A taxonomy of relationships between images and text

Emily E. Marsh and Marilyn Domas White

*College of Information Studies, University of Maryland,
College Park, Maryland, USA*

Abstract *The paper establishes a taxonomy of image-text relationships that reflects the ways that images and text interact. It is applicable to all subject areas and document types. The taxonomy was developed to answer the research question: how does an illustration relate to the text with which it is associated, or, what are the functions of illustration? Developed in a two-stage process – first, analysis of relevant research in children's literature, dictionary development, education, journalism, and library and information design and, second, subsequent application of the first version of the taxonomy to 954 image-text pairs in 45 Web pages (pages with educational content for children, online newspapers, and retail business pages) – the taxonomy identifies 49 relationships and groups them in three categories according to the closeness of the conceptual relationship between image and text. The paper uses qualitative content analysis to illustrate use of the taxonomy to analyze four image-text pairs in government publications and discusses the implications of the research for information retrieval and document design.*

Documents are complex tools created to communicate messages to readers. Among the elements used in developing documents are prose, images, font types, color, and spatial relationships. The message is a wedding of these components, and the interplay among all elements is a critical concern to people who need to convey information effectively. This paper focuses on an increasingly important aspect of this problem – the relationships between text and images. In writing about the design of Web sites, where images are widely used in conjunction with prose, Karen A. Schriver (1997, pp. 406-7) says:

> We have almost no knowledge about how best to structure the relationships between words and pictures on the Web ... too much time in document design circles is being spent arguing over the details of HTML code ... It's the interplay of prose and graphics that ought to concern us.

The paper creates a taxonomy of relationships between images and text that can be used for analyzing the way that images and text interact. It is applicable to all subject areas and all types of documents. Writers and illustrators can use it as a tool as they create documents, and researchers can analyze the relationships between image and prose to identify and predict the effects of combinations once documents have been published. The taxonomy was

derived to answer the research question: How does an illustration relate to the text with which it is associated, or, alternatively, what are the functions of illustration? The function of an illustration is defined as an image's functional relationship to a relevant text string. It is a conceptual variable that refers directly to the way that ideas expressed within text are conveyed through a relevant illustration and is not related to physical characteristics, style, or file format.

Because several fields claim interest in understanding the relationships between text and images – advertising, education, journalism, and information studies, for example – a significant purpose of this taxonomy is to develop a common language for researchers and practitioners in these fields to use in discussing the relationships.

Literature review

The relevant literature supporting this paper is not only dispersed across fields but also at times hidden within documents that do not focus directly on the relationship between images and text. Some researchers have developed lists of functions ordered by some principle (for example, Hidderly and Rafferty, 1997; Levin, 1981; Levin *et al.*, 1987; Levin and Mayer, 1993; Schwarcz, 1982). In the literature in which the functions are not directly addressed, the functions generally emerge from other objectives, for example, understanding the impact of images on people's retention of text (see Levin *et al.* (1987) for a meta-analysis of studies with this focus). Methodologically, it consists of experimental research (see, for example, David, 1998), literature reviews (see Levin, 1981), content analyses of documents (see Walma van der Molen, 2001), and essays or books based primarily on observation and introspection (see Fang, 1996). The latter predominate.

Table I identifies briefly the research that has identified relationships included in the taxonomy. It groups them by subject field to show the range of fields addressing this issue. Children's literature researchers (Bodmer, 1992; Fang, 1996; Nikolajeva and Scott, 2000; Schwarcz, 1982) are interested in illustrations in non-textbook material written for children, e.g. children's picture books. Dictionary designers focus on using illustrations to convey definitions or usage effectively (see, for example, Landau, 1989; Zgusta, 1989). Educational researchers emphasize the instructional value of images and look at the use of illustrations in instructional texts for all age groups, for example, Woodward's (1993) content analysis of photographs in high-school textbooks. Journalists are interested in newspaper layout and how images affect the retention of information. David (1998), for example, did an experimental study to see how news pictures affected retention of abstract and concrete news. Library and information studies researchers have studied the information value of images in multimedia (Berenstein, 1997); and the "perspective-inducing function" of text illustrations (Peeck, 1993); and how

No.	Researcher[a]	Date	Subject[b]	A. Little relationship	Taxonomy areas B. Close relationship	C. Extend beyond text
1	Bodmer	1992	ChLit	A1	B5, B5.1	C1, C2
2	Fang	1996	ChLit		B1, B2.3, B5.1	C2, C2.2
3	Nikolajeva and Scott	2000	ChLit	A2.1	B1.7, B5.2	C2, C2.2, C3.1
4	Schwarcz	1982	ChLit	A2.1	B4.1	C2, C2.2, C3.1, C.3.2, C.3.3
5	Hancher	1992	Dict		B5, B5.1	
6	Ilson	1987	Dict		B1.7, B3.1, B3.2, B5.1	C3.2, C3.2.1
7	Landau	1989	Dict	A1	B1.6	
8	Zgusta	1989	Dict	A1		
9	Brody	1980	Educ	A1.1, A3, A3.2	B1.3, B1.6, B2, B2.1, B3.1, B4	C1.1, C1.2, C2.1, C3.2, C3.2.1
10	Duchastel	1978	Educ	A1	B5, B5.1	C3
11	Levin	1981	Educ	A1	B1, B1.7	C1, C3
12	Levin and Mayer	1993	Educ		B1.1, B2, B4.1, B4.2, B5.1	C3, C3.2, C3.2.1
13	Levin, Anglin and Carney	1987	Educ			C1.2
14	Woodward	1993	Educ	A1		
15	David	1998	Jour	A1	B1, B1.1, B1.2, B2, B5	C1, C3
16	Kress and van Leeuwen	1998	Jour	A2.1	B1, B5.2	C2, C2.2
17	Walma van der Molen	2001	Jour	A2.1	B1	C2.2
18	Wanta	1988	Jour		B2.4	
19	Berenstein	1997	LID	A1, A2	B1.1, B1.5, B2.3	C1.2
20	Hidderly and Rafferty	1997	LID	A2	B1.1.1, B2, B2.1, B2.2, B2.3, B3	
21	Peeck	1994	LID		B2.4	
22	Schriver	1997	LID		B2.3, B2.4, B5.2	C2, C2.2
23	Stam	1988	LID	A2.2	B1, B1.1, B1.2, B1.4, B1.6, B2.2, B3.1, B5, B5.1	C1.1, C2
24	Tanner and Larson	1994	LID	A3.1	B1.4	

Notes: [a] Complete bibliographic references are given in References. [b] ChLit=children's literature; Dict = dictionary design; Educ = education; Jour = journalism; LID = library or information design

art historians use illustrated texts (Stam, 1988). Information designers have also focused on selecting and using illustrations in technical communication (Tanner and Larson, 1994). At some point, all studies in Table I name or discuss the relationships between images and text; the last three columns indicate how their relationships are incorporated into the taxonomy.

In the following sections, the article first describes the two-stage methodology used to develop a taxonomy of image functions. The second section presents the taxonomy itself and briefly discusses the structure and nature of the terms included. The third section uses qualitative content analysis to demonstrate using the taxonomy to analyze image-text pairs from several different types of documents. The fourth section discusses the implications of the taxonomy for design and retrieval. The final section briefly summarizes the paper and suggests additional research.

Methodology

The taxonomy was developed in two stages. The first stage collocates and identifies, integrates, and organizes concepts describing relationships between images and text already described by other researchers in several different fields. This process is the commonly accepted procedure for classification and thesaurus development (Soergel, 1974). Table I identifies briefly the 24 articles in the literature review. In the second stage, the taxonomy derived from the first stage was applied to a sample of image-text pairs in Web pages to test its sufficiency and appropriateness for identifying relationships in this environment and to find any new relationships that needed to be represented in the taxonomy.

First stage

In the first stage, the following procedures were used to develop the functions:

(1) Identify all functions specified by articles from the literature review.

(2) Collapse syntactic variants expressing closely similar concepts under a single term. For example, describe, describing, and description were combined under the single term, "describe".

(3) Combine semantic variants of the same concept under a single, preferred term. For example, "reinforce", "repeat", "symmetrical", and "transcribe" were grouped under the single term "reiterate".

(4) Group the concepts by the degree of relationship the image has with the text.

(5) Develop scope notes using the original definitions from articles in the literature review.

(6) Use the semantic variants of the same concept (as noted in no. 3) as lead-in terms.

In the fourth step, the concepts were categorized at the highest level into three groups: (A) functions expressing little relation to text; (B) functions expressing a close relation to the text; and (C) functions going beyond the text. This set of higher categories maps the functions according to a principle of semantic organization, an approach used in similar groupings in the literature (see, for example, David, 1998; Hancher, 1992; Hodnett, 1988).

Second stage
The analytical approach used in the second stage of development, qualitative content analysis, is also suggested for applying the taxonomy generally to images and texts (see the "Examples" section) so it is described here in some detail. Unlike quantitative analysis, which emphasizes the objective content of the message, qualitative content analysis focuses on the meaning and other rhetorical elements in the message.

Over the last few years, researchers have acknowledged and demonstrated the value of qualitative analysis (Altheide, 1996; Kellehear and Fook, 1997; Kiser, 1997; Morgan, 1993; Thompson, 1996; Thompson, 1999; Trochim, 1999), yet they have generated only a few precise definitions of the purpose and procedures of the methodology. Glaser and Strauss (1967) point out that "the major emphasis of qualitative content analysis is to capture the meaning, emphasis, and themes of messages". Other authors refer to themes (Trochim, 1999), major ideas (Trochim, 1999), patterns and the position in social content (Thompson, 1996), and the rhetorical structure of texts (Kiser, 1997). In contrast, quantitative content analysis, a more widely used content analysis method, focuses on the objective content of text and collects quantitative data, often to measure the extent or frequency of certain occurrences within text. Usually the data are collected as predefined categories of variables. Studies of TV violence that count acts of violence reflect this approach.

The focal concepts of qualitative analysis are elusive and difficult to measure with precision and rigor. Phillips and Hausbeck (2000, p. 187) argue that qualitative content analysis calls for interpretation "that goes beyond mere counting of occurrences and instead analyzes more subtle aspects of textual construction, layout, and content". The interpretation is based on "careful, analytical, theoretically informed discourse analysis" (PROMISE: Projects for Multicultural and Interdisciplinary Study and Education, n.d.). Miller and Crabtree (1992), in their editing analysis style of qualitative research, describe an essentially inductive process in which the analyst approaches the data, reaches an understanding of its essential parts, and formulates a conception of its meaning in terms of a larger context, perhaps a theory. In their style, which was used in this study, the researcher reads the text closely, looks for, and reorganizes meaningful segments of text to create a summary that "reveals the interpretive truth in the text" (Miller and Crabtree, 1992, p. 21), then categorizes the text segments and determines the patterns and themes that connect them. A researcher can use a

pre-existing schema, such as the taxonomy developed in stage 1, as a template but recognizes that it may have to be adjusted based on the critical analysis. This process is similar to axial coding (Strauss and Corbin, 1990).

In the second stage of development, the taxonomy based on stage 1 was used as a template to code a broad sample of image-text pairs[1]. An image-text pair consists of an image and its related text segment. For this study, images consist only of drawings and photographs relevant to the document's rhetorical goals; omitted are tables because they are text-based, irrelevant images, such as borders, and true-to-life animations because they are highly complex media types with their own internal rhetoric that cannot be conceived as direct representations of textual content (Morrison *et al.*, 2000). Graphics that are really images of words, for example, a menu illustrated by a jpg file, were converted into text segments. Text segments can be any size but, in this context, consist of the string of words directly related to the image as determined by correspondence between the content of words and image, not mere physical proximity.

The image-text pairs in the study represent all image-text pairs ($n = 954$, based on 767 text segments and 363 images) found in 45 Web pages. The Web pages make a quota sample drawn from three categories of 100hot.com (n.d.), a Web catalog that lists the most popular Web sites as determined by its sample base of over 100,000 users. The three categories, which represent genres of Web pages, were: pages with educational content designed for children, online newspapers, and retail business pages. These categories were selected because their parent disciplines, education, journalism, and advertising, are interested in the systematic use of images; thus, the pages are likely to contain diverse image-text relationships rich in meaning. A total of 15 pages were randomly selected from each category. Omitted were any pages that were inactive, subsumed by another Web entity, or inaccessible, composed exclusively of either images or text, or were not written in English.

General rules for applying the taxonomy were to code each image-text pair with as many illustration functions as necessary since a single illustration may have several functions, and to code each function at the most specific level possible, for example, A3.2 instead of A3. The taxonomy was used as a template; it served as a beginning point in the analysis but was modified and expanded as necessary, based on the form of the data and the direction taken by the analytical process. Any new functions were then processed according to the relevant steps in stage 1 to integrate them into the existing taxonomy. Table II thus represents functions identified through both a top-down approach (based on concepts in the relevant research literature) and a bottom-up approach (derived from applying the taxonomy to image-text pairs)[2].

Taxonomy

The taxonomy, including all 49 functions developed in the two stages, is presented briefly in Table II and more completely (with scope notes, lead-in

A Functions expressing little relation to the text	B Functions expressing close relation to the text	C Functions that go beyond the text
A1 Decorate	*B1 Reiterate*	*C1 Interpret*
A1.1 Change pace	B1.1 Concretize	C1.1 Emphasize
A1.2 Match style	B1.1.1 Sample	C1.2 Document
A2 Elicit emotion	B1.1.1.1 Author/Source	*C2 Develop*
A2.1 Alienate	B1.2 Humanize	C2.1 Compare
A2.2 Express poetically	B1.3 Common referent	C2.2 Contrast
A3 Control	B1.4 Describe	*C3 Transform*
A3.1 Engage	B1.5 Graph	C3.1 Alternate progress
A3.2 Motivate	B1.6 Exemplify	C3.2 Model
	B1.7 Translate	C3.2.1 Model cognitive process
	B2 Organize	C3.2.2 Model physical process
	B2.1 Isolate	C3.3 Inspire
	B2.2 Contain	
	B2.3 Locate	
	B2.4 Induce perspective	
	B3 Relate	
	B3.1 Compare	
	B3.2 Contrast	
	B3.3 Parallel	
	B4 Condense	
	B4.1 Concentrate	
	B4.2 Compact	
	B5 Explain	
	B5.1 Define	
	B5.2 Complement	

Table II.
Taxonomy of functions
of images to the text

terms, and sources) in the Appendix. Only four (8.2 percent) of the functions were identified in the analysis of Web pages in atage 2: A1.2 Match style; B1.1.1 Sample; B3.3 Parallel; and C3.2.2 Model physical process.

The completeness of the taxonomy for identifying relationships between text and images can only be assessed indirectly. It is based on analyzing image-text relationships in diverse environments: television news, children's picture books, textbooks, technical communication, newspapers, dictionaries, and multimedia, and Web pages. In stage 2, 30 functions were used to code the image-text pairs in the sample of Web pages. The fact that only four additional concepts were established in the analysis of Web pages argues for the completeness of the taxonomy.

These functions were grouped in two levels, based on semantic analysis. Hancher's (1992) relationship types of dependence, independence, and interdependence and David's (1998) three-pronged notion of decorative, representative, and organizing functions parallel respectively the A, B, and C groups above. Similarly, Walma van der Molen (2001, pp. 485-6) suggests judging correspondence between text and pictures in television news on a "continuum that ranges from … indirect or partial correspondence …

through] direct semantic overlap ... [to] indirect or partial correspondence to divergent or conflicting text-picture relations". His approach, like Kress and van Leeuwen's (1998), provides another way of looking at the functions using a principle of semantic relationship. Hodnett's (1988) survey of book illustration, however, identifies three functional designations: to decorate, to inform, and to interpret, that most closely parallel the three used in the taxonomy.

Examples illustrating application of the taxonomy
The following examples from US government publications demonstrate use of the taxonomy to identify functional relationships between an image and relevant text. These examples are not from the stage 2 analysis but have been selected to show the analytical approach and provide more evidence of a range of functions[3]. US government publications were selected as sources because of the publication variety (in subjects, audiences, and formats) and because they are not usually copyrighted so illustrations from them can be reproduced without permission. Two of the publications are reproduced in entirety on the Web and in print, whereas the others appear only on the Web. The analytical approach used is qualitative content analysis, described in the methodology section.

Examples A and B: war crimes
Figure 1 appears twice in the Web site for the US Rewards for Justice program, sponsored by the US Department of State. Examples A and B refer to its use on the War Crimes page and on the Rwandan Genocide page respectively. These examples show that the same image's functions can change in relation to different text segments. By clicking on "War Crimes in Rwanda" (the hypertext link noted in the text below), the user is connected to a set of options, one of which is Rwandan Genocide (see the second text segment).

Figure 1.
Rwandan war crimes

Source: US Rewards for Justice (n.d.a.)

War Crimes [page title]
The links to the left and below will take you to pages covering various international war
crime incidents where crimes have been committed against humanity.
[heading for war crimes in another country]
War Crimes in Rwanda [link] (US Rewards for Justice, n.d,a).

On the Rwandan Genocide page the photograph is the only illustration for the
second text segment noted below:

Rwandan Genocide [page title]
In, 1994, during a 100-day period approximately 800,000 Rwandan Tutsis and moderate
Hutus were brutally murdered in the Rwandan genocide (US Rewards for Justice, n.d,b).

In the photograph, the orderly pile of bones, sorted by type and stacked, is
disturbing. In both examples, it evokes an emotional response from the user
(A2 Elicit emotion). Coupled with the image of a Rwandan man wearing a
medical mask to avoid contamination, it also concretizes the notion of "war
crimes in Rwanda" (on the War Crimes page) and "Rwandan genocide" (the
second page) by showing the outcome of those crimes in terms of the victims
(B1.1 Concretize). In example B, the text segment is longer and, among other
content, describes the war crime of genocide of the Tutsis and Hutus,
specifying both the length of the massacre and the numbers killed. The picture
relates directly to that description showing the outcome of the massacre and
indicating the number of people killed by the piles of bones and skulls. The
image and text complement each other to convey the horror of the actions (B5.2
Complement). In example B as well, the image intensifies the description, thus
giving a sense of importance to the described event (C1.1 Emphasize).

Example C: radio waves

Figure 2 reproduces an illustration from a chapter on radio waves for grades
8-12 in a teacher's guide from the National Aeronautics and Space Agency. The
text segment related to this illustration is:

The frequency of each of these waves is what determines whether or not it is absorbed or able
to pass through the atmosphere. Low-frequency waves do not travel very far through the
atmosphere and are absorbed rather quickly. Higher frequency waves are able to pass
through the ionosphere and escape into space while the low frequency waves reflect off the
ionosphere and essentially "skip" around the earth. The diagram below will help illustrate
this (US National Aeronautics and Space Agency, 2003, p. 33).

This illustration is a complex image that, among other things, puts the three
types of radio waves into perspective, focusing on their physical properties but
relating them to more concrete, familiar objects in space, e.g. the satellite for the
ionosphere, and to the earth itself. Careful reading of the text and matching its
components to elements within the image shows that the image is very closely
related to the text, so many of its coded functions fall within the B column in
Table II. It translates the written form into a visual form fairly directly (B1.7
Translate). In the process it allows the reader to see the waves in their true
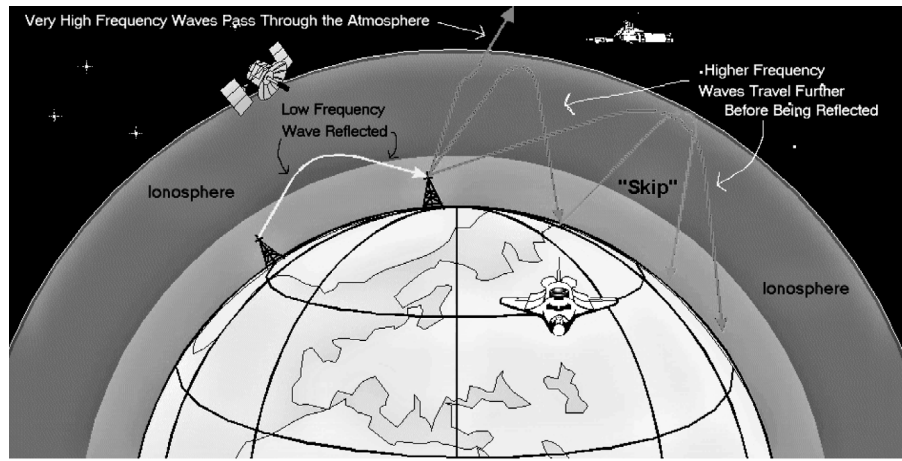
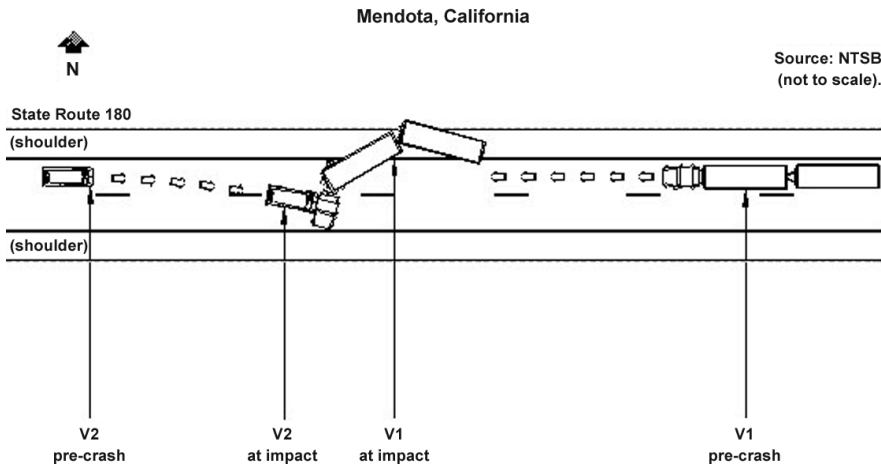**Source:** US National Aeronautics and Space Agency (2003)

relations and thus induces perspective (B2.4 Induce perspective). It relates the waves by explicitly comparing them (B3.1 Compare) and thus makes the text more intelligible but still follows the text segment very closely. The image also shows a relationship that extends beyond the text (see column C). In showing the waves in relation to the earth and to the different flying objects, the image models the waves and their properties (C3.2.2 Model a physical process) helping the reader to see waves that are not themselves directly observable.

Example D: highway accident
Figure 3 diagrams an accident occurring between a van and a semi trailer truck and is part of the documentation produced by the National Transportation Safety Board in its subsequent investigation. The figure, labeled "Accident diagram" is referred to in the text segment, entitled "Description of the Accident":

> A witness behind the van stated that the van's driver drove into the opposing lane of State Route 180 traffic at about 55 mph to pass two cars. (See figure 1.) The truckdriver stated that he saw the van's headlights through the fog traveling toward him in his lane. He said he braked, skidded, and jackknifed into the opposing lane. The van struck the right side of the tractor. Eleven of the 12 van occupants were killed. The right-rear van passenger sustained serious injuries, and the truckdriver sustained minor injuries (US National Transportation Safety Board, 2002, p. 1).

This diagram is an abstract model of the physical actions involved in the collision; it shows both vehicles' positions pre-clash and at impact (see labels within the figure) (C3.2.2). The arrows show the movement of each vehicle and the paths each took immediately prior to the collision. In the case of the truck, the image conveys one impact of the collision on the vehicle (jackknifing). The diagram allows the user to contrast the paths of the two vehicles (B3.2

Mendota, California

Source: NTSB
(not to scale).

State Route 180

(shoulder)

(shoulder)

V2
pre-crash

V2
at impact

V1
at impact

V1
pre-crash

Source: US National Transportation Safety Board (2003)

Figure 3.
Highway accident

Contrast) and concentrates elements in the description, focusing on the most critical information (B4.1 Concentrate). The image omits some elements central to the event, however. It does not portray fog, for example, passing cars as a rationale for the van's being in the wrong lane, and the consequences of the accident on the passengers and drivers. The diagram does not attempt to convey motivations of the drivers, weather, road conditions, and consequences of the accident. As a result, the diagram directs the reader's attention to certain details, thus regulating his attention (A3 Control) and influencing and developing his perspective on the accident (B2.4 Induce perspective).

Example E: congressional district map
Figure 4 appears in the Web site for the *National Atlas*, produced by the US Department of the Interior. It is one of three illustrations on a page indicating three map product lines available for printing at this Web site. Supporting text for this figure is as follows:

> Congressional Districts – 108th Congress.
> The National Atlas prepared simple maps of each and every District of the 108th Congress expressly for use on the World Wide Web. These maps include the latest District boundaries that resulted from the 2000 decennial census ... Each map illustrates the District boundary and includes roads, streams, and cities for easy reference. Our free Congressional District maps are provided in two formats; one that's ideal for viewing on screen and another for generating high quality color prints (US Department of the Interior, National Atlas.gov, n.d.).

This map of the first congressional district in New Mexico engages the reader and directs his attention to the product line (A3.1 Engage), not to the district itself. It constitutes a sample (B1.1.1 Sample) of the congressional district maps described in the text. No reference is made in the text to the specific district; the

Figure 4.
Congressional district
map
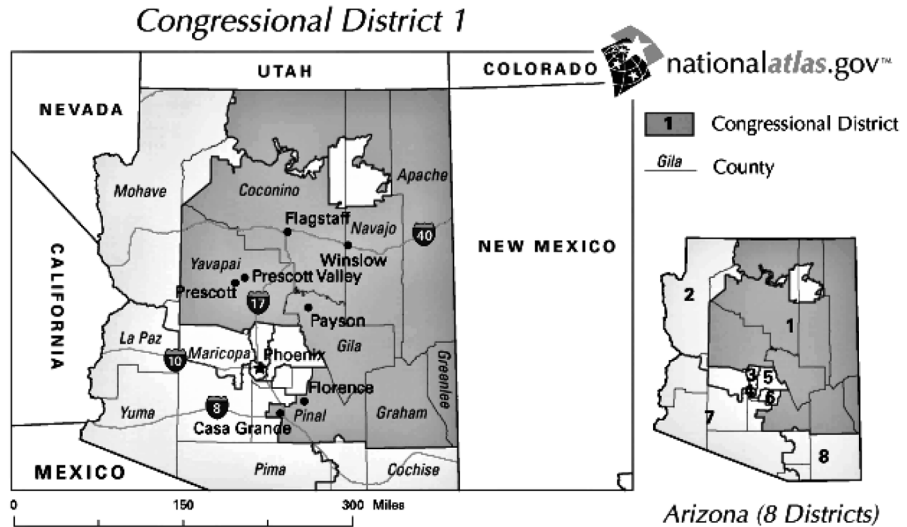
Source: US Department of the Interior National Atlas.gov. (n.d.)

text refers instead to the product line's characteristics, objectives, and content, clearly shown in the sample. Because the map is not an "ideal" map, it is not coded as exemplifying a congressional district map (B1.6 Exemplify).

Example F: Dewie the turtle
Figure 5 reproduces the only illustration at a quiz on computer security and Internet privacy for children at the US Federal Trade Commission's (n.d.) (FTC) Web site. The supporting text for the illustration is minimal: "Are you a safe cyber surfer?"

Figure 5 is an example of personification. In Dewie the turtle, which appears in other poses at this site, the FTC creates a character that attracts children (A3.1 Engage). Dewie helps the agency to personify its message (B1.2 Humanize) to sensitize children to concerns about Internet privacy and security. By showing Dewie surfing on the ocean, the figure uses a parallel action in a different context to relate the idea of surfing on the Internet (B3.3 Parallel).

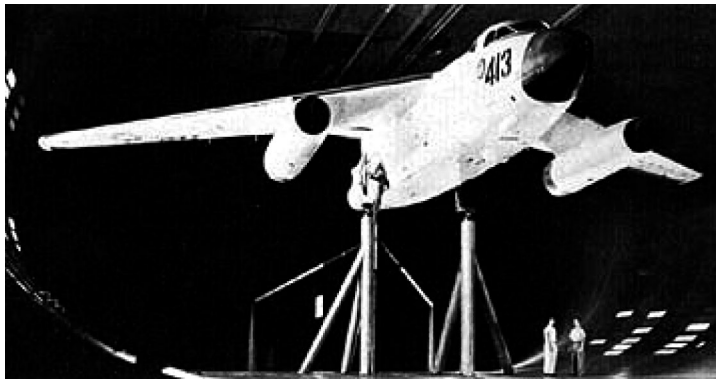*Example G: Timeline: Ames wind tunnel*
The photograph (Figure 6) of an airplane in a wind tunnel is one of eight photographs illustrating a timeline showing accomplishments in aeronautics in the 1940s sponsored or participated in by the US National Aeronautics and Space Administration (n.d.). The complete timeline is extensive (ten pages and 53 illustrations) and covers 1915 through 2003. It is divided, for more recent years, into decades. The following text accompanies this picture:

Figure 5.
Dewie the turtle

Source: US Federal Trade Commission (n.d.)

Figure 6.
Timeline: Ames wind
tunnel

Source: US National Aeronautics and Space Administration (n.d.)

The First Century of Flight: NACA/NASA Contribution to Aeronautics (at top of page).
1940 (along the margin of the page to indicate the division by decades)
1944 – Ames 40 × 80 full-scale wind tunnel became operational. It allowed whole aircraft to
be wind-tunnel-tested, as compared to models at low flight speeds, and expanded testing
capabilities to larger and faster aircraft.

In this timeline, a photograph or drawing that portrays the accomplishment
accompanies each date. The timeline itself (with all dates and photographs)
locates each specific accomplishment in time (B2.3 Locate). In the Ames wind
tunnel example shown in Figure 6, scale, which is emphasized in the text, is
difficult to show in the photograph, which shows a real, entire, large airplane in
a wind tunnel (see the evidence of the building's structure at the top of the

photograph). As a result, the two modes of presentation (text and photograph) complement each other to convey the nature of the accomplishment (B5.2 Complement). The photograph also attracts the reader's attention (A3.1 Engage) and changes pace in the timeline, moving the reader from one accomplishment to another (A1.1 Change pace).

This set of examples has illustrated 17 different functions (sometimes more than once), most of which are closely related to the text (Table II, column B). In three instances, the images went beyond the text to emphasize or model a physical process. The functions, with references to examples in parentheses, are:

- A   Functions expressing little relation to text.
- A1.1   Change pace (example G).
- A2   Elicit emotion (A, B).
- A3   Control (D).
- A3.1   Engage (E, F, G).
- B   Functions closely related to the text.
- B1.1   Concretize (A, B).
- B1.1.1 Sample (E).
- B1.2   Humanize (F).
- B1.7   Translate (C).
- B2.3   Locate (G).
- B2.4   Induce perspective (C, D).
- B3.1   Compare (C).
- B3.2   Contrast (D).
- B3.3   Parallel (F).
- B4.1   Concentrate (D).
- B5.2   Complement (B, G).
- C   Functions that go beyond the text.
- C1.1   Emphasize (B).
- C3.2.2 Model physical process (C, D).

Implications for design and retrieval

That no commonly accepted method exists for describing relationships between images and text is striking, especially considering the length of time

interest has existed and the range of disciplines involved in their study. A. Kibedi Varga (1989, pp. 31-2) sums up the problem, "Facing the tremendous variety of phenomena which belong to word-and-image research, one would like to find some very general categories or headings which would allow a clear and comprehensive classification of all these phenomena". The taxonomy in this paper is a step toward such a classification. The three main headings for the 49 relationship types described by this study introduce a degree of order to a topic that is addressed often but specified infrequently.

Many applications can be developed quickly and with little effort using this conceptual system. The first opportunities arise from the beginning of the creative process for illustrated documents, both in print and electronic form. For writers concerned more with the text end of the image-text relationship, recognizing that images can fulfill several functions in various combinations provides them with conceptual support during the creative process. In addition, the taxonomy allows them to communicate successfully with artists commissioned to illustrate their works. On the other side of this equation stands the illustrator who can use the taxonomy to question clients about their texts. In sum, this image function taxonomy offers new ways to understand and combine images with text both for writers and illustrators.

Another application of the taxonomy is in the development of Web-based electronic documents. During the design process, Web designers create mockups of page layouts to test their appearance and functionality. These diagrams are intended to give a high-level perspective of a given page's form, not actual content, hence their common designation as "wireframes" (Rosenfeld and Morville, 2002). Wireframes contain no content – only placeholders to show the eventual location of document elements such as headlines, graphics, captions, and navigational and search aids. Designers draw, test, and revise these models throughout the design process to communicate effectively with clients. Placeholders indicating the existence of functional relationships can enrich Web document wireframes containing graphics and linked text. These indicators would not contain any specific relationship types, given the lack of actual content, instead only labels from the taxonomy at a generic level (for example, a wireframe can specify that an image-text pair placed in the center of a given home page should always share a close, type B relationship). Once a Web page is ready for actual content and html coding, image functions for particular text strings can be included as metadata tags, similar to those proposed by the Dublin Core Metadata Initiative (2003).

More traditional information storage and retrieval methodologies can also benefit from the taxonomy. For example, a hypertext version of the functions displayed in classified order can be linked to a database of image-text pairs. Such a design can be the architecture for an integrated system whereby users can access examples of image-text pairs with specific relationship functions, either by browsing through the hierarchy and clicking terms, or actual retrieval

using a search feature, or both used in combination. Visual information professionals prefer this type of integrated system using search and display of functions that show images in their original context, that is, linked to text from original documents (Stam, 1984, 1988, 1989). This brief description of a hypertext-based system gives a sense of the potential value a full implementation of the taxonomy can offer the searcher in need of images placed in relation to text. The information-rich displays it would produce would facilitate the in-depth, conceptual, qualitative analysis considered the most helpful for giving users of visual information the kind of value-added presentation that fits their needs for relevance, clarity, and meaning (Bates, 1998; Stam and Giral, 1998).

Finally, a classification of illustration function can provide content analysis researchers with the basis for a codebook for document analysis and writers with a foundation for style guide of image use in the creation of such documents.

Conclusion and suggestions for additional research

This paper has developed a taxonomy of relationships between images and text that can be used for analyzing the way that images and text interact through a two-stage process. The two-stage process involved systematic analysis of relevant research and subsequent application of the preliminary taxonomy to images in Web pages. The taxonomy that resulted from these steps identifies 49 functions that images play in relationship to relevant text. These functions vary in their closeness to the text and are grouped in the taxonomy into three major categories: those with little relationship to the text, those with close relationship, and those whose functions are based on, but transcend, the text. Useful for all subject areas and document types, the taxonomy develops a common, standardized language for expressing relationships between images and text. Because interest in how image and text interact is dispersed across several disciplines, a common language is especially useful. It facilitates and promotes developing a coherent, non-duplicative body of research, theory, and practice.

What additional research needs to be done or can be done with this taxonomy? The first suggestion relates to validating the taxonomy. As developed, the taxonomy relies on researchers looking back at documents and inferring the functions between the images and texts. Complementary research focuses on the design process and asks: To what extent are these functions the actual functions that graphic and Web designers consider in choosing images to illustrate texts? Research is currently in progress (Marsh, in progress) that asks how and why these users select images to use in illustrating text. As part of this research, the user group will react to the taxonomy and judge the accuracy, completeness, and structure of the thesaurus.

Other research relates to use of the taxonomy for retrieval purposes. How can the taxonomy be used to enhance the retrieval of images? The Implications section has already mentioned several ways and others can be developed, but they need to be tested to show their usefulness for the intended user group and their effectiveness in retrieval. Since the functional relationship expressed is between image and text, what level of granularity should be considered for retrieval purposes, for example, just the coded image-text pair or it in conjunction with broader contexts, such as intermediate groupings within the documents or the document itself, for example, a Web page?

In addition, the taxonomy can be used to assess the use of images comparatively across documents, as was done in stage 2, for example, assessing the functions played by illustrations in scientific texts and perhaps relating the functions to the effectiveness of the texts for learning or retention of information. Other research can focus on judging the relative usefulness of different versions of documents with images serving varying functions.

Notes

1. Dagobert Soergel originally suggested image-text pairs as the contextual unit of analysis for applying the taxonomy in a real application.

2. The analytical results that relate to the occurrence of the concepts in Web pages are presented in Marsh (2002) and Marsh and Soergel (in progress).

3. Additional examples of coded image-text pairs from Web pages are included in Marsh (2002). Only one (Marsh, 2002, pp. 106-14) presents the qualitative content analysis that preceded the coding. Additional examples are also available in Marsh and White (2003).

References

100hot.com (n.d.), available at: www.100hot.com (accessed May 27, 2003).

Altheide, D.L. (1996), *Qualitative Media Analysis*, Sage Publications, Thousand Oaks, CA.

Bates, M.J. (1998), "Indexing and access for digital libraries and the Internet: human, database, and domain factors", *Journal of the American Society for Information Science*, Vol. 49 No. 13, pp. 1185-205.

Berenstein, P. (1997), "Moving multimedia: the information value in images", *Searcher*, Vol. 5 No. 8, pp. 40-9.

Bodmer, G.R. (1992), "Approaching the illustrated text", in Sadler, G.E. (Ed.), *Teaching Children's Literature: Issues, Pedagogy, Resources*, Modern Language Association of America, New York, NY, pp. 72-9.

Brody, P.J. (1980), "Research on pictures and instructional text: difficulties and directions", Report IR00-9-028, paper presented at the annual meeting of the Association for Educational Communications and Technology (ERIC Document Reproduction Service ED 196 422).

David, P. (1998), "News concreteness and visual-verbal association: do news pictures narrow the recall gap between concrete and abstract news?", *Human Communication Research*, Vol. 25 No. 2, pp. 180-201.

Dublin Core Metadata Initiative (2003), *DCMI Metadata Terms*, Online Computer Library Center, Inc., Dublin, OH, available at: http://dublincore.org/documents/dcmi-terms (accessed May 28, 2003).

Duchastel, P.C. (1978), "Illustrating instructional texts", *Educational Technology*, Vol. 18 No. 11, pp. 36-9.

Fang, Z. (1996), "Illustrations, text, and the child reader: what are pictures in children's storybooks for?", *Reading Horizons*, Vol. 37 No. 2, pp. 130-42.

Glaser, B.G. and Strauss, A.L. (1967), *The Discovery of Grounded Theory*, Aldine DeGruyter, New York, NY.

Hancher, M. (1992), "Illustrations", *Dictionaries*, Vol. 17 No. 5, pp. 79-115.

Hidderly, R. and Rafferty, P. (1997), "Democratic indexing: an approach to the retrieval of fiction", *Information Services and Use*, Vol. 17 No. 3, pp. 101-9.

Hodnett, E. (1988), *Five Centuries of English Book Illustration*, Scolar Press, Aldershot.

Ilson, R. (1987), "Illustrations in dictionaries", in *The Dictionary and the Language Learner: Papers from the EURALEX Seminar at the University of Leeds, 1-3 April 1985*, M. Niemeyer, Teubingen, pp. 193-212.

Kellehear, A. and Fook, J. (1997), "Lassie come home: a study of 'lost pet notices'", *Omega*, Vol. 34 No. 3, pp. 191-202.

Kiser, E. (1997), "Comment: evaluating qualitative methodologies", *Sociological Methodology*, Vol. 27, pp. 151-8.

Kress, G. and van Leeuwen, T. (1998), "Front pages: the critical analysis of newspaper layout", in Bell, A. and Garrett, P. (Eds), *Approaches to Media Discourse*, Blackwell Publishers, Inc., Malden, MA, pp. 186-282.

Landau, S.I. (1989), *Dictionaries: The Art and Craft of Lexicography*, Cambridge University Press, Cambridge.

Levin, J.R. (1981), "On functions of pictures in prose", in Pirozzolo, F.J. and Wittrick, M.C. (Eds), *Neuropsychological and Cognitive Processes in Reading*, Academic Press, New York, NY, pp. 203-28.

Levin, J.R. and Mayer, R.E. (1993), "Understanding illustrations in text", in Britton, B.K., Woodward, A. and Binkley, M. (Eds), *Learning from Textbooks: Theory and Practice*, Lawrence Erlbaum Associates, Hillsdale, NJ, pp. 95-113.

Levin, J.R., Anglin, G.J. and Carney, R.N. (1987), "On empirically validating functions of pictures in prose", in Willows, D.M. and Houghton, H.A. (Eds), *The Psychology of Illustration*, Volume 1: Basic Research, Springer-Verlag, New York, NY, pp. 51-114.

Marsh, E.E. (2002), "Rhetorical relationships between images and text in web pages", unpublished doctoral dissertation, University of Maryland, College Park, MD.

Marsh, E.E. (in progress), "Selecting and using images to illustrate text: designers".

Marsh, E.E. and Soergel, D. (in progress), "The rhetoric of Web pages".

Marsh, E.E. and White, M.D. (2003), "A taxonomy of relationships between images and text: additional examples illustrating its application", available at: www.glue.umd.edu/~whitemd/taxonomy_ex.html (accessed August, 2003).

Miller, W.L. and Crabtree, B.F. (1992), "Primary care research: a multimethod typology and qualitative road map", in Miller, W.L. and Crabtree, B.F. (Eds), *Doing Qualitative Research*, Sage Publications, Newbury Park, CA, pp. 3-28.

Morgan, D.L. (1993), "Qualitative content analysis: a guide to paths not taken", *Qualitative Health Research*, Vol. 3 No. 1, pp. 112-21.

Morrison, J.B., Tversky, B. and Betrancourt, M. (2000), "Animation: does it facilitate learning?", in Butz, A., Kruger, A. and Oliver, P. (Eds), *Smart Graphics: Papers from the 2000 American Association for Artificial Intelligence Spring Symposium*, Technical Report SS-00-04, AAAI Press, Stanford, CA, pp. 53-60.

Nikolajeva, M. and Scott, C. (2000), "The dynamics of picturebook communication", *Children's Literature in Education*, Vol. 31 No. 4, pp. 225-39.

Peeck, J. (1993), "Increasing picture effects in learning from illustrated text", *Learning and Instruction*, Vol. 3 No. 3, pp. 227-38.

Phillips, J. and Hausbeck, K. (2000), "Just beneath the surface: re-reading geology, re-scripting the knowledge/power nexus", *Women's Studies Quarterly*, Vol. 28 No. 3, pp. 181-202.

PROMISE: Projects for Multicultural and Interdisciplinary Study and Education (n.d.), "Content analysis of introductory geology textbooks", available at: www.scsv.nevada.edu/~promise/content.html (accessed May 29, 2003).

Rosenfeld, L. and Morville, P. (2002), *Information Architecture for the World Wide Web: Designing Large-Scale Web Sites*, 2nd ed., O'Reilly, Cambridge, MA.

Schriver, K.A. (1997), *Dynamics in Document Design*, Wiley Computer Publishing, New York, NY.

Schwarcz, J.H. (1982), *Ways of the Illustrator: Visual Communication in Children's Literature*, American Library Association, Chicago, IL.

Soergel, D. (1974), *Indexing Languages and Thesauri: Construction and Maintenance*, Melville Publishing Company, Los Angeles, CA.

Stam, D.C. (1984), "How art historians look for information", *Art Documentation*, Vol. 3 No. 4, pp. 117-9.

Stam, D.C. (1988), "Art historians and their use of illustrated texts", in *Scholarly Resources in Art History: Issues in Preservation*, Commission on Preservation and Access, Washington, DC, available at: www.clir.org/pubs/reports/cpaarth/cpaarth.html#historians (accessed May 27, 2003).

Stam, D.C. (1989), "Tracking art historians: on information needs and information-seeking behavior", *Art Libraries Journal*, Vol. 14 No. 3, pp. 13-16.

Stam, D.C. and Giral, A. (1998), "Introduction", *Library Trends*, Vol. 37 No. 2, pp. 117-9.

Strauss, A. and Corbin, J. (1990), *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*, Sage Publications, Newbury Park, CA.

Tanner, B. and Larson, P. (1994), "'Worth a thousand words': choosing and using illustrations for technical communication", *Technical Communication*, Vol. 41 No. 1, pp. 150-3.

Thompson, I. (1996), "Competence and critique in technical communication", *Journal of Business and Technical Communication*, Vol. 10 No. 1, pp. 48-81.

Thompson, I. (1999), "Women and feminism in technical communication", *Journal of Business and Technical Communication*, Vol. 13 No. 2, pp. 154-79.

Trochim, W.M.K. (1999), *Research Methods Knowledge Base*, 2nd ed., available at: http://trochim.human.cornell.edu/kb (accessed May 21, 2003; version current as of August 2, 2000).

US Department of the Interior National Atlas.gov (n.d.), *Atlas Printable Maps, Congressional Districts – 108th Congress*, available at: http://nationalatlas.gov/printable.html (accessed August 1, 2003).

US Federal Trade Commission (n.d.), *Are You a Safe Cyber Surfer?*, available at: www.ftc.gov/bcp/conline/edcams/infosecurity/popups/safesurf_quiz.html (accessed August 3, 2003).

US National Aeronautics and Space Administration (n.d.), *The First Century of Flight: NACA/NASA Contribution to Aeronautics, 1940*, available at: http://history.nasa.gov/centtimeline/pages/nine-40.html (accessed August 3, 2003).

US National Aeronautics and Space Agency (2003), "Radio waves (grades 8-12)", in US National Aeronautics and Space Agency (Ed.), *Live from the Aurora, Educators' Guide 2003*, US National Aeronautics and Space Agency, Washington, DC, available at: http://sunearth.gsfc.nasa.gov/sunearthday/2003/educators_guide2003/pdf/rwaves_acts.pdf (accessed May 27, 2003), p. 33.

US National Transportation Safety Board (2002), "Highway accident brief: collision with jackknifed semitrailer, Mendota, California, November 16, 1997", NTSB Report no. HAB-02-11, adopted on 9/17/02, National Transportation Safety Board, Washington, DC, available at: www.ntsb.gov/publictn/2002/HAB0211.htm (accessed May 4, 2003), www.ntsb.gov/publictn/2002/HAB0211.pdf (accessed May 4, 2003).

US Rewards for Justice (n.d,a), *War Crime*, available at: www.rewardsforjustice.net/english/warcrimes/War_Crimes_index.htm (accessed May 4, 2003).

US Rewards for Justice (n.d,b), *Rwanda Genocide*, available at: www.rewardsforjustice.net/english/warcrimes/rwanda_genocide.htm (accessed May 4, 2003).

Varga, A.K. (1989), "Criteria for describing word-and-image relations", *Poetics Today*, Vol. 10 No. 1, pp. 31-53.

Walma van der Molen, J.H. (2001), "Assessing text-picture correspondence in television news: the development of a new coding scheme", *Journal of Broadcasting and Electronic Media*, Vol. 45 No. 3, pp. 483-98.

Wanta, W. (1988), "The effects of dominant photographs: an agenda-setting experiment", *Journalism Quarterly*, Vol. 65 No. 1, pp. 107-11.

Woodward, A. (1993), "Do illustrations serve a purpose in US textbooks?", in Britton, B.K., Woodward, A. and Binkley, M. (Eds), *Learning From Textbooks: Theory and Practice*, Lawrence Erlbaum Associates, Hillsdale, NJ, pp. 115-34.

Zgusta, L. (1989), "Ornamental pictures in dictionaries", in James, G. (Ed.), *Lexicographers and Their Works*, University of Exeter, Exeter, pp. 215-23.

Appendix. Taxonomy of image-text relationships

Each entry includes: a number indicating its order in the taxonomy; the name of the concept (function or relationship); a scope note (SN) that explains or defines the concept; lead-in terms based on other researchers' names for the same concept; and sources for the concept. Occasionally, the scope note includes a separate note headed "Note:" that relates the concept to other concepts as necessary to insure appropriate use. The numbers in the sources sections refer to research listed in Table I. If the term is based on the stage 2 analysis in this paper, it is marked EM. Any missing information is labeled "Not given".

A Functions expressing little relation to text

A1   Decorate
                SN: make the text more attractive without aiming to produce any real effects on the reader's understanding or memory. Note: can be applied even if other functions are used.
                    Lead-in terms: attract interest; draw attention; gain attention.
                    Sources: 1, 7, 8, 10, 11, 14, 15, 19.

A1.1   Change pace
　　　SN: interrupt continuity by shifting to a different activity.
　　　　Lead-in terms: novel stimulus; provide something new or unusual.
　　　　Sources: 9

A1.2   Match style
　　　SN: image and text match along same stylistic dimension.
　　　　Lead-in terms: not given.
　　　　Sources: EM.

A2   Elicit emotion
　　　SN: encourage emotional response from reader through display of content
　　　or style that is especially arresting or disturbing.
　　　　Lead-in terms: tone or mood inducing; create a specific emotional
　　　atmosphere designed to engage the reader.
　　　　Sources: 19, 20.

A2.1   Alienate
　　　SN: create tension between image and text through contrast in style or
　　　mood. Note: for content-based differences, use B3.2 (contrast).
　　　　Lead-in terms: clash; contradict; counterpoint; text and image collaborate
　　　to communicate meanings beyond the scope of either one alone; diverge.
　　　　Sources: 3, 4, 16, 17.

A2.2   Express poetically
　　　SN: suggest the spiritual qualities or effects of the object depicted.
　　　　Lead-in terms: not given
　　　　Sources: 23,

A3   Control
　　　SN: exercise restraining or directing influence.
　　　　Lead-in terms: direct; regulate activities or course.
　　　　Sources: 9.

A3.1   Engage
　　　SN: hold the attention of the reader. Note: if method of engagement is
　　　primarily emotional, then use A2 (elicit emotion).
　　　　Lead-in terms: not given.
　　　　Sources: 24.

A3.2   Motivate
　　　SN: encourage some response from reader. Note: if desired response if
　　　emotional in nature, then use A2 (elicit emotion).
　　　　Lead-in terms: not given.
　　　　Sources: 9.

B Functions expressing close relation to the text

B1   Reiterate
SN: restate with minimal change or interpretation.
Lead-in terms: direct; double, reinforce; one source provides a visual or textual restatement of another; repeat; symmetrical; transcribe.
Sources: 2, 11, 15, 16, 17, 23.

B1.1   Concretize
SN: make explicit. Use for captions especially. "Concretize" a textual reference to a thing or concept. Note: by definition, this action produces a shorter or less detailed account than B1.3 (describe) and a less complex account than C1 (interpret) or C3.2 (model).
Lead-in terms: make the unseen visible.
Sources: 12, 15, 19, 23.

B1.1.1   Sample
SN: give a sense of the concept by providing an example, although not a paragon or ideal.
Lead-in terms: not given.
Sources: EM.

B1.1.1.1 Author/Source
SN: use when author or source of image is given within a caption.
Lead-in terms: not given.
Sources: 20.

B1.2   Humanize
SN: represent elements within the text in the form of a living being; thereby making the text more accessible.
Lead-in terms: not given.
Sources: 15, 23.

B1.3   Common referent
SN: text and image share same symbolic source of meaning.
Lead-in terms: not given.
Sources: 9.

B1.4   Describe
SN: represent or give an account by definition; concretize. By definition, this action produces a longer or more detailed account than B1.1. Note: do not use for captions that merely identify.
Lead-in terms: discuss; present in detail for examination or consideration; convey ideas; impart or communicate by statement, suggestion, gesture, or appearance.
Sources: 23, 24.

B1.5   Graph

SN: translate numeric data into a visual representation.
    Lead-in terms: not given.
    Sources: 19.

B1.6   Exemplify
    SN: present a paragon that captures the essential meaning of a concept. Note: this code is applied to an image-text pair when it is used by the advertiser to present a given product line.
        Lead-in terms: not given.
        Sources: 7, 9, 23.

B1.7   Translate
    SN: convert from one form to another.
        Lead-in terms: symmetry; text and image repeat the same content; represent; reinforce by repeating written content in visual form.
        Sources: 3, 6, 11.

B2   Organize
    SN: form into a coherent unity or functioning whole. Includes advance organizers.
        Lead-in terms: make coherent; organize or structure text.
        Sources: 9, 12, 15, 20.

B2.1   Isolate
    SN: select and separate from others.
        Lead-in terms: not given.
        Sources: 9, 20.

B2.2   Contain
    SN: keep within limits. Includes Venn diagrams, flowcharts, timelines, and advance organizers.
        Lead-in terms: not given.
        Sources: 20, 23.

B2.3   Locate
    SN: set or establish in a time or place.
        Lead-in terms: establish setting; indicate time, place, era; stage setting; images or text introduces or sets the scene for the other.
        Sources: 2, 19, 20, 22.

B2.4   Induce perspective
    SN: encourage reader to see things in their true relations or relative importance. Note: more complex than B2.3 (locate).
        Lead-in terms: agenda setting; stage setting.
        Sources: 18, 21, 22.

B3   Relate

SN: these terms refer to processes intended to bring out concepts contained wholly within the text. Functions classed under C2.1 (analogize) and C2.2 (contrast) use concepts outside the text to explain and interpret.

Lead-in terms: not given.

Sources: 20.

B3.1 Compare

SN: make explicit intended elements of comparison between objects depicted in text.

Lead-in terms: not given.

Sources: 6, 9, 23.

B3.2 Contrast

SN: make explicit intended elements of contrast between objects depicted in text.

Lead-in terms: not given.

Sources: 6.

B3.3 Parallel

SN: image shows the same action or state in a different context than that presented in the text.

Lead-in terms: not given.

Sources: EM.

B4 Condense

SN: reduce to essential elements.

Lead-in terms: simplify; reduce to basics or essentials.

Sources: 9.

B4.1 Concentrate

SN: bring the most critical information to the reader's attention. Note: reduces text more than B4.2 (compact).

Lead-in terms: reduce. Some aspects of text are ignored to emphasize others.

Sources: 4, 12.

B4.2 Compact

SN: represent succinctly. Note: reduces text less than B4.2 (concentrate).

Lead-in terms: summarize; tell or reduce to an abstract, abridgement, or compendium; make concise.

Sources: 12.

B5 Explain

SN: make plain or understandable. Note: use only when original text is followed closely; if external constructs are used to explain by means of comparison or contrast, use terms under C2.1 (analogize) or C2.2 (contrast).

Lead-in terms: clarify; make comprehensible.

Sources: 1, 5, 10, 15, 23.

B5.1   Define
       SN: determine or identify the essential qualities or meaning.
       Lead-in terms: elucidate; explicate.
       Sources: 1, 2, 5, 6, 10, 12, 23.

B5.2   Complement
       SN: one mode helps the other to convey the message.
       Lead-in terms: not given.
       Sources: 3, 16, 22.

C Functions that go beyond the text

C1   Interpret
       SN: provide illustrations of complex ideas in concrete form. Note: provides a
       lesser degree of interpretation than C3.2 (model).
       Lead-in terms: not given.
       Sources: 1, 11, 15.

C1.1   Emphasize
       SN: provide force or intensity of expression that gives impressiveness or
       importance to something.
       Lead-in terms: underscore; make evident.
       Sources: 9, 23.

C1.2   Document
       SN: provide factual or substantial support.
       Lead-in terms: instruct; provide with authoritative information or advice.
       Sources: 9, 13, 19.

C2   Develop
       SN: set forth or make clear by degrees or in detail.
       Lead-in terms: amplify; expand (as a statement) by the use of detail or
       illustration or by closer analysis; elaborate; expand something in detail;
       enhance, one medium amplifies the meaning of the other; expand, express
       at length or in greater detail; extend, provide additional details; specify,
       bring out information about the story embedded in the text; supplement.
       Sources: 1, 2, 3, 4, 16, 22, 23.

C2.1   Compare
       SN: emphasize points of similarity between image and text. Note: unlike B3
       (relate), the point of reference rests outside the original meaning of the text.
       Lead-in terms: analogize; create a resemblance in some particulars
       between things that are otherwise different.
       Sources: 9.

C2.2   Contrast

SN: emphasize points of difference between image and text. Note: unlike B3 (relate), the point of reference rests outside the original meaning of the text.
Lead-in terms: clash, contradict, text and image are in opposition for a rhetorical purpose; counterpoint, add new information, creating a wholly different narrative thrust; different viewpoint, differing as to a position from which something is considered or evaluated; diverge; oppose, add new information to provide depth or insight; juxtapose.
Sources: 2, 3, 4, 16, 17, 22.

C3   Transform
SN: recode into concrete form; relate components to each other; provide organization to facilitate recall. Note: introduces more interpretation than C1 (interpret), C2.1 (analogize), and C2.2 (contrast).
Lead-in terms: code; perform a mnemonic function and make the text more readily available for processing into memory.
Sources: 10, 11, 12, 15.

C3.1   Alternate progress
SN: the text and illustrations "take turns" in progressing the story.
Lead-in terms: not given
Sources: 3, 4.

C3.2   Model
SN: provide a description or analogy used to help visualize something that cannot be directly observed. Note: provides a greater degree of interpretation than C1 (interpret). Use this when a more specific model code does not or cannot capture the function shown.
Lead-in terms: correspond; provide links between reader's pre-existing knowledge and new ideas within text; create metaphor.
Sources: 4, 6, 9, 12.

C3.2.1 Model cognitive process
SN: provide visual representation of abstract process.
Lead-in terms: not given.
Sources: 6, 9, 12.

C3.2.2 Model physical process
SN: provides visual representation of material or mechanical process.
Lead-in terms: not given
Sources: EM.

C3.3   Inspire
SN: using the text as a starting point, the illustration veers away to introduce new content that adheres to the spirit of the original story.
Lead-in terms: not given.
Sources: 4.