APPLIED THEORY

SUMMARY

- **Discusses the intrinsic properties of visual codings, as opposed to verbal ones**
- **Establishes that visual representations, because they are limited to conveying the concrete, are no panacea: words convey the abstract better than a thousand pictures**

Verbal Versus Visual: A Word Is Worth a Thousand Pictures, Too

JEAN-LUC DOUMONT

A picture is worth a thousand words, or so the saying goes—a saying debated by some but accepted pragmatically by most. Do we not all remember some little drawing or other that came in handy to clarify an otherwise plainly unintelligible discourse? Professionally, experienced technical communicators know the benefit of adding illustrations to the text of their technical publications. With increasingly better tools available for their production, pictures seem to have a bright future indeed.

Still, not all pictures are created equal, and the power of visual communication is sometimes misunderstood, if not misused. For example, in an effort to "humanize" graphical representations, a recent *Technical communication* article (Dragga and Voss 2001) proposed well-meant, but ineffective or inappropriate displays by failing to take into account basic principles of visual communication. Pictures are no panacea; some words may convey concepts better than a thousand pictures, too.

After clarifying the fundamental differences between verbal codings (text) and visual ones (pictures), this article establishes what visual codings are good and less good at doing, before offering pragmatic recommendations for their optimal use. It then applies these concepts to the case of the so-called "cruel" graphs discussed by Dragga and Voss.

PICTURES VERSUS WORDS

Mental processes are complex, multiple, and still poorly understood. Pragmatically, they can be modeled along various axes. Some processes are rational, able to manipulate more abstract, more intangible concepts, whereas others are intuitive, almost unconscious, and possibly innate. Some processes are clearly sequential, having a beginning, a middle, and an end, and therefore appear slow, while others seem—for all practical purposes—instantaneous or simultaneous. From a communication point of view, some processes are symbolic, able to manipulate a human language or language-like references or constructs, while others are not (see, for example, Deacon 1997).

More frequently, technical communicators simply distinguish between verbal and nonverbal processes. *Verbal*, here, refers to languages, with the symbolic arbitrariness of their semantics and the sequential logics of their syntax. *Nonverbal*, then, refers to largely nonsymbolic and nonsequential codings, regardless of whether they are perceived by the ears (for example, *vocal* communication, as encoded in variations of the tone, rate, and volume of the voice) or by the eyes (for example, *visual* communication, as encoded in a picture, a gesture, or a facial expression). *Verbal* and *nonverbal*, in this sense, are perhaps best understood as *symbolic* and *nonsymbolic* in the remainder of this article.

Verbal and nonverbal processes are complementary. To a point, they can be performed independently and do not seem to compete for the same intellectual resources. For example, players of visual computer games such as Tetris (Figure 1) can keep on chatting with onlookers while engaging in purely visual reasoning against the clock (making decisions about where and in which orientation to place the falling shapes). Similarly, students can doodle and still not miss a word the teacher is saying. And drivers can talk on their cellular phones and keep their cars on the road. Sequential processes, however, are mutually exclusive: students chatting among themselves, for example, can no longer listen to the teacher, and drivers engaged in a

Manuscript received 10 August 2001; revised 17 December 2001; accepted 18 December 2001.

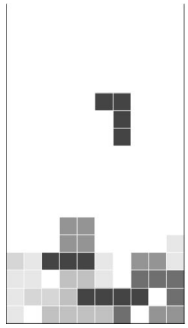

Figure 1. Verbal and visual processes can be performed independently. One can play Tetris (moving and rotating shapes for best fit) and chat with onlookers at the same time.

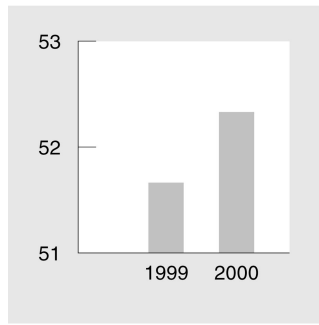

Figure 2. Visual codings have more impact than verbal ones. Despite the correct left scale, one walks away from this graph feeling that 2000 was twice 1999.

phone conversation are less able to react to unexpected circumstances that require reasoned decisions. Nonverbal processes, by contrast, can more easily coexist.

Visual refers here to the coding, as processed by the brain, not to the channel, as perceived by the senses. In this sense, so-called "visual aids" (slides) used by speakers to "illustrate" their presentations are seldom visual: if they display nothing but text, they are clearly verbal. (As such, they are not "aids" either, at least not to the audience, as they compete with what the speaker says for the audience's verbal attention.) True visual aids are grasped in a few seconds. Often, they are absorbed nonverbally: when asked later what was on them, audience members find themselves at a loss for words and start making gestures, often redrawing the illustration in the air.

Nonverbal codings, being intuitive, have more impact than verbal ones. In case of conflict, they usually prevail over the verbal ones (Mehrabian 1981): we believe tone of voice and body language more than words, possibly because we know (rationally) that nonverbal signals are less controlled by the speaker, more probably because we absorb the nonverbal signals without the critical filter we apply to words. (In other words, we do not reason about the conflict: somehow, we simply "feel" that the person is lying. Nonverbal processes are, in some sense, more remote from consciousness. Some of them are largely innate: newborns, for example, can interpret tone of voice.)

This conflict between verbal and nonverbal communication is put to good use in irony and humor, when we let people know nonverbally that we do not mean what we say verbally. The superior credibility of nonverbal codings also explains (if need be) the hopelessness of do-as-I-say-but-not-as-I-do approaches: a technical manual showing (nonverbally) what *not* to do and indicating (verbally) that this is a "don't" is more likely to make readers learn the wrong gesture than to make them recognize it as wrong. Sports coaches similarly know that students will simply mimic them, often mirror-imaged, rather than listening to what they are asked to do.

Because they are powerful, visual codings are dangerous. Carefully designed, they work wonders; poorly designed, they backfire. Graphs can thus be intentionally or accidentally deceptive—for example, if they suggest a length representation but lack a zero reference point. If one bar of a chart looks twice as long as another, it strongly suggests a value twice as large for the encoded variable, no matter what the rational information along the scale says (Figure 2). Graphical integrity has received much attention over the last decades (see, for example, Tufte 1983).

Visual codings, like any sign meant to represent an object, can be interpreted at three different levels—icon, index, and symbol—a theory formalized a century ago by Charles Sanders Peirce. At the iconic (or literal) level, a picture of an elephant means an elephant. At the indexical (or metaphorical or analogic) level, a picture of an elephant means something large or heavy like an elephant. At the symbolic (or conventional) level, a picture of an elephant means the Republican Party in the U.S. or a famous brand of chocolate in Belgium. While remaining a tangible sign (it does look like an elephant), the symbolic elephant is closer to a word than to a picture; in a sense, it is no longer a visual coding as defined above, but has become a verbal one. Similarly, hieroglyphics and sign language are clearly languages, that is, verbal codings (Pinker 1994).

WHAT PICTURES ARE GOOD AND NOT GOOD AT DOING

Visual codings, because they are intuitive and global, are clearly superior for conveying intuitive or global information. Facial expressions, for example, convey emotions infinitely more subtly than words do (we perceive and interpret minute changes in facial features that we can often not describe). Maps convey relative positions infinitely more rapidly than words do. Drawings describe

Figure 3. Visual codings are intrinsically ambiguous. This pictogram is difficult to interpret unambiguously without the accompanying text ("Turn automobile engine off").

concrete objects infinitely more clearly than words do. It is no wonder, then, that visual codings such as ";–)" emerge in otherwise plain-ASCII worlds such as e-mail to convey meaning that relies on intuition, such as humor.

Visual codings, by contrast, do a poor job of expressing abstract concepts. A given visual representation illustrates but *one instance* of a concept so easily expressed in words. How would you, for example, render visually the concept "soft drinks," to identify the corresponding aisle of a food supermarket? A photograph of an actual soft-drink container suggests a brand, not a generic concept. A drawing can be more generic, but nevertheless suggests specific types of container (cans, bottles, . . .) rather than a specific type of beverage. Moreover, the colors and shapes used in the drawing are likely to suggest brands, even if unintentionally. A generic sparkling beverage in a generic plastic cup might not suggest a brand, but it suggests sales at a soda fountain, not in closed containers. Nonsymbolic, visual representations are condemned to be concrete.

Visual codings, moreover, lack the accuracy that words are endowed with through conventional association of meaning. Like Rorschach tests, they are intrinsically ambiguous: everyone might see something different and yet be unable to imagine anyone else seeing anything else. In a short exercise I often conduct in my training programs, a pictogram showing a key in a red circle crossed by a red diagonal (Figure 3) has been interpreted by training participants simply as "No keys allowed" or "Do not lock the door," but also as "Metal objects not allowed" and "The solution does not exist," depending on whether the key is interpreted iconically, indexically, or symbolically. At gas stations (where I found it), the pictogram is used to mean "Turn automobile engine off."

Words, in a sense, are worth a thousand pictures. They can express abstract concepts unambiguously and concisely, even if not intuitively. The words *soft drinks*, for example, designate all possible soft drinks and thus transcend all pictures that can show only specific soft-drink containers of specific brands. Words can convey concepts that cannot be expressed by nonsymbolic codings, such as the negation or the hypothesis. In pictograms, the negation requires a purely conventional (symbolic) sign, such as the red circle with red diagonal of Figure 3.

WHEN AND HOW TO USE PICTURES

Visual codings, like other codings, should serve the purpose of the act of communication that they contribute to. Used indiscriminately, for example as a panacea, as a mere attention getter, or as a way to shun translation expenses, they may well detract from the document's intent. Their intrinsic properties, as outlined above, suggest the following three practical recommendations:

1. When being visual, be truly visual.

2. Select a level of pictorialness that suits the desired level of abstractness.

3. Combine visual and verbal codings to get messages across.

First, effective illustrations are truly visual, not verbal. In other words, they do not rely on symbolic association and can thus be interpreted correctly without a verbal step. As a counterexample, separate legends, because they are symbolic, turn graphs into something to read, not something to look at (Bertin 1973). After briefly viewing a pie chart with separate legend (Figure 4a), viewers easily remember the largest segment as being "the red one" or "the one on the bottom" (something they have seen at a glance), but can often not associate it with its meaning in the legend (something they needed to—but perhaps did not—read), especially if they were listening to a speaker while viewing the chart on screen. A chart with direct labeling (Figure 4b) works better: the association between pie segment and matching word is now immediate and intuitive because it is based on spatial proximity, not arbitrary color coding.

Visual communication tolerates isolated words better than text. The symbolic key used in a separate legend is verbal in nature because it is both symbolic and sequential: knowing that, in a pie, *red* means *Europe* is arbitrary. When interpreting such a chart, viewers often go through an explicit verbal step, which they sometimes even voice out loud ("Let's see now, what does *red* mean?"). Well-positioned isolated words, while still symbolic by nature, relate to what they describe in a nonsymbolic way (proximity, not color) and do not force a sequential process; hence, they respect, as well as words can, the nonverbal nature of the illustration.

Truly visual codings thus suggest a preference for iconic or possibly indexical associations over symbolic ones. Most of us are familiar with the international pictograms for men's and women's restrooms, sketching a male and a female, respectively (Figure 5). The association with the sexes is iconic, but the association with restrooms is at best indexical, if not plainly symbolic (conventional). In some countries (and increasingly so in the U.S., in combi-

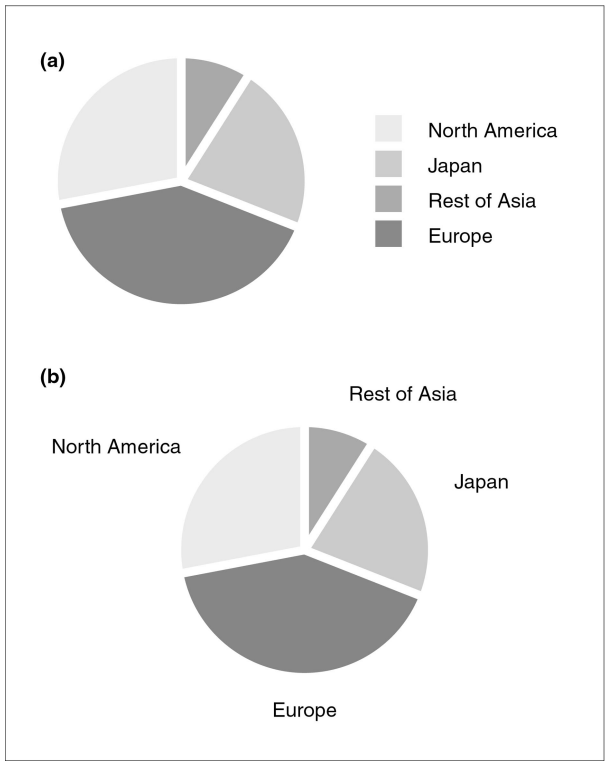

Figure 4. Truly visual graphs avoid the symbolic association of separate legends (a) and use direct labeling (b) instead.

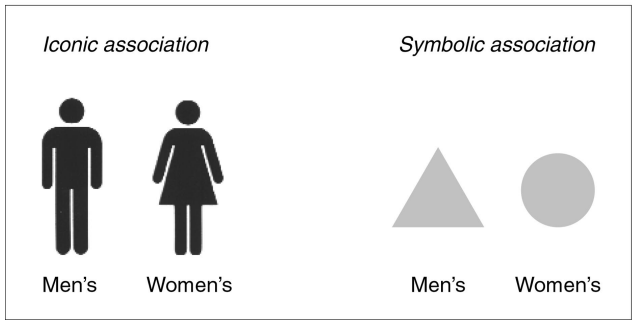

Figure 5. Truly visual codings prefer iconic or indexical associations (left) to symbolic ones (right).

nation with the male and female icons), men's and women's restrooms are indicated by a triangle and a circle, respectively. Such pictograms are clearly verbal: intuitive guessing is ruled out—you have to "know the semantics" to understand them.

Second, while all visual codings are ill-suited to expressing abstract concepts, drawings nevertheless do better than photographs (Figure 6). They should thus be preferred over photographs when being abstract is part of the purpose, and vice versa. Careful technical illustrators actually invest much time and effort in improving the abstract character of their illustrations. If they draw a human hand manipulating a piece of equipment, they will strive to make the hand ageless, genderless, and raceless—something a photographed hand can hardly achieve. Drawings can also "abstract" much of the irrelevant visual information present in a photograph and therefore be less distracting. By contrast, the visual richness of a photograph can more easily convey an atmosphere, something advertising agencies know well.

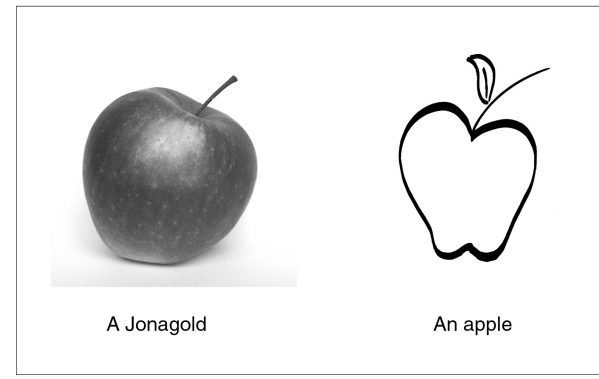

Figure 6. Drawings do a better job of expressing abstract concepts than photographs. (Words do better still, even when the concept is as minimally abstract as "apple.")

Neither effective technical drawings nor effective photographs come cheap. With the advent of inexpensive digital cameras and color printers, however, writers of user's guides now seem tempted to prefer photographs (something they can readily produce themselves) to drawings (something they must obtain from a third party such as another employee, a contractor, or a clipart source). In such a decision, available technology unfortunately prevails over more fundamental arguments, often at the cost of the document's effectiveness.

Third, effective documents use both verbal and nonverbal codings, not one or the other. Because they match distinct intellectual processes, the two codings ideally complement each other to provide an even richer vision. Visual codings can be complemented by words in two ways. First, they benefit from being properly labeled with words but, as stated above, careful labeling hardly modifies their visual character. Second, they benefit from being properly commented with sentences, stating unambiguously the (abstract) message that they are meant to illustrate.

Effective redundancy between verbal and visual codings suggests that messages be, as much as possible, conveyed *both* verbally and nonverbally. It thus goes further

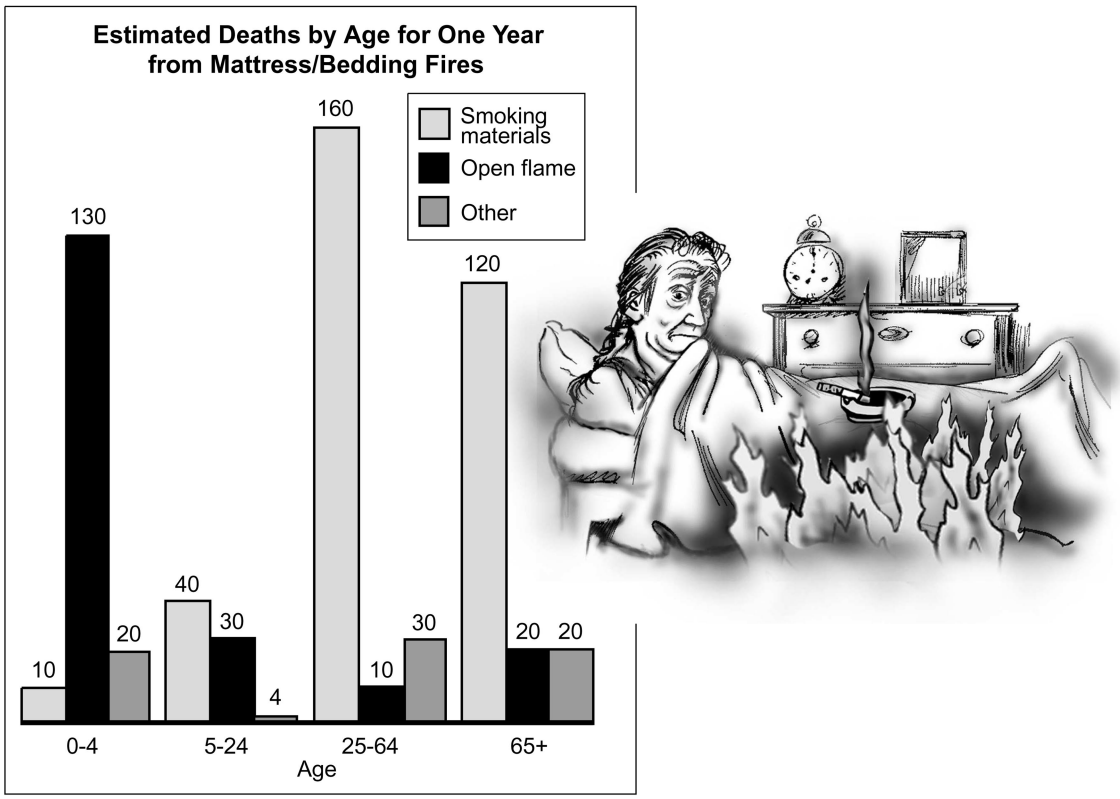

Figure 7. To humanize the "cruel column graph [that] offers an emotionless tally of human fatalities" (left), Dragga and Voss propose to add "a bed, a person, and a cigarette [to] put flesh, blood, and tragic foolishness behind the cold statistics" (right). Because of the intrinsic properties of visual codings, however, the added illustration detracts from the graph much more than it enhances it.

than using verbal codings only for one part of the contents and visual codings only for the other part. It works through both *compensation* (readers who do not understand the text may understand the graph and vice versa) and *collaboration* (readers who understand both the text and the graph get more out of the combination).

AN EXAMPLE: THE CASE OF CRUEL GRAPHS

In a recent issue of *Technical communication*, Sam Dragga and Dan Voss point out what they call "the inhumanity of technical illustrations" through a sampling of "cruel" graphs—graphical representations that (they say) fail to convey the horror behind the numbers they are coldly encoding (Figure 7). Their statements about the ethics of visual communication and their proposed graphical solutions call for at least three comments as an illustration of the points made above.

First, the largely shared view that professional communication is purpose-driven suggests that we design visual codings in a way that best serves this purpose. If visual representations of human suffering help readers make better informed decisions, then they are fully justified, for they do contribute to more effective communication. Such effectiveness, in itself, requires accuracy, if not honesty (no data distortion, no graphical lies). So do ethics, of course, but ethics are a largely subjective matter, rooted in culture, if not religion. Still, I fail to see how any ethic creates the need to depict visually the horror of the situation discussed: in what sense, for example, do such visual representations show more respect for any of the people involved (those who suffered, the readers, or the authors)? Many people I know would not want their suffering on display.

Second, the qualification "cruel" that Dragga and Voss attribute to emotionless graphs seems misused at best. If "cruel" means "causing, or of a kind to cause, pain, distress, etc.," then an emotionless graph is infinitely less cruel than a visual display of human suffering. The authors' proposal to include a photograph of the patient in medical records strikes me, in fact, as needlessly cruel.

Third, and maybe most important, the solutions pro-

posed by Dragga and Voss (assuming we accept the need to "humanize" the graphs) work poorly, because of the very nature of visual codings. Showing "a bed, a person, and a cigarette [to] put flesh, blood, and tragic foolishness behind the cold statistics" (Figure 7) offers *one very specific example* and thus fails to capture the generality of the accidents discussed. To me, the character depicted in the proposed solution looks surprised and worried, but by no means in pain or in mortal danger. Because I do not smoke (in bed or elsewhere), I fail to see the relevance of the display to me and may lose interest, whereas the graph also reports other accidents, such as from an open flame, to which I might more easily relate. Moreover, the added drawing includes many details to look at: all the time I spend looking at those largely irrelevant details is time I do not spend understanding the data—or understanding a message about the data. Among such details are the age, gender, and race of the individual shown, something to which many audiences are increasingly sensitive.

Photographs are a similar example of the dangerous power of rich visual communication. To discuss further the proposal of Dragga and Voss, one picture of a patient says so much . . . and so little at the same time. Two pictures of the same patient, taken at slightly different times, may well say very different things about him or her, and thus have very different impacts on viewers (imagine the patient in pain in one picture and resting quietly in the other). Which one is most ethical? Which one most closely "tells the truth" about the patient? More importantly, which one will help a medical doctor do his job better (for example, make a better decision about the patient)? Each picture could be more limiting than it is enlightening. Similarly, when I receive and use photographs for identification (for example, when I want to learn ahead of time the name of the participants of an upcoming training program), I later find that few people resemble the "image" that I had formed of them on the basis of their picture.

FINAL WORDS

When I was a graduate student, my advisor always asked two questions whenever I would show him a quick and dirty graph at our weekly research group meeting. The first was "What am I looking at?"—an indication that my graph was improperly labeled. A few words, positioned where the viewer needs them most, would solve that problem. His second, and the more incisive (though never derogatory), was "Why are you showing this to me?"—an indication that my graph was improperly commented. A few statements, conveying the *so what* rather than the *what*, would solve that problem. In other words, my advisor did not want me to show a visual coding for the sake of showing a visual coding; he wanted me to focus on a purpose instead and use whatever coding—or combination of codings—best allowed me to get the message across. Only then would I have my full added value.

The quest for effectiveness should admittedly not make us lose our humanity. Still, what makes us human is more than simply empathy toward suffering members of our species—or of others. Maybe our most distinguishing human characteristic, certainly compared with other animal species, is our ability to generate, manipulate, and communicate abstract concepts, among others, through symbolic codings. In this sense, dispassionate graphs are most human, too. T**C**

REFERENCES

Bertin, Jacques. 1973. *Sémiologie graphique*. Paris, France: Mouton-Gauthiers-Villars.

Deacon, Terrence. 1997. *The symbolic species: The coevolution of language and the human brain*. London, UK: Penguin Press.

Dragga, Sam, and Dan Voss. 2001. "The inhumanity of technical illustrations." *Technical communication* 48:265–274.

Mehrabian, A. 1981. *Silent messages: Implicit communication of emotions and attitudes* (2nd ed.). Belmont, CA: Wadsworth.

Pinker, Steven. 1994. *The language instinct: How the mind creates language*. New York, NY: William Morrow.

Peirce, Charles Sanders. "Logic as semiotic: The theory of signs." In J. Buchler, ed. 1995. *The philosophical writing of Peirce*. New York, NY: Dover Books.

Tufte, Edward R. 1983. *The visual display of quantitative information*. Cheshire, CT: Graphics Press.

JEAN-LUC DOUMONT teaches and provides advice on professional speaking, writing, and graphing. He also trains trainers and facilitates any process that requires structuring and effective communication. For over 15 years, he has helped audiences of all ages, backgrounds, and nationalities structure their thoughts and construct their communication. He holds an engineering degree from the Université catholique de Louvain and a PhD in applied physics from Stanford University. Contact information: JL@JLConsulting.be