** When a line appears beneath a word or phrase on the Web, it is a symbolic image.

and to analyze how they are used to produce meaning by contemporary image-makers” (p. 1). Kress and van Leeuwen’s book was a first step in establishing visual social semiotics for imagery in Western culture, and analysts in this field are now adding to the framework through applied research and additional theory. (See Kress and van Leeuwen 2001 for the development of an overarching social semiotic theory for all types of multi-modal communications.)

Essentially, the Kress and van Leeuwen framework recognizes that an image performs, simultaneously, three kinds of meta-semiotic tasks to create meaning. These tasks are called the *representational* metafunction, *interpersonal* metafunction, and *compositional* metafunction.

This section of the article addresses the major elements of each metafunction, briefly analyzes different Web site images according to each metafunction, and provides a list of useful questions for basic analysis of the metafunction. At the end of the article, one of the images is used to demonstrate how the three metafunctions work together to create visual meaning for viewers.

I have included a variety of types and styles of images to help professional communicators working in different fields who wish to undertake visual analyses. It is important to note that these analyses apply visual social semiotic theory from the perspective of Western culture in general and North American culture in particular. As such, it may not be relevant to other traditions that have developed different conventions of imagery and reading.

### The representational metafunction

The representational metafunction is about the people, places, and objects within an image—the represented participants (RPs)—and answers the question “What is the picture about?” Table 1 outlines basic structures and processes of this metafunction.

**An action image** According to Kress and van Leeuwen, “When participants are connected by a vector, they are

**TABLE 1: BASIC STRUCTURES AND PROCESSES OF THE REPRESENTATIONAL METAFUNCTION**

Structures	Processes
<p><b>Narrative:</b> Narrative images allow viewers to create a story about the RPs because the images include vectors of motion.</p> <p><b>Conceptual:</b> Conceptual images do not include vectors. Rather, RPs tend to be grouped together to present viewers with the “concept” of who or what they represent.</p>	<ul style="list-style-type: none"> <li>◆ <b>Action:</b> The narrative is created by vectors that can be bodies, limbs, tools, weapons, roads, and so forth.</li> <li>◆ <b>Reactional:</b> The narrative is created by eyelines (acting as vectors) between RPs.</li> <li>◆ <b>Classificatory:</b> RPs as “kind of” something or some group (that is, they are members of the same class). Advertisements for beauty products often have classificatory images such as a group of models (for instance, Revlon models).</li> <li>◆ <b>Analytical:</b> RPs are displayed in terms of a “part-whole” structure. The “whole” is a Carrier who possesses “parts” called Attributes. The Supreme Court building in Figure 2 is a Carrier, and its architectural components are its Attributes. A pie chart is an analytical image in which the chart is the Carrier and its segments are Attributes. Diagrams are also analytical processes.</li> <li>◆ <b>Symbolic:</b> RPs are important for what they “mean.” A motorbike in an advertisement can, for example, be analytical (that is, asking the viewer to check out its attributes), but it is also symbolic of virility. Abstract shapes such as triangles, squares, and circles also fall in this category.</li> </ul>

represented as *doing* something to or for each other. . . . These vectors are formed by depicted elements that form an oblique line, often a quite strong diagonal line . . . ” (1996, pp. 56–57).

The photograph in Figure 6 demonstrates the effective use of vectors to create action and, therefore, narration. It is the main image on the home page for the American Podiatric Medical Association, and the text following is the first paragraph appearing beneath it.

#### Paying Careful Attention to Details

*The foot is a complex structure made up of 26 bones, thirty-three joints, 107 ligaments, and 19 muscles and tendons. They deserve the focused attention and care that the members of the American Podiatric Medical Association (APMA) can provide. For feet of all ages, from the young foot to older feet the members of the APMA provide high quality healthcare. Details are carefully attended to in meeting the requirements of the healthcare needs of your feet. Good foot health is vital to your overall health.*

Several vectors appear in this image. The two strongest are the diagonal lines created by the adult’s hand and the



**Figure 6.** Strong vectors create a meaningful narrative (<http://www.apma.org/>).

newborn’s leg. Both vectors begin at a corner of the photograph and meet approximately in the middle, creating a powerful interaction between the major RPs—the foot and hand. The smaller vectors created by the toes and fingers provide additional action elements so that the picture has a naturally dynamic rather than static feel. Given the text, viewers have no problem understanding the narrative: that the newborn is a patient and that the hand belongs to a caring podiatrist who pays careful attention to detail, no



matter how tiny.

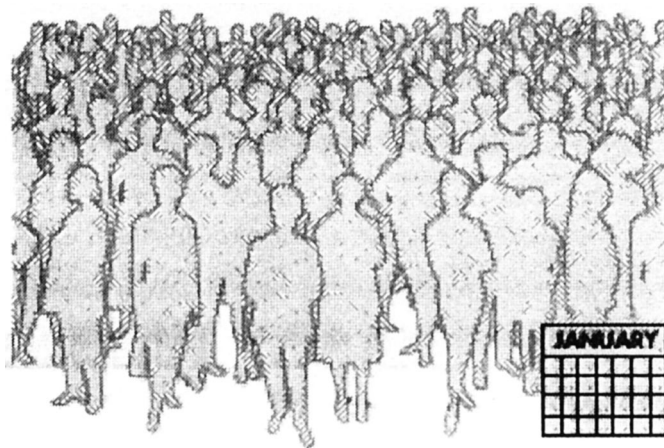
The representational metafunction can include embedding (that is, containing more than one process) in the same way that a sentence can have a main clause and embedded relative clauses such as “The man *who wore the hat* was walking with an umbrella.” In Figure 6, the out-of-focus circle in the background—the hand entrance to an incubator—is a conceptual structure, that is, a symbolic process, embedded within the action process. An incubator is found in hospitals and is, therefore, symbolic of medical intervention to promote human health.

This embedded image supports the text statement that “For feet of all ages, from the young foot to older feet the members of the APMA provide high quality healthcare.” Further, Kress and van Leeuwen point out that “Circles and curved forms generally are elements we associate with an organic and natural order. . . . The world of organic nature is not of our making, and will always retain an element of mystery” (1996, p. 53). Note how the text demonstrates practitioners’ specialized knowledge of the “organic and natural order” by unveiling some of its mystery: “The foot is a complex structure made up of 26 bones, thirty-three joints, 107 ligaments, and 19 muscles and tendons.”

**A conceptual image.** Conceptual images do not involve action or reaction on the part of RPs but represent “participants in terms of their more generalized and more/less stable and timeless essence, in terms of class, or structure, or meaning” (Kress and van Leeuwen 1996, p. 79).

Figure 7 is the main illustration from the home page of The Breast Cancer Site and is accompanied by the text: “43,000 will die this year.” The use of a faceless crowd, a *classificatory* process, encourages the viewer to think of breast cancer in the abstract, that is, about the many who will die and the social, political, and economic consequences of the disease. The figures, which are only outlines, are similar to those seen in graphs, giving a technical/scientific resonance to the image. (It is also interesting that this facelessness of the figures allows the viewer to insert his/her face, thereby adding a personal, and frightening, resonance.) The calendar in the right-hand corner is an indexical sign, serving as a type of “death-watch countdown.” The rectangular shapes of both the image and this sign add to the scientific/technical quality of this image since “In our society, squares and rectangles are the elements of the mechanical, technological order . . .” (Kress and van Leeuwen 1996, p. 52).

Although visual social semiotics assisted in the analysis of Figure 7, the theoretical framework is not sufficient to deal with the many issues involved in the design of appropriate and accurate charts, diagrams, and so on. Professional communicators who must deal with technical illus-



**Figure 7.** This “slice” of a faceless crowd encourages the viewer to consider breast cancer in the abstract (<http://www.thebreastcancersite.com/cgi-bin/WebObjects/CTDSites>).

trations in their projects would be well served by becoming acquainted with the extensive literature on the subject. (Kostelnick 1998 provides a good introduction to the different standards for data displays as well as identifying leaders in the field of technical illustration.)

#### Useful questions for analysis of the representational metafunction

1. Who are the represented participants (RPs) in the image? Include both human and non-human objects.
2. Are there any vectors in the image that indicate action? If so, what kind of story does this action tell?
3. Are the human RPs looking at each other, creating eyeline vectors? If so, what does this tell me about the history of these people?
4. If there are no vectors, what is the image trying to tell me in terms of social/cultural concepts? What types of conventional thinking do different objects evoke in me?
5. Is the image a complex one with more than one process embedded within it? If so, how do these embedded processes add to my overall understanding of the image?
6. In terms of the overall document/Web site, does the choice of image RPs best enhance its intent and that of the text?

#### The interpersonal metafunction

This metafunction is about the actions among all the participants involved in the production and viewing of an

**TABLE 2: BASIC FEATURES AND PROCESSES OF THE INTERPERSONAL METAFUNCTION**

Features	Feature Processes
<b>Image Act and Gaze:</b> The image act involves the eyeline of the RP(s) in relation to the viewer.	<ul style="list-style-type: none"> <li>◆ Demand: The RP is looking directly at the viewer. A demand generally causes the viewer to feel a strong engagement with the RP.</li> <li>◆ Offer: The RP is looking outside the picture or at someone or something within the image. In this case, the RP becomes an object of contemplation for the viewer, creating less engagement than that of the <i>demand</i>.</li> </ul>
<b>Social Distance and Intimacy:</b> Social distance is determined by how close RPs in an image appear to the viewer, thereby resulting in feelings of intimacy or distance.	<p>The viewer can see an RP in six different ways.</p> <ul style="list-style-type: none"> <li>◆ Intimate distance: The head and face only</li> <li>◆ Close personal distance: The head and shoulders</li> <li>◆ Far personal distance: From the waist up</li> <li>◆ Close social distance: The whole figure</li> <li>◆ Far social distance: The whole figure with space around it</li> <li>◆ Public distance: Torsos of several people</li> </ul>
<b>Perspective—The Horizontal Angle and Involvement:</b> This angle refers to the relationship between the position of the RP(s) and the viewer.	<ul style="list-style-type: none"> <li>◆ The frontal angle: When an RP is presented frontally to the viewer. This angle creates stronger involvement on the part of the viewer as it implies that the RP is “one of us.”</li> <li>◆ The oblique angle: When an RP is presented obliquely to the viewer. This angle creates greater detachment since it implies that the RP is “one of them.”</li> </ul>
<b>Perspective—The Vertical Angle and Power:</b> There are two possible vertical-angle relationships: 1) that of the RP(s) and the viewer, and 2) that between RPs within an image.	<ul style="list-style-type: none"> <li>◆ High angle: The RP “looking up” has less power.</li> <li>◆ Medium angle: The RP “looking horizontally” has equal power.</li> <li>◆ Low angle: The RP “looking up” has less power.</li> </ul>

image (that is, the creator, the RPs, and the viewer), and answers the question “How does the picture engage the viewer?” Table 2 outlines basic features of this metafunction.

**Creating strong viewer involvement.** Figure 8 is the only image on the home page of the Stand for Children Web site and appears below the following text:

*Stand for Children is America's only nationwide grass-roots voice for children. Our members take action at the national, state, community, and neighborhood levels.*

*From early childhood education and the schools to after-school programs and health services, we take on the issues affecting our kids.*

The human face is one of the most powerful resources in visual imagery because people are “hard-wired” from infancy to study faces and their expressions. However, faces can be presented in many different ways. This photograph effectively combines four aspects of the interpersonal metafunction to engender strong viewer involvement with the image: visual *demand*, intimate distance, frontal angle, and medium vertical angle.





**Figure 8.** This head shot creates strong viewer engagement (<http://www.stand.org>).

- ◆ **Visual demand** Kress and van Leeuwen say that a *demand* has two functions: “In the first place it creates a visual form of direct address. It acknowledges the viewers explicitly, addressing them with a visual ‘you’. . . . [Also,] the participant’s gaze (and the gesture, if it is present) demands something from the viewer, demands that the viewer enter into some kind of imaginary relation with him or her” (1996, p. 122). The child’s direct look and lack of smile strongly beseech the viewer to acknowledge the needs of children. That the child is female and non-Caucasian is interesting, suggesting that the producers wished to add gender and racial meanings not included in the text.
- ◆ **Intimate distance** Spatial distances are related to emotions of intimacy and distance. Kress and van Leeuwen explain that “The relation between the human participants represented in images and the viewer is once again an imaginary relation. People are portrayed as *though* they are friends, or as *though* they are strangers” (1996, p. 133). In Figure 8, the producers

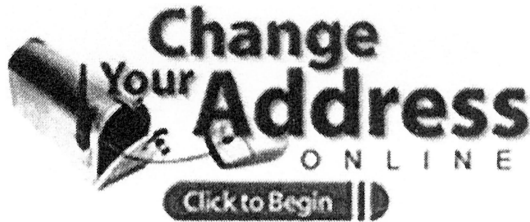
have chosen to bring the child as close to the viewer as possible (even cutting off the top of her head) to create a strong sense of viewer affinity.

- ◆ **Frontal and medium vertical angles** Kress and van Leeuwen note that “The horizontal angle encodes whether or not the image-producer (and hence, willy-nilly, the viewer) is ‘involved’ with the represented participants or not. The frontal angle says, as it were: ‘what you see here is part of our world, something we are involved with’ (1996, p. 143). Simultaneously, the medium vertical angle between the viewer and child indicates equality. The combination of the two angles promotes intense involvement with the child and, by association, with American children. The angles suggest not only that children are equal to adults but that their issues and problems are equally as important.

When analyzing any image, it is important to remember that the producers had other semiotic choices. For example, the producers of this site could have chosen to photograph the child in profile (at an oblique angle presenting a visual *offer*), or with a group of children (public distance), or placed at a low angle to the viewer (less equality). It is questionable whether any of these choices would have resulted in an image as rhetorically supportive of the accompanying text.

**An interesting visual trend** As discussed in Table 2, the vertical angle is associated with power relationships. I have noted an interesting trend in imagery when producers wish to encourage readers/users to incorporate new technologies into everyday practices. The images representing the use of the technology are frequently at a low visual angle to the viewer. In semiotic terms, this means that the viewer holds the power over the technology, not only in terms of choosing to use the technology, but also of being able to understand, control, and manage it. For example, Figure 9 from the U.S. Postal Service encourages users to go online to change their addresses, whereas Figure 10 tells us to “Eliminate check writing and pay anyone anytime online.”

Figure 11, from a site promoting telemedicine, is interesting because it demonstrates two different vertical angles and power relationships. The first is the user-over-technology relationship because the perspective of the photograph ensures that viewers look down on the scene itself. The second is implicit in the vertical angle between the doctor (higher; more powerful) and the patient (lower; less powerful). This angle plus the embedded symbolic representation of a figure in a white coat and stethoscope is designed to reinforce beliefs about the important knowledge possessed by healthcare providers and their resulting high power status in our society.



**Figure 9.** The low visual angle is intended to encourage users to go online (<http://www.moversguide.com/>).



**Figure 10.** The low angle suggests that online banking can be easily learned and managed (<http://moneycentral.msn.com/banking/home.asp>).

### Useful questions for analysis of the interpersonal metafunction

1. Does the image include human RPs? If so, what type of image act is taking place, a *demand* or an *offer*? Theoretically, an object can create a *demand*—for example, a car placed so that its headlights appear to be looking at us. However, *demands* and *offers* seem most powerful when they involve an actual human face.
2. If the image act is a *demand*, how does it affect me? And is it accompanied by any gestures or expressions that make it more forcible?
3. If the image act is an *offer*, why has the producer of the image chosen to make the RP an object of study?
4. How close do I feel to the RPs in the image? Does the closeness make me feel as if the RPs are friends or strangers? In either case, why has the producer of the image chosen to evoke these feelings within me?
5. What do I notice about the perspective in the image? What horizontal and vertical angles have been used?
6. How does the horizontal angle affect my sense of involvement with the RPs?
7. How does the vertical angle add to my knowledge of power relations between myself and the RP and between the RPs themselves?
8. What other semiotic resources could the producer have used to create a different impression?



**Figure 11.** The low visual angle of the total image suggests that patients should not be overawed or intimidated by new medical technologies, while the vertical angle between physician and patient reinforces the status of the medical profession (<http://www.mgh.org/telehealth/vision.html>).

### The compositional metafunction

The compositional metafunction answers the question “How do the representational and interpersonal metafunctions relate to each other and integrate into a meaningful whole?” Composition in imagery is the equivalent of syntax in language—a set of rules that enable the signs of language (that is, words) to be arranged grammatically so that they make sense to the reader.

In other words, the composition of an image or the lay-out of a print or Web page represents visual syntax. If the signs are not put together in a rule-oriented way, viewers will see a hodge-podge of images rather than a coherent whole. Table 3 outlines the basic systems and elements of this metafunction.

### Revisiting Figure 6

Let us revisit Figure 6 and its text to examine how the three metafunctions work together to make a visual message for the viewer, and why the image and text work so effectively together as a rhetorical unit. On the Web page, the image



### Paying Careful Attention To Details

The foot is a complex structure made up of 26 bones, thirty-three joints, 107 ligaments, and 19 muscles and tendons. They deserve the focused attention and care that the members of the American Podiatric Medical Association (APMA) can provide. For feet of all ages, from the young foot to older feet the members of the APMA provide high quality health care. Details are carefully attended to in meeting the requirements of the health care needs of your feet. Good foot health is vital to your overall health.

**Figure 12.** The image plus text create a strong rhetorical unit.

appeared above the first paragraph in the text, as shown in Figure 12.

I have already discussed the representational metafunction of the photograph, and the way in which the RPs and their vectors create a strong narrative structure (that is, a podiatrist taking good care of a newborn). The interpersonal metafunction for this image is “realized” through the use of intimate distance, the horizontal frontal angle, and the medium vertical angle, which heighten viewers’ sense of personal involvement and identification with the RPs. The compositional metafunction effectively integrates the first two metafunctions through a variety of elements.

- ◆ **Information value** The photograph uses all three types of information value systems to get the rhetorical message across to viewers. Within the image, there is a clear left/right system. The foot on the left represents the given (that is, a patient requiring care) while the hand on the right represents what the Web site wishes to offer as a new idea to users—that podiatry care is a significant aspect of personal healthcare. The image also has a center/margin system with the hand holding the foot as the nucleus of information. And, finally, the image combined with the text creates an ideal/real system. The ideal of the positive, caring doctor-patient relationship is visualized in the photograph above the text while the real—facts about feet and practitioners—is incorporated in the text below the image.

- ◆ **Salience** Size, focus, and foreground/background are the processes of salience that contribute to the rhetoric of the photograph. The hand is the largest RP, emphasizing the importance of the doctor. The vectors are in focus and in the foreground, whereas the circle of the incubator and the rest of the background are out of focus. The viewer is, therefore, compelled to pay attention to the action inherent in central image of the clasp.

- ◆ **Framing** Although the photograph has no actual lines as a frame, the use of color (which of course cannot be seen in the black-and-white version of the photo reproduced here) sets it apart from the whiteness of the Web page on which it is placed and frames it as a separate item. Also, the colors within the image fall within the white-cream-brown spectrum of the palette, creating a continuity that holds the pictorial elements together. The text sits directly below the photograph with the slogan “Paying Careful Attention to Details,” acting as the photograph’s caption and the lead-in to the paragraph itself.

- ◆ **Modality** Photographs, by the nature of their technology, suggest a reality that is far stronger than that of drawings, illustrations, and paintings. As Shapiro notes, “Of all modes of representation, [photography] is the one most easily assimilated into the discourses of knowledge and truth for it is thought to be an unmediated simulacrum, a copy of what we considered the ‘real’” (1988, p. 124). In fact, we find photographs to be so real that Shapiro believes they pacify us to the point of unquestioning acceptance of societal norms and conventions. In Figure 12, therefore, the producer of the image uses the photograph’s high modality so that the ideal will be depicted, not as “what might be,” but as “what is” (that is, the truth).

In sum, the systems of the compositional metafunction in Figure 12 play a significant role in integrating the two other metafunctions so that the rhetorical messages of the image and text combine almost seamlessly and come across “loud and clear” to the viewer.

### Useful questions for analysis of the compositional metafunction

1. How have the RPs been placed to provide information, and why has the producer of the image chosen this placement?
2. Which RPs are more salient than others, and how does this salience affect the impact and meaning of the image?
3. How are the RPs held together or separated within an image, and why?



**TABLE 3: BASIC SYSTEMS AND ELEMENTS OF THE COMPOSITIONAL METAFUNCTION**

System	Elements
<p><b>Information Value:</b> The placement of RPs allows them to take on different information roles.</p>	<ul style="list-style-type: none"> <li>◆ Left/Right: RPs on the left side of an image have the value of being “given” knowledge while RPs on the right are “new.”               <ul style="list-style-type: none"> <li>◆ Given = familiar, commonsense</li> <li>◆ New = an issue, a problem, a solution</li> </ul>               (Note: This value is based on how we read in Western cultures, that is, from left to right. This does not necessarily apply to cultures in which reading occurs from right to left or in columns.)             </li> <li>◆ Top/Bottom: RPs at the top of an image have the value of being “ideal” while RPs below represent the “real.”               <ul style="list-style-type: none"> <li>◆ Ideal = emotive, imaginary, what might be, often the pictorial elements of an image</li> <li>◆ Real = factual, informative, down to earth, practical, often textual elements in an image</li> </ul> </li> <li>◆ Center/Margin: RPs in the center provide the nucleus of information to which surrounding elements are subservient.</li> </ul>
<p><b>Salience:</b> Salience refers to the ability of an RP to capture the viewer’s attention.</p>	<ul style="list-style-type: none"> <li>◆ Size: The larger the RP, the greater the salience.</li> <li>◆ Sharpness of focus: Out-of-focus RPs have less salience.</li> <li>◆ Tonal contrast: Areas of high tonal contrast have greater salience.</li> <li>◆ Color contrast: Strongly saturated colors have greater salience than “soft” colors.</li> <li>◆ Foreground/Background: An RP in the foreground has greater salience than an RP in the background.</li> </ul>

**(continued)**



TABLE 3: (Continued)

<p><b>Framing:</b> How RPs are framed affects whether they are seen as connected or separate.</p>	<ul style="list-style-type: none"> <li>◆ <b>Framelines:</b> The lines within the image that divide RPs or hold them together.</li> <li>◆ <b>Pictorial framing devices:</b> The stronger the lines around the image, the greater the connection.</li> </ul>
<p><b>Modality:</b>◆ Modality refers to how we feel about the visual message's validity and reliability. Images with higher modality appear more real than those with a lesser modality. However, the "realness" of imagery can be problematic. For example, although Figure 6 has very low modality according to the framework (that is, the figures are not real, but abstracted), the message behind the image may hold great validity for viewers.</p>	<p>Modality markers: The visual cues that indicate "realness" generally run along a spectrum of possibilities.</p> <ul style="list-style-type: none"> <li>◆ Color saturation, differentiation, and modulation:                             <ul style="list-style-type: none"> <li>◆ Full color = high modality</li> <li>◆ Black-and-white = low modality</li> </ul> </li> <li>◆ Contextualization:                             <ul style="list-style-type: none"> <li>◆ Fully conceived background = high modality</li> </ul> </li> </ul>
<p><b>System</b></p>	<p>Elements</p> <ul style="list-style-type: none"> <li>◆ Background completely absent = low modality</li> <li>◆ Depth:                             <ul style="list-style-type: none"> <li>◆ Deep perspective = high modality</li> <li>◆ No perspective = low modality</li> </ul> </li> <li>◆ Illumination:                             <ul style="list-style-type: none"> <li>◆ Fullest representation of light and shade = high modality</li> <li>◆ Absence of light and shade = low modality</li> </ul> </li> </ul> <p>◆ Although Kress and van Leeuwen (1996) do not include modality in any of the metafunctions, I have followed the lead of Jewitt and Oyama (2001, pp. 151–153) by placing it in the compositional metafunction.</p>

4. How does the use of color or lack of it affect the rhetorical message of the image?

5. How real does the image appear to the reader, and does this sense of reality affect the validity of its message and that of the accompanying text?

6. Are there other ways this image could have been organized that would strengthen its message and more effectively enhance its accompanying text?

## FINAL WORDS ABOUT GETTING STARTED

Kress and van Leeuwen state that

*Social semiotics is an attempt to describe and understand how people produce and communicate meaning in specific social settings, be they "micro" settings such as the family or settings in which sign-making is well institutionalized and hemmed in by habits, conventions and rules. But social semiotics, sign-making in society, is so varied an activity that any attempt to capture it in a general theory must look crude by comparison with the richness of the actual semiotic world. (1996, p. 264)*

Given this caveat, I believe that professional communicators will find visual social semiotics an effective tool for

understanding many conventions found in Western imagery that, despite people's differences in age, ethnicity, gender, and so on, evoke generally uniform reactions. However, learning to use the framework *en masse* is a formidable, daunting process because it is complex and introduces a great deal of new terminology. I have found that the best approach is not to attempt to use the whole framework immediately, but to focus on one or two aspects of the framework when starting to analyze images. Here are my suggestions for getting started.

- ◆ Begin with simple images rather than complex ones—for example, the icons chosen for a GUI or head shots that involve only one person. Move to more complex images and then page designs as you develop your skills.
- ◆ Use only one metafunction in early analyses. I suggest the interpersonal metafunction because it is the one whose effect you and your colleagues will feel most immediately on viewing an image.
- ◆ Use only one or two elements of the chosen metafunction in early analyses. Add elements as you gain the confidence to undertake more in-depth study.
- ◆ When analyzing the representational metafunction, dig deep. The rapidity of human visual perception is

such that you understand what an image is about immediately on sight. Chances are you will not be accustomed to taking an image apart to analyze its separate elements.

- ◆ Your visual virtuosity also means that you are likely to understand many of the elements of the compositional metafunction at an intuitive level. Make this knowledge explicit by explaining it to others.
- ◆ Remember that not every aspect of every metafunction is important in analyzing an image for its impact on a reader/user. For example, the most notable feature of the photograph of the U.S. Supreme Court in Figure 2 is the interpersonal aspect of power “realized” through the vertical angle.
- ◆ Continually ask yourself why a producer chose a particular semiotic resource. Just as you instinctively know how to use words to evoke certain feelings in readers/users, producers often rely on instinct in their image choices and may not be aware at a conscious level of the sign conventions they have employed.
- ◆ Reflect on the development of images and design as a project moves through its production phases, noting how colleagues discuss images and how illustrators and designers respond to their suggestions. And if usability testing is undertaken, note how users react to the rhetorical messages embedded in the combination of text and imagery. The more you learn about others’ social/cultural “lenses,” the more sensitive you will become to the different potentials for meaning-making with images.
- ◆ As you gain confidence and can handle a variety of images and page designs, consider ways in which you might, as Dragga and Voss suggest, propose that a team humanize technical illustrations by adding imagery to pie charts, bar graphs, diagrams, and so on, and “genuinely integrating words and pictures instead of simply juxtaposing the two on the page or screen” (2001, p. 270).
- ◆ Practice visual social semiotics to “get the hang” of it. Magazine advertising is a fertile field for analysis as marketers rely on either meeting our expectations regarding sign conventions or breaking them to attract our attention.

Throughout my schooling, I was told that “a little bit of knowledge is a dangerous thing”—the implication being that people making judgments without in-depth knowledge of a particular subject will make mistakes. My experience in visual social semiotics is the reverse. Even a small understanding has been extremely useful in my work. Therefore, I believe that professional communicators will find value in the framework whether they decide to focus on one aspect or delve deeply into theory. **TC**

## REFERENCES

- Aristotle. 1954/2001. “Aristotle’s rhetoric.” W. R. Roberts, trans. <http://www.public.iastate.edu/~honeyl/Rhetoric/index.html>
- Chandler, Daniel. 2001. “Semiotics for beginners.” <http://www.aber.ac.uk/media/Documents/S4B/semiotic.html>
- Dragga, Sam, and Dan Voss. 2001. “Cruel pies: The inhumanity of technical illustrations.” *Technical communication* 48:265–274.
- Hammerich, Irene, and Claire Harrison. 2002. *Developing online content: The principles of writing and editing for the Web*. New York, NY: John Wiley & Sons.
- Horn, Robert. E. 1999. “Information design: Emergence of a new profession.” In *Information design*, ed. R. Jacobson, Cambridge, MA: The MIT Press, pp. 15–33.
- Kostelnick, Charles. 1998. “Conflicting standards for designing data displays: Following, flouting, and reconciling them.” *Technical communication* 45:473–482.
- Kress, Gunther, and Theo van Leeuwen. 2001. *Multimodal discourse: The modes and media of contemporary communication*. London, UK: Arnold.
- Kress, Gunther, and Theo van Leeuwen. 1996. *Reading images: The grammar of visual design*. London, UK: Routledge.
- Jewitt, Carey, and Rumiko Oyama. 2001. “Visual meaning: A social semiotic approach.” In *Handbook of visual analysis*, eds. T. van Leeuwen and C. Jewitt. London, UK: Sage Publications, pp. 134–156.
- Lemke, Jay. 1990. *Talking science: Language, learning and values*. Norwood, NJ: Ablex Publishing Corporation.
- O’Toole, Michael. 1994. *The language of displayed art*. London, UK: Leicester University Press.
- Shapiro, Michael J. 1988. *The politics of representation: Practices in biography, photography, and policy analysis*. Madison, WI: University of Wisconsin Press.
- Schriver, Karen. 1997. *Dynamics in document design*. New York: John Wiley & Sons.

Tyner, Kathleen R. 1998. *Literacy in a digital world: Teaching and learning in the age of information*. Mahway, NJ: Lawrence Erlbaum Associates, Publishers.

**CLAIRE HARRISON**, president of CANDO Career Solutions, Inc., is an award-winning writer and editor with 25 years' experience working for government and corporate clients. She recently

completed an MA with Distinction in applied language studies at Carleton University, Ottawa, ON, specializing in writing, rhetoric, social semiotics, and systemic functional linguistics. She is the co-author with Irene Hammerich of *Developing online content: The principles of writing and editing for the Web* and has published papers on hypertext, professional writing, and Critical Discourse Analysis (CDA). Contact information: can.do@rogers.com