Grouping, Simile, and Oxymoron in Pictures: A Design-Based Cognitive Approach

Norman Y. Teng
Institute of European and American Studies
Academia Sinica, Taiwan

Sewen Sun
Institute of Telecommunications
National Chung Cheng University, Taiwan

Researchers have identified 2 distinctive types of pictorial displays, namely, pictorial metaphor and pictorial simile, and offered theoretical explanations of them. Regarding the distinction between pictorial metaphor and pictorial simile, we argue that symmetric image alignment of pictorial components depicting things at the object level is the principal design factor that sets pictorial simile apart from pictorial metaphor and links pictorial simile to pictorial grouping. Based on the idea of symmetric image alignment, an attempt is made to identify and explain a further type of pictorial display, namely, pictorial oxymoron. We propose a hypothesis, called the “image grouping hypothesis,” to provide an integrated account of pictorial grouping, pictorial simile, and pictorial oxymoron. The hypothesis can be summed up as follows: Symmetric image alignment of pictorial components depicting things of the same kind is apt for expressing pictorial grouping; symmetric image alignment of pictorial components depicting things of different kinds is apt for expressing pictorial simile; and symmetric image alignment of pictorial components depicting things that can be seen as incompatible with each other is apt for expressing pictorial oxymoron.

Pictures in ads have begun to shape what might be called “pictorial tropes” in our everyday discourse. Their designs are often clever, creative, sometimes surprising and explorative, yet they are easy to consume, and designedly so. Researchers in this
field have identified two distinctive types of pictorial tropes, namely, pictorial metaphor and pictorial simile, and offered theoretical explanations of them (e.g., Forceville, 1994, 1996; Kaplan, 1990, 1992; Messaris, 1997). It is worth mentioning that many aspects of pictorial metaphor and pictorial simile investigated in relation to advertising also show up in film and art (for relevant studies, see Carroll, 1994, 1996; Forceville, 1999b, 2002; Whittock, 1990). This is to be expected, for the basic design of the pictorial tropes is based on the cognitive resources we all share, as is shown later in this article. The task we seek to accomplish is to separate out pictorial simile from pictorial metaphor in cognitive terms and to provide a broader scheme in which grouping, simile, and oxymoron in pictures can be systematically identified and coherently explained.

According to Forceville (1996), pictorial simile is similar to pictorial metaphor, in that both kinds of tropes involve a cognitive process of understanding and experiencing one kind of thing (the target) in terms of another (the source). (This formulation is adapted from Lakoff & Johnson, 1980, p. 5; see also Gibbs, 1994; Johnson, 1987, 1993; Lakoff, 1987b, 1993, 1996; Lakoff & Johnson, 1999; Lakoff & Núñez, 2000; Lakoff & Turner, 1989; Sweetser, 1990; Teng, 1999; Turner, 1987, 1991.) They differ in their visual manifestations. Pictorial simile has image juxtaposition of the source and the target as its defining feature. Pictorial metaphor, on the other hand, has image integration of the source and the target as its defining feature. This design-based difference has the cognitive effect that pictorial similes are weaker than pictorial metaphors in expressing figurative messages. This is certainly a welcome feature, for the cognitive effect echoes that of a verbal simile and a verbal metaphor, at least when they are expressed in their respective paradigmatic forms, namely, “A IS LIKE B” and “A IS B.” For example, the verbal simile “She’s like a walking dictionary” is weaker than the verbal metaphor “She’s a walking dictionary” in regard to their figurative effects. The former suggests a similarity between the source (dictionary) and the target (the person “she” refers to), with the target retaining its relatively cognitive independence from the source. The latter suggests that the target is metaphorically transformed into a walking dictionary. This design-based, cognitive effect allows us to use verbal simile and its relation to verbal metaphor as a reference point for understanding pictorial simile.

Forceville (1996) made the following observation about the difference between verbal similes and verbal metaphors:

Metaphors are experienced as stronger than similes because the short-cut of the metaphorical ‘is’ constitutes a less explicit and hence more surprising link than the simile’s ‘is like,’ the metaphorical ‘is’ to a greater degree suggests transformation of the primary subject [the target] by the secondary subject [the source] than does simile’s ‘is like.’ (p. 142)
Here we set aside the first part of this observation, for it seems to us that verbal similes can sometimes be surprising too. One might try a different tack and argue that verbal similes are different from verbal metaphors simply because verbal similes are always trivially true, in the sense that the statement “A is like B” is always correct for some attribute or other, no matter how irrelevant it may be. This argument, however, cannot be correct when we take into consideration psychological facts of how similarities are actually processed.

Thus clarified, the issue comes to this. We agree that both pictorial simile and pictorial metaphor involve a cognitive process of understanding one kind of thing in terms of another. We also agree that what distinguishes them from each other are their visual manifestations, the ways they are pictorially represented. But we contend that merely stating that pictorial simile is weaker than pictorial metaphor is unsatisfactory. As is shown in this article, the ways they are represented tap into different cognitive resources that support our activities of reading and making images. A design-based hypothesis, called the “image grouping hypothesis,” is proposed, which shows how pictorial simile is cognitively anchored to, and grounded in, pictorial grouping and thus sets pictorial simile apart from pictorial metaphor.

Based on an initial formulation of the image grouping hypothesis, an attempt is made to identify a further type of pictorial trope, namely, pictorial oxymoron. We then, by extending the image grouping hypothesis to pictorial oxymoron, explain how oxymoron works in pictorial terms. A formulation of the hypothesis in a more general form is then put forward.

Certain preliminary remarks are in order before we proceed. First, image reading is context sensitive. Knowing that the images are advertisements certainly will affect one’s interpretations of them (Phillips, 1997). We take this as a background setting and focus on the design patterns that further constrain image reading in such background settings (see Forceville, 1996, chap. 4, for an account of such background settings). Second, we take image reading and image making as two sides of the same coin from a design-based, cognitive point of view. To foreground certain points, we sometimes write as if we were prescribing what should be done to achieve certain effects. Nonetheless, our primary concern is describing the design patterns that are already implicit in the practice of reading and making images. Third, the working hypothesis that guides this study is that the design patterns are cognitively based, cross-cultural invariant structures, in that people with different cultural models can share the invariant structures or at least can smoothly incorporate the structures into the cognitive tools they use in reading images (see Maalej, 2001, for a cross-cultural account of pictorial metaphor related to Forceville’s theory). But this does not mean that people with different cultural models will come up with the same interpretations of the images. They might not even take the images as instances of pictorial metaphor or pictorial simile when the invariant structures have not yet been incorporated into their cognitive repertoire. Regarding this possibility, we need to make a distinction between explicit and im-
plicit presentations of metaphors and similes. An explicit presentation strongly suggests that a metaphoric construal is in order, but an implicit presentation only suggests that a metaphoric construal is possible but not necessary (Forceville, 1999b, pp. 191–194). Based on this distinction, we take it that the images still count as pictorial metaphors or pictorial similes even if they have not yet been explicitly recognized as such. Finally, we do not assume that our interpretations of the examples discussed in the following are the correct ones. Viable alternative interpretations are always possible. And our claim is a modest one: The design patterns constrain, but do not determine, how one interprets the images.

PICTORIAL GROUPING

Let us start with pictorial grouping. One distinctive feature of pictorial grouping is that the depicted entities are symmetrically aligned with one another with respect to size, orientation, and distance between them. The symmetric image alignment can be achieved in the following manner: Abstract the entities from the contexts in which they are normally situated; depict the entities as being similar in size and orientation and placed at equal distance from each other along a chosen axis, with the depth of the pictorial space reduced or absent and the background being plain and neutral. Overall, this design pattern has the following cognitive effect: The depicted entities do not interact with each other but are decontextualized so that they do not form an integrated scenario but are related to each other in standing for themselves as exemplifying their own type. This effect fits very well our everyday understanding of what a category is.¹

A number of points need to be clearly stated. First, this discussion is about how a category can be pictorially represented, rather than how people group things into the same category. There are many views of what a category is, for example, the family resemblance or prototype view (e.g., Rosch & Mervis, 1975), the exemplar view (e.g., Medin & Schaffer, 1978), the explanation-based view (e.g., Keil, 1989), and the cognitive model view (e.g., Lakoff, 1987a, 1987b). Our discussion is neutral regarding the question of which of these views is the correct understanding of human categorization. Second, further variations on the design pattern of pictorial grouping are in use, too. For example, the depicted entities can be laid out along the horizontal or vertical axes or both in the pictorial space. Diagonal alignment is also common; and when depth is strongly suggested, the diagonal is often slightly tilted toward the depth dimension. Third, the pictorial space is not the physical region or

¹Kress and van Leeuwen (1996, chap. 3) analyze pictorial grouping under the heading of conceptual representations and base their analysis on social semiotics. We adapt their concepts and put them to use in cognitive terms. For an account of the cognitive reorientation for image analysis, see Forceville, 1999a.
surface defined by the physical properties of the picture. Rather, it is the space represented by the pictorial arrangement or implied by the depicted scene in the picture. Because the depth of the pictorial space is often reduced or absent, the chosen axes often coincide with the corresponding axes of the physical surface of the picture. Fourth, the design pattern does not require that the symmetric image alignment be precise. An approximately symmetric image alignment is sufficient, in that selective attention to the overall pattern filters out the locally slight deviations as irrelevant to the symmetric image alignment thus detected. Finally, a subtle but important distinction must be made between a design pattern specified only schematically and a design pattern actually detected in its richly detailed, concrete instances. Mistake the former for the latter and conceptual confusion will result. Symmetric image alignment specified only schematically cannot have any cognitive effect, and, indeed, it cannot be said to have any effect at all unless it is pictorially realized in actual usage events. Actual instances of pictorial grouping are one sort of instance in which symmetric image alignment can be detected. Other sorts of instance are possible, as will become evident. It is thus important to keep in mind that symmetric image alignment detected in actual usage events can result in pictorial grouping only when the pictorial components laid out in this manner do in fact depict things of the same kind.

The technique of pictorial grouping is commonly used in advertisements, where photographs show a set of commercial products laid out in this manner. People can also be laid out in this manner. Nevertheless, the basic message is the same: The products or the persons belong to the same category, in the sense that the products are sold under the same brand name, the persons all use the products (e.g., the message “The most unforgettable women in the world wear Revlon” from a Revlon advertisement in Women’s Weekly, December 1987; see Kress & van Leeuwen, 1996, pp. 80–81) or whatever. It is clear that examples of pictorial grouping are ample, especially in print advertisements. Given the enormous influences of advertising and the pervasive uses of pictorial grouping in contemporary societies, we assume that the design pattern of pictorial grouping is already deeply entrenched in how we read and make images. This fact allows us to use the design pattern of pictorial grouping as a reference point with relation to which other design features of pictorial displays can be brought to the fore and put into proper perspective, as we shall see.

It is worth making a brief digression here to compare the perceptual grouping that Gestalt theorists advocate with the pictorial grouping we propose here. Gestalt theorists identify many different factors (for example, similarity of size, color, and orientation, spacing between visual elements, and so on) that govern which visual elements are perceived as going together in one’s perceptual experience. The visual elements that are perceived as going together are not cognized as belonging to the same category. They simply go together in larger groups that are perceptually significant in one way or another (see Palmer, 1999, chap. 6, for a detailed account
of perceptual grouping). In contrast, the pictorial grouping we advocate here is an artifact that is intended to produce the cognitive effect of inviting the viewers to see the depicted entities as belonging to the same category. Notice that the design pattern of pictorial grouping features a number of cognitive factors akin to those of perceptual grouping, such as similarity of size, similarity of orientation, and equal spacing, but the factors operate at the level of whole objects, rather than at the level of more primitive visual elements. This suggests that the design pattern and the pictorial components that depict entities at the object level can be kept distinct from each other, which in turn suggests that we can hold the design pattern as an invariant and vary the pictorial components to achieve new cognitive effects.

**PICTORIAL SIMILE**

The foregoing account of the difference between perceptual grouping and pictorial grouping marks the beginning of the descriptive task of what more we can do with symmetric image alignment. Here is a first move to get on with the task: Depict two entities of different kinds in a way that makes them visually recognizable as things of different kinds (see Lopes, 1996, for an account of pictorial recognition), and align them in a way that conforms to the design pattern of pictorial grouping. This should create a cognitive dissonance in image reading, in that the design pattern invites us to see the two entities as belonging to the same category, yet the entities are depicted as things of different kinds. The cognitive dissonance prompts us to find out or construct a novel interpretation of the picture if it is evident that we cannot dismiss the picture as merely falsely grouping things of different kinds into the same category. This effect makes the visual design particularly suitable for expressing similes in pictorial terms. It should be noted that it is not our contention that viewers will always first detect some anomalies in the picture and then proceed to interpret it as a pictorial simile. If the contextual setting of the picture is appropriate and the design pattern of pictorial simile is already entrenched in the viewers' cognitive systems, they may directly interpret the picture as a pictorial simile without first taking it as an anomalous picture.

To further elaborate and clarify the ideas discussed so far, an array of examples are considered. Let us begin with the following example (Forceville, 1996, p. 140): an ad for a Dutch beer, containing two photos of the same size and symmetrically aligned along the horizontal axis. On top is a photo of the Great Wall rolling over a mountainous region in China. Below it is a photo of a Dutch landscape with a small wall of full beer crates gently crawling along an undulating ground surface. The beer crates are assembled and photographed from an angle that makes their overall outlook similar to that of the Great Wall. According to Forceville (1996, p. 137), the picture contains a simile, which can be interpreted as follows: The Dutch beer
resembles the Great Wall of China in having, or deserving, a similar kind of national fame, prestige, or pride.

We agree with that interpretation. But what requires explanation here is how the pictorial representation can motivate such an interpretation. First, we know from the context that the Dutch beer is what the picture is about. After all, the ad is about the Dutch beer, rather than the Great Wall. It is also clear from the context that the ad cannot be dismissed as merely falsely grouping the Dutch beer and the Great Wall into the same category. Having the contextual setting in place, how might we expect the symmetric image alignment to affect the image reading?

It is commonplace that grouping objects into the same category may make them look similar to each other (Goodman, 1970). Now, the Dutch beer is what the picture is about, and it differs from the Great Wall in kind. Given the setting just described, linking the Dutch beer to the Great Wall in terms of symmetric image alignment would invite the viewers to see the Dutch beer as being similar to the Great Wall. Furthermore, the perceptual similarity between the two photographed scenes in the picture reinforces this invitation. Framing it this way should make it clear why symmetric image alignment can be an effective visual design for pictorial simile. The ad indeed contains a simile, with the Great Wall being the source, the Dutch beer the target, and the symmetric image alignment functioning like a connector linking the target to the source.

It is worth emphasizing here that a design-based, cognitive analysis of how similes can be represented in pictorial terms should not be taken to be an equivalent of a cognitive analysis of how similes are processed and comprehended. Though our analysis shows that pictorial simile is related to, and motivated by, pictorial grouping from a design-based point of view, we do not thereby contend that pictorial simile is essentially nonstandard categorization or class inclusion. For even if the symmetric image alignment functions like a connector linking the source and the target, the source and the target still are visually encoded as things of different kinds, and the latter is to be understood in terms of the former (for an account of the class-inclusion approach, see Glucksberg & Keysar, 1990, 1993; for critical evaluations of that approach, see Chiappe & Kennedy, 2001; Kennedy, 1997; Kennedy & Chiappe, 1999).

One may enrich the symmetric composition of a pictorial simile by introducing certain relations into the two depicted entities. Consider an ad for the magazine Business Weekly (see Figure 1): a portrait of a politician (or a businessman) and that of a shark are placed next to each other along the vertical axis, with the portrait of the politician on the left and that of the shark on the right. The visual design of this picture conforms to the design pattern of pictorial simile, with a fine touch added to it. The background of the two portraits is not quite plain and neutral, but made up of pages of a magazine. The Chinese characters at the right bottom corner indicate that the magazine is Business Weekly. The magazine is about business and politics, rather than marine life. This contextual information helps to fix the target,
and it does not really matter whether the target is on the left or on the right as long as it can be identified in a certain way (see Kress & van Leeuwen, 1996, pp. 181–188, for an account of the significance of the left/right arrangement; see Forceville, 1999a, pp. 171–172, for a critical review of that account). The visual design allows us to form the interpretation that politicians are like sharks (or, to elaborate it a bit, that politicians are to ordinary people as sharks are to other forms of marine life). The pictorial simile, under this interpretation, contains the message that the magazine watches politicians (and business people) critically.

A further elaboration on the design pattern of pictorial simile is possible. When the depth of the pictorial space is strongly suggested, the symmetric image alignment can be arranged along the depth dimension, placing the depicted entities in the same pictorial space. This alignment, unlike the previous two, typically has the source and the target depicted as participants involved in the same coherent scenario. Another of Forceville’s (1996, p. 138) examples, a German ad for swimwear, can illustrate this point. This ad features a girl in a bathing suit with a dolphin positioned next to her along the depth dimension of the same pictorial space. The girl and the dolphin seem to act together in the process of diving into the water, which suggests a coherent scenario of which the girl and the dolphin are a part. It should be noted that though the visual design of this picture still conforms to the design pattern of pictorial simile, the viewers might interpret this picture as depict-

![FIGURE 1](image_url) An advertisement for *Business Weekly* (Advertising agency: Ogilvy & Mather, Taiwan). Copyright 1997 by Ogilvy & Mather, Taiwan. Reprinted with permission.
ing a coherent scenario without taking it to be a pictorial simile. Notice, however, that several striking similarities between the girl and the dolphin are put into perspective: The girl’s well-fitting swimsuit and the dolphin’s smooth skin, the manners in which they dive into the water, and the graceful positions they assume are all depicted as similar to each other. All these striking aspects of resemblance between the two invite the viewers to take the symmetric image alignment seriously and interpret the picture as a pictorial simile.

Introducing story-like scenarios into pictorial simile makes it possible for contemporary image-makers to experiment with more sophisticated compositions of pictorial simile. Consider an ad for the protection of animals (see Figure 2), which depicts a human hand and a white front paw placed on the Bible as if a person and an animal are taking the oath together. Notice that the symmetric image alignment in this picture cannot be defined with reference to the pictorial space as a whole but has to be defined in terms of the spatial orientation of the Holy Bible in it. The hand and the paw are acting out in a similar way and seem to be mirror images of each other along the middle line of the Bible. This visual design, though a bit more complicated, still conforms to the design pattern of pictorial simile. The paw metonymically stands for animals, the hand metonymically stands for humans, and the symmetric alignment of them invites the viewer to see (at least, partly) the sta-

![FIGURE 2](image-url) An advertisement for the protection of animals (Advertising agency: Lowe, Taiwan). Copyright 1998 by Lowe, Taiwan. Reprinted with permission.
tus of animals as equal to that of humans. The figurative message of the whole picture can then be read as follows: Animals are like us and deserve our solemn promise to treat them with respect. Moreover, the story-like scenario of taking the oath together suggests a further reading that animals are like us and deserve to be protected by law.

THE IMAGE GROUPING HYPOTHESIS

We have argued that pictorial simile is related to, and motivated by, pictorial grouping, and the pictorial devices constrain the ways one reads and makes images. Our analyses of the previous examples show how we can have an integrated account of pictorial simile and grouping, which, we propose, can be generalized into an image grouping hypothesis. An initial formulation of the hypothesis goes as follows:

1. Pictorial grouping: Symmetric image alignment defined in terms of pictorial space, or in terms of the spatial orientation of things in that space, is apt for expressing pictorial grouping when the alignment functions as a connector linking elements from the same category.
2. Pictorial simile: The symmetric image alignment is apt for expressing pictorial simile when it functions as a connector linking two elements from different categories.
3. Pictorial simile with a story-like scenario: A story-like scenario can be introduced into the symmetric image alignment on condition that visual aspects of resemblance between the source and the target are emphasized, keeping the symmetric image alignment salient or making it more pronounced.

Note that the image grouping hypothesis outlined here is not about categorizing or having a thought for simile. Rather, it is about how a category and a simile can be pictorially represented. And the hypothesis shows how pictorial simile is anchored to pictorial grouping, with symmetric image alignment as their common defining feature.

We mentioned that pictorial metaphor has image integration of the source and the target as its defining feature, which makes it stronger than pictorial simile in expressing figurative messages. Let us now see how pictorial metaphor can be related to, or contrasted with, pictorial simile with respect to the image grouping hypothesis. Consider an ad for fresh fruit (see Figure 3), in which a banana and gadgets of a Swiss army knife are integrated into a whole, with the banana as the main body and the gadgets its functioning parts. This visual design suggests that the banana, as the main body of this integrated whole, is the subject being assessed or evaluated, and the gadgets, portrayed as its functioning parts, are descriptions of
the features attributed to the main body. We know that Swiss army knives are famous for their convenience and are affordable for many people across the world. The metaphoric picture, accordingly, invites us to have the following interpretation: The bananas distributed by the company have many different nutrients (functioning parts), you can get one when you want it (convenience), and they are not expensive (affordability).

One might think of the integration as obtained from transforming a symmetric image alignment of the banana and the gadgets into a part–whole relation. This line of thinking, if sustainable, would allow us to see how the visual design of pictorial metaphor could be motivated by that of pictorial simile, which, as has been shown, is anchored to the visual design of pictorial grouping. A design-based, integrated account of pictorial metaphor, pictorial simile, and pictorial grouping thus might be expected. Notice, however, that image integration of pictorial metaphor obliterates the visual design of pictorial grouping. Symmetric image alignment is destroyed. The image grouping hypothesis thus cannot be extended to pictorial metaphor. That sets pictorial metaphor apart from pictorial simile and pictorial grouping from the design-based perspective. The visual design for pictorial metaphor requires a different set of cognitive operations for its effective use in image making. Because the part–whole relations are essentially involved in designing pictorial metaphors, the global features of size, orientation, and position of the objects in the pictures are not
sufficient. Instead, how the parts are depicted and integrated into a whole must play a more important role.

Take again the ad for bananas discussed previously as an example. The part–whole relation in the picture is incongruent with what we may actually experience: The banana is too large to be fitted into the position of the main body of a Swiss army knife, and the gadgets cannot actually be affixed to the banana and still function properly. We know from the context that the picture is an ad for bananas, and it is evident that we cannot dismiss it as merely misrepresenting what it is about. Given this contextual setting, the picture prompts the viewers to come up with a novel interpretation of it. Presumably, the picture invites the viewers to develop a metaphoric viewpoint, taking the bananas as the target, the Swiss army knife as the source, and understanding the former in terms of the latter.

For our purposes, a full-scale analysis of pictorial metaphor is not required. The previous description, we hope, is sufficient to show that visual designs of pictorial metaphor are grounded in our perception of parts and how the parts are (incongruently) integrated into a whole, rather than the grouping effect we found in the visual designs of pictorial simile and pictorial grouping.

Summarizing thus far, Forceville’s (1996) observation that pictorial simile has image juxtaposition as its defining feature is on the right track. It turns out that image juxtaposition is only a special case of a more general design pattern for pictorial simile and pictorial grouping. Forceville’s observation that verbal simile and its relation to verbal metaphor can provide a reference point for understanding pictorial simile is correct in some respects. Yet, it turns out that the cognitive underpinnings of the visual designs of pictorial simile and pictorial metaphor are fundamentally different. Our analysis shows that symmetric image alignment of the source and the target in a pictorial space is apt for expressing a simile in pictorial terms. And it turns out that symmetric image alignment is the principal factor that sets pictorial simile apart from pictorial metaphor and links it to pictorial grouping. The image grouping hypothesis is a generalization over the cases we examined. It accounts for how a category and a simile can be pictorially represented and how pictorial simile is anchored to pictorial grouping with the symmetric image alignment described at an abstract level as their common defining feature.

PICTORIAL OXYMORON

Armed with the image grouping hypothesis, we are ready to see a new type of pictorial trope in practice, namely, pictorial oxymoron. Let us start with an example (see Figure 4), an ad from Benetton’s 1989 ad series, which features a white hand and a black hand handcuffed together, with the copy line “United Colors of Benetton” printed on it. (The copy line had by then been adopted as the official company trademark. All the ads discussed in the following also have the company trademark
printed on each of them and can be found on Campaign Images of United Colors of Benetton: http://www.benetton.com.) The two hands are symmetrically aligned in the pictorial space. The symmetric image alignment invites the viewers to group the two hands together as belonging to the same category. The handcuffs somehow tell a story of how the two hands are linked to each other and thus reinforce the grouping effect of the symmetric image alignment. Viewing the picture in a social setting where race issues are serious, the contrast between black and white stands out prominently and yields a self-opposing tendency that makes the link between them paradoxical. This is an oxymoron in the realm of pictorial representation, yoking together two incompatibles to produce a quick paradox in pictorial terms. Moreover, the concrete image of the handcuffs locking the two hands together presents race issues in an irreverent and provocative way, which prompts the viewers to further think about or explore possible meanings implicated in the image, say, the metaphoric meaning suggested by the bondage of handcuffing, and perhaps an implied commitment to tolerance and a respect for diversity as suggested by the phrase “United Colors.” It is noteworthy that the exploration is open-ended and could yield results that are mutually contradictory.

Another ad, Priest and Nun (Figure 5), also nicely illustrates the open-ended character and the paradoxical effect. This ad shows a priest (dressed in black)
and a nun (dressed in white) kissing. Its visual design conforms to that of pictorial oxymoron. It has succeeded in attracting the attention of the public and arousing strong reaction at least from people to whom religious beliefs are of deep concern. It is provocative and demands a rethinking of, say, what love is and what its relation to customs and religious beliefs can be. Interpreting the image is an open-ended enterprise, and the image is intended to engage or challenge viewers' sensibilities and intelligence and prompt them to come up with novel interpretations. The Italian magazine *L'Espresso* (February 14, 1991) called the picture “A Kiss from God.” Benetton quickly followed suit and dubbed it *A Kiss from God*.

What then makes pictorial oxymoron possible and cognitively effective? By extending the image grouping hypothesis to pictorial oxymoron, we can see pictorial oxymoron as a clever redeployment of pictorial grouping, which can be analyzed as follows:

1. Depict two things in a way that conforms to the visual design of pictorial grouping; this should have the cognitive effect of inviting the viewers to group the two things into the same category.

2. Highlight the dimension along which the two things can be seen as incompatible with each other. (This is what makes pictorial oxymoron different from pictorial...
simile, though both are based on pictorial grouping.) The dimension is well chosen when it touches on serious, social issues that deeply concern or divide the public.

3. Introduce a thematic link into the picture, connecting the two things together in a way that not only reinforces the grouping effect but also prompts the viewers to further explore possible meanings implicated in the image.

4. An oxymoron in pictorial terms can thus be created. It should be accessible to wide numbers of people, because the design is based on the cognitive resources we share. Furthermore, the sense of a quick paradox of combining the incompatibles with a thematic link should make the image striking and to an extent provoking.

A good number of images in Benetton’s advertising campaigns conform to this design pattern: for example, (a) an image of two bakers, one white and the other black, both being covered with white flour; (b) an image of two miners, one white and the other black, both being smeared with soot; (c) an image of a white child with angel-blond curls and a black child with its hair styled in a pair of pointed horns that recalls a devil, but the white child seemingly bearing a devilish smile and the black child looking innocent; and so forth. All these images have two incompatible elements symmetrically placed on the two sides of the vertical axes of the pictorial spaces.

Further elaborations are possible. Consider a fine piece from Benetton ads, dubbed Piano Duet, which features an elegant layout of a black woman’s hands and a white girl’s hands on top of a row of black and white keys of a piano. Layers of symmetric image alignment of black and white and their alternate thematic contrasts call for a free play of our imagination and understanding. This is a subtle variation on the design pattern described previously. The contrast between black and white hands, the contrast between black and white keys, the contrast between black and white hands and black and white keys, and the layers of symmetric image alignments among them all are integrated into a whole that invites the viewers to construct a viewpoint in which explosive and serious race issues are comparatively connected to the orderly layout of black and white keys on a piano.

Other elaborations and variations on the design pattern are still possible. An ad, another one from Benetton, dubbed Soldier (Figure 6), features a black soldier who, machine gun on his shoulder, firmly holds a human thigh bone. The thigh bone and the machine gun form a virtual triangle, along a middle line of which the two are symmetrically aligned. Another ad (see Figure 7) features a black woman breast-feeding a white child, whom she’s holding in her arms. The black woman’s breast, the white child, and the black arms are alternately placed next to each other in an elegant, symmetric manner along the depth dimension.

It would be useful to draw a connection between our analysis of pictorial oxymoron and some of the research on oxymoron in language use. According to Shen (1987), an oxymoron can be construed as based on a number of semantic features with a hierarchical structure in which the lowest semantic feature bears

most of the semantic load in distinguishing the lexical item in question from its neighboring item. In this view, the lexical item “man” is defined by the semantic features “… +animate, +adult, +male,” the lexical item “woman” is defined by the semantic features “… +animate, +adult, –male,” and it is the lowest feature (i.e., “+/–male”) that distinguishes “man” from “woman.” “Man” and “woman” are antonyms in this sense. An oxymoron then can be construed as comprising two terms that are antonyms, in that the only difference between the two terms consists of a change in the “+/–” sign of their lowest feature, all other features being identical (e.g., “a feminine man” and “living death”). This is one type of oxymoron, which Shen called “direct oxymoron.” Another type of oxymoron, which is more complicated and is called “indirect oxymoron,” arises out of the operation of combining two terms whose lowest, distinguishing semantic features are at different hierarchical levels. “The silence whistles” is an indirect oxymoron, for example. The antonym of “silence” is lexically realized by the word “sound.” The second term of this oxymoron is not “sound,” however, but the lexical item “whistles,” which shares the same feature list as does “sound,” with an additional feature “+sharpness” at the level next to “+sound.” “Whistles” is a hyponym of “sound” in this sense. An indirect oxymoron then can be construed as a combination of two terms in which the second term is the hyponym of the first term’s antonym. Shen observed that indirect oxymoron is much more frequently in use than direct oxymoron in modern Israeli poetry and in classical literary dictionaries. Gibbs and Kearney (1994) took up Shen’s distinction between direct and indirect oxymoron. Their study showed that understanding an oxymoron requires that people access relevant world knowledge to constrain their interpretations, and an indirect oxymoron is easier to understand than a direct oxymoron. Indeed, they suggest, a direct oxymoron should be more difficult to understand for people do not generally conceptualize concepts in completely contradictory terms.

Fitting our analysis of pictorial oxymoron into this framework, it is clear that all the pictorial oxymora in the ads for Benetton are direct oxymora. This phenomenon contrasts sharply with Shen’s (1987) observation of oxymora in literary language and raises an interesting question about why there is such a sharp contrast. Here are our suggestions. First, indirect pictorial oxymora are rare, because it is very hard to depict things in which the distinction between levels of conceptual organization can be pictorially realized. An indirect oxymoron, comprising one term that is the hyponym of the other term’s antonym, requires exactly that the distinction be somehow realized. Second, even though it is difficult to interpret a combination of two directly conflicting terms, image-makers manage to overcome the difficulty by introducing a concrete theme to the images, which provides sufficient structures as cues that prompt the viewers to set up the right scheme necessary for interpreting the picture. If this conjecture is on the right track, we may further argue that, not coincidentally, the second term of a
The elaboration introduces a more concrete theme that renders the combination of conflicting terms at the more abstract level intelligible. For example, the second term of “sweet sorrow” is a hyponym of the first term’s antonym; it introduces a more concrete theme, which in turn can be described as a mixture of sweetness and bitterness.

**THE IMAGE GROUPING HYPOTHESIS RECONSIDERED**

Our initial formulation of the image grouping hypothesis covers only pictorial grouping and pictorial simile. We go further and show that the hypothesis can be extended to pictorial oxymoron. This suggests that the hypothesis can be formulated at a more abstract level. Let us call the pictorial components that depict entities at the object level *object-components* and the sites of object-components at an abstract level where the invariant structure of symmetric image alignment is to be found *elaboration sites* (abbreviated as *e-sites*). Now, the image grouping hypothesis in a more general sense can be formulated as follows:

1. Symmetric image alignment with e-sites elaborated by object-components depicting things of the same kind is apt for expressing pictorial grouping.
2. Symmetric image alignment with two e-sites elaborated by two object-components depicting things of different kinds is apt for expressing pictorial simile.
3. Symmetric image alignment with e-sites elaborated by object-components depicting things that can be seen as incompatible with each other is apt for expressing pictorial oxymoron.

As can be seen from this more abstract formulation, symmetric image alignment with e-sites to be elaborated by some object-components is the core of the hypothesis. Furthermore, its purported grouping effect is central to how we may exploit the alignment in reading and making images. Besides pictorial simile and pictorial oxymoron, we surmise, the exploitation of the alignment is potentially open-ended; it all depends on how the alignment may be used in actual usage events. For example, symmetrically aligning a political candidate with known criminals may have the effect of inviting the viewers to see the candidate as hav-

---

2The term *elaboration sites* and its abbreviation are appropriated from cognitive grammar. The following example may roughly illustrate how the concept of e-sites is used in cognitive grammar. The landmark of [UNDER] is an e-site, which is to be elaborated by a more specific component structure. [UNDER THE TABLE] is obtained when the e-site is elaborated by [THE TABLE]. See Langacker, 1987, pp. 304–306.
ing a lot of criminal contacts and connections by prompting the viewers to group the candidate and the criminals into the same category. This technique was used in the 2000 presidential election in Taiwan, especially when issues of the so-called “black money” politics were raised and hotly debated and accusations tossed around. This usage may become entrenched and proliferate in what we do in reading and making images through repetition, variation, and elaboration. In this sense, this formulation of the image grouping hypothesis is not complete, but it sets out a pattern for how we may use the core of the hypothesis to explain new ways of reading and making images whose designs are importantly based on symmetric image alignment.

CONCLUSION

We show here how grouping, simile, and oxymoron can be understood in pictorial terms and how an integrated account of them can be made. The image grouping hypothesis is a design-based, cognitive account that integrates pictorial grouping, pictorial simile, and pictorial oxymoron into a coherent set, with symmetric image alignment as their core design pattern, which is potentially extendable to subtler ways of reading and making images. This design pattern, in tapping the cognitive resources we share, makes pictures based on it easy to access and can be understood almost on first contact. People with normal perceptual capacities and the background knowledge of the themes expressed in the pictures can readily see in them the design-based, figurative messages. The task we attempt to accomplish can be considered to be part of making explicit, from a cognitive perspective, the design-based, structural invariances that are already implicit in our practice of reading and making images.

It is useful to step back and take a broader view of what general guidelines a design-based analysis as set up in this study may provide. We start our investigation by taking verbal simile and its relation to verbal metaphor as a reference point for understanding pictorial simile. We agree that understanding similes and metaphors is fundamentally a conceptual matter. Taking similes and metaphors this way has a liberating effect, in that linguistic expressions are seen only as one among many types of device for expressing similes and metaphors. This makes room for a serious study of pictorial simile and pictorial metaphor, and our analysis shows that pictorial simile differs from pictorial metaphor from a design-based point of view and that there is an intricate interaction between the expressive resources and what may be expressed by them. Furthermore, the design-based analysis can be generalized to pictorial oxymoron. A comparison of pictorial and verbal oxymoron further shows that an investigation into the interrelationships of pictorial and linguistic expressive resources can be a stimulating and rewarding project.
ACKNOWLEDGMENTS

We are very grateful to Charles Forceville for detailed and helpful written comments on earlier drafts. We thank Raymond Gibbs and an anonymous reviewer for their valuable comments and suggestions.

This article is a substantially revised version of two unpublished papers that were joined together. One paper, entitled “Pictorial Simile” and authored by Norman Y. Teng, was presented at the Fourth International Conference on Researching and Applying Metaphor, held in Tunis, Tunisia, in April 2001. Thanks to the discussants (especially Charles Forceville, who chaired the session) for providing critical comments and suggestions on that occasion. We are also indebted to Zouhair Maalej for helpful comments and Susan Gordon for suggestions on ways of improving the writing. The other paper, entitled “Oxymoron in Pictures” and co-authored by Norman Y. Teng and Sewen Sun, was presented at Transculturalism: An International Conference, held in Taipei, Taiwan, in July 2001. We are indebted to Ying-hsiung Chou, Wen-ching Ho, Nikos Papastergiadis, Te-hsing Shan, and Werner Sollors for their helpful comments and suggestions on that occasion.

REFERENCES

Gibbs, R. W., Jr., & Kearney, L. R. (1994). When parting is such sweet sorrow: The comprehension and appreciation of oxymora. Journal of Psycholinguistic Research, 23, 75–89.


