Points of View and Pieces of Time: A Taxonomy of Image Attributes

VAL LARSEN
DAVID LUNA
LAURA A. PERACCHIO*

Visual persuasion is effective to the extent that manipulations of attributes in the ad system elicit desired responses in the human system. This article identifies several ad system attributes—angle of vision, cutting rate, and camera motion—and proposes metrics for specifying the range of values for each attribute. Using the resource-matching hypothesis as a theoretical framework, it discusses human system effects of the ad system attributes and, thus, integrates the semiotics and information-processing approaches to advertising.

Scott (1994) has argued that research on visual persuasion has been fundamentally misdirected because consumer researchers have not engaged the image. Consequently, both academics and practitioners have an incomplete and fragmented understanding of the role of images in advertising. This is unfortunate because ad space devoted to images generally dominates that devoted to words (Childers and Houston 1984; McQuarrie and Mick 2003).

If images are so important, why have consumer researchers not developed a more sophisticated understanding of what they are and how they work? A likely explanation for the neglect is the chasm that has historically separated the two academic cultures (Snow 1959). The best research on the human system that processes and responds to visual stimuli has been done in the scientific tradition of experimental psychology while the best research on the ad system within which images are constructed has been done in the humanistic tradition, preeminently in semiotics. To date, most consumer researchers have approached visual persuasion focusing exclusively on the human system. Very little research has attempted to join psychology with semiotics though the combination has great potential (McQuarrie and Mick 2003).

The failure to combine experimental work on the human system with semiotic work on the ad system has reduced the contribution of both research traditions. Based on an inadequate theory of the image, research on psychological responses to visual stimuli may have limited relevance to advertising practice (Scott 1994). Conversely, in the absence of empirical testing of affective, cognitive, and behavioral responses, the validity of carefully drawn aesthetic and semiological distinctions also remains in question (McQuarrie and Mick 2003).

Fortunately, consumer research has developed both an openness to and familiarity with semiotic and other forms of text analysis while enhancing its already strong competence in experimental psychology. Consumer researchers are, thus, well positioned to bridge the chasm separating Snow's two cultures by yoking together research traditions that can be and, indeed, already have been shown to be powerfully synergistic (McQuarrie and Mick 2003). This article's ultimate goal, therefore, is to reconcile two traditions in consumer research: semiotics and information processing. We join the traditions by developing a taxonomy of image, or icon, attributes and by specifying in a series of research propositions the implications of those attributes for information processing and ad effectiveness. Our taxonomy combines and extends the classification schemes of film theorist Christian Metz (1974) and media aesthetician Herbert Zettl (1990). Thus, this article advances prior work on visual persuasion through the review, extension, and application of a variety of semiotic and psychological concepts.

To frame our analysis of images, we will compare visual communication with verbal communication. Words are formed from sounds whose properties are determined by where and how the stream of air is impeded in the mouth. A sentence is then formed by a sequence of words, whose order is prescribed by syntactic rules. In this article, we

*Val Larsen is associate professor of marketing at James Madison University, Harrisonburg, VA 22807 (larsenwv@jmu.edu). David Luna is associate professor of marketing at Baruch College, City University of New York, New York, NY 10010 (david_luna@baruch.cuny.edu). Laura A. Peracchio is professor of marketing at the University of Wisconsin—Milwaukee, Milwaukee, WI 53201 (lperacch@uwm.edu). The authors thank Joe Sirgy, Julie Ozanne, David Trafimow, Steven Anderson, Scott Hendricks, James Littlefield, the editors, associate editor, and three reviewers for their helpful comments on this research.

suggest that the properties of an image, like those of a sound, are determined by a set of basic visual attributes.

One respect in which our taxonomy extends prior research in semiotics is that we focus on the most basic units of meaning in visual communication. Previous studies have centered on higher-level processes. For example, Scott (1994) works at a relatively high level of abstraction—the level of style, genre, and cultural competence—in developing a critique of traditional advertising research. McQuarrie and Mick (1996, 1999) work at a lower level as they rigorously describe and test the effects of visual rhetorical figures. In this article, we work at a still lower level of abstraction, focusing on describing fundamental, constitutive attributes of the icons others discuss in the context of higher-order meaning-making processes. We develop the micro-analytic foundation Saint-Martin (1990, p. 5) has called for in saying that "No endeavor to describe visual language can [be adequate] if it does not provide a preliminary level of description, analogous to phonology in verbal linguistics, that can explain how primary elements are joined together to form larger units." The basic image attributes we discuss are camera angle, which is a feature of every image, and camera movement and camera cuts, which are features of most moving images. Thus, a rigorous description and theoretically grounded awareness of the role these attributes play in constituting different icons can provide insight into how images are processed, what they mean, and how they persuade.

We begin by describing a recent information-processing model that can be used to understand the human system. We explain some of the model's predictions with regard to key ad response measures. Then, we combine semiotics and information processing by describing an integrated set of icon attributes within the ad system and exploring how those attributes influence the human system. In doing so, we formulate several propositions that relate icon attributes to psychological processes and persuasion.

THE RESOURCE-MATCHING HYPOTHESIS AND ICON ATTRIBUTES

There are two fundamental constructs in the Resource Matching Hypothesis (RMH): available cognitive resources and required cognitive resources. Available resources (AR) refer to mental capacities an individual uses for processing a particular task. Required resources (RR) refer to capacities needed to adequately process that task. The RMH suggests that, for persuasion to be maximized, there must be a balance between the resources available to process an ad stimulus and the resources that such processing requires. If the resources required to process the ad exceed available resources or if the available resources exceed those required to process the ad, a resource imbalance occurs (Anand and Sternthal 1990). Such resource imbalances have detrimental effects on persuasion because when there is a resource imbalance, consumers either fail to process persuasive content (when $RR > AR$) or they generate more idiosyncratic thoughts that negatively influence persuasion (when $RR < AR$) relative to conditions in which there is a balance between available and required resources (Anand and Sternthal 1990).

We propose that the manipulation of icon attributes within the ad system can increase or decrease icon salience, which may contribute to a balance or imbalance between resource requirements and availability, thereby determining the persuasiveness of an advertising image. We follow Fiske and Taylor (1984) in suggesting that an icon is salient when it stands out from its immediate context (e.g., when a camera cut juxtaposes semantically inconsistent icons), from the perceiver's prior experience or expectations (e.g., when an unusual camera angle dramatically alters an object's appearance), or from other foci of attention (e.g., when an image morphs into something different while the background remains constant).

Ad system factors such as an unusual camera angle can lead to higher salience because they are inconsistent with viewers' expectations. Salience, in turn, leads to greater focused attention on the image (Cave 1999; Fiske and Taylor 1984; Nothdurft 1993; Theeuwes 1992; Turatto and Galfano 2001). Attention is defined as the general distribution of mental activity to the tasks being performed by an individual (Moates and Schumacher 1980). Attention is a finite resource that can be allocated in various degrees to processing the primary task (e.g., an ad) or to secondary tasks such as generating unrelated thoughts or engaging in social interaction. The more (less) attention individuals dedicate to a task, the more (less) resources they make available to process it, within presumed resource and capacity constraints (MacInnis and Jaworski 1989; Niepel et al. 1994). Because salience leads to attention and greater attention generally leads to directing higher levels of cognitive resources, or capacity, toward the image, salience is an important determinant of the allocation of available resources.

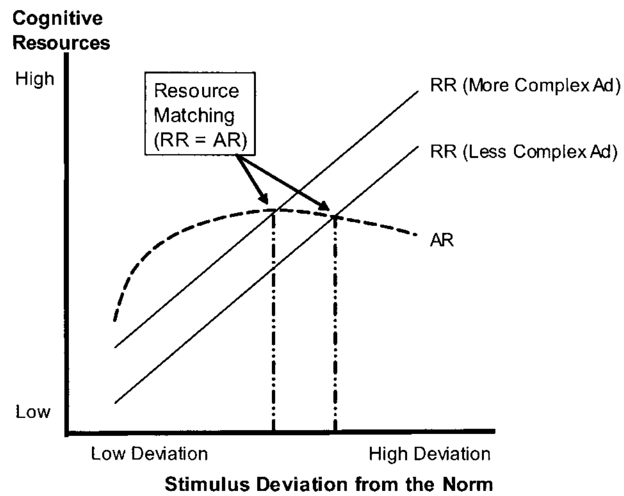
However, another process may result from an increase in salience. An image that deviates from an expected norm (and that is, therefore, relatively salient) may also require an elevated level of resources to be fully processed. Hence, although salient deviations from a norm may direct available resources toward a stimulus, they may also increase the level of resources required to process the stimulus. Potential increases in required resources will generally not have an upper limit. Consider, for example, the effects of increasing the number of camera cuts in a video image. As the number of shifts in point of view steadily increase, the cognitive resources required to process the stimulus should also increase. Since the number of cuts can increase indefinitely, there would seem to be no upper limit to this increase in required resources. But increasing the number of cuts will not produce a steady and unbounded reallocation of available resources. The resources available for stimulus processing are bounded by the resource requirements of secondary tasks and the limit on an individual's cognitive capacity (Johnston and Heinz 1978; Miller 1956).

With respect to the attributes we discuss, salience will generally be the factor that produces changes in the allocation of resources available for image processing. Previous research has shown that stimulus salience leads to viewer attention (Cave 1999; Nothdurft 1993) and that it may influence persuasion positively if the stimulus is not negatively valenced (Berlyne 1958, 1974; Fiske and Taylor 1984). Ultimately, however, when an individual's cognitive capacity limit is reached, the increment in available resources must flatten out. Since the increase in required resources can continue indefinitely, required resources will generally exceed available resources at extreme levels of stimulus deviation from a norm. For example, at very high cutting rates, viewers would experience cognitive overload. And as the image becomes an unintelligible blur of quick cuts, attention may be directed to more tractable secondary stimuli. Thus, available resources may not only flatten out but decline at extreme levels of stimulus deviation from the norm.

Ad or image complexity is a factor that would increase the resources required to process the image. For example, complexity could be increased or reduced by varying the amount of information presented in the ad, providing qualitative versus quantitative arguments, or placing the copy in a print ad in certain formats that are more or less easily processed (Anand and Sternthal 1989; Peracchio and Meyers-Levy 1997). When the ad is relatively undemanding, available resources exceed required resources at the lowest level of stimulus deviation. Thus, the initial spike in available resources increases the gap between required and available resources and may diminish persuasion. Resource matching may be achieved only at moderately high levels of stimulus deviation as available resources approach capacity limits and stop increasing. In a more complex ad, the initial spike in available resources could produce a match between available and required resources at a lower level of stimulus deviation from the norm because the required resources function would be shifted higher (see fig. 1). Having introduced the basic principles of the RMH, we will now develop our taxonomy of icon attributes.

FIGURE 1

INTERACTION OF DEVIATION AND STIMULUS COMPLEXITY



NOTE.—RR is Required Resources; AR is Available Resources. The lines depicted in this graph are for illustration purposes only. Many factors could cause slopes to vary.

THE AD SYSTEM: TOWARD A TAXONOMY OF ICON ATTRIBUTES

The following ad system taxonomy is not exhaustive. We focus on three attributes that are ubiquitous in print and audiovisual advertising, have received research attention, and have a substantial amount of consumer research potential. We link these ad system attributes to the human system using the RMH. Several research propositions will suggest how ad system factors may affect resource matching and, thereby, influence persuasion.

The Verbal Linguistics Analogy

Our taxonomy of icon attributes is grounded in semiotics, a research tradition that provides a solution to a common problem in the development of a taxonomy: classification schemes are inherently rooted in the intuition of similarity, with objects falling into the same class to the extent that they resemble each other on some dimension. But there is no limit to the number of dimensions on which objects may be similar, so taxonomic categories can multiply uncontrollably (Pepper 1970).

Semiotics solves this problem by uncovering hidden foundations on which the meaningfulness of signs rests. In its characteristic mode of analysis, semiotics posits a relationship between a relatively stable underlying system that is defined by a limited number of logically related dimensions and a larger, more changeable set of surface manifestations of that system (Mick 1986; Saint-Martin 1990). For example, although each of the many sounds in a language may seem to be chaotically unique, they are actually formed by the systematic combination of a limited set of underlying phonetic attributes. The quality of a sound is affected, for instance, by the place and manner of articulation. In English, the *d* and *t* are both made by using the tongue to completely occlude the flow of air in the middle of the mouth at the alveolar ridge but the *g* and *k* are formed by occluding the air at the velum in the back of the mouth. A slightly more extended analysis could show that the complete inventory of English sounds (and of sounds in all other languages) is produced by combinations of a very limited set of underlying attributes (Fromkin and Rodman 1978).

Like sounds, sentence syntax may at first glance appear to be randomly varied. But all well-formed English sentences are reducible to one (or some combination) of just 10 basic sentence patterns (Kolln 1990). Surface differences in syntax nevertheless affect meaning. Let us illustrate the importance of syntax with two related sentences: (*A*) The hungry but cautious customer crept into the cafe. (*B*) Into the cafe the customer crept, hungry but cautious. These two sentences do not differ in their basic content. At a deep structural, sentence-prototype level, they are the same sentence (Chomsky 1957). The sentences probably do differ, however, in the cognitive demands they make on a reader and in the amount of attention they focus on sentence elements. Sentence B is harder to process than sentence A because it deviates from normal sentence word order. To understand it, a reader must mentally map it back into the normal, deep-structural prototype, a task that is not necessary for sentence A, which already matches the prototype. At the same time, sentence B directs additional attention to the customer attributes, hunger and caution. These attributes have probably become more salient, being both out of their normal position (the inconspicuous middle of a phrase) and situated at the relatively emphatic end of the sentence.

Deep Structure and Icons

In the discussion that follows, we hope to demonstrate that the semiotic approach that has advanced the study of phonetics can produce similar advances in the study of iconic images. Just as all phonetic sounds can be defined by their place and manner of articulation, so all visual icons are characterized by attributes that can be precisely specified on spatial and temporal dimensions. For these iconic attributes, there is an experiential norm, a value that generally minimizes both cognitive demands and icon salience. This referential norm functions as a stable, deep-structural prototype or standard against which an icon's degrees of deviation on the dimension may be measured. The existence of an iconic referential norm is consistent with psychological studies finding that images are stored in the form that individuals normally encounter them (Tarr and Pinker 1989).

Something is a sign if it stands for some other thing. Signs are classified as icons if they imitate their referent, the thing they signify. Since all icons imitate a referent, the referent can function as a norm or prototype for the signs that represent it. Referents are intrinsically more stable than the iconic signs that signify them, for a given referent may be signified by many different icons that imitate it with widely varying degrees of closeness (Metz 1974). Because it is relatively stable, the referent can function as a deep-structural norm against which degrees of iconic deviation may be measured and classified. The measurement of degrees of deviation from referent norms is possible because icons and their referents share certain dimensions. In the sections that follow, we focus, in particular, on three shared spatial dimensions and one shared temporal dimension.

ICONOLOGY

In this section, differences between icon space and our ordinary experience of referent space will be discussed in subsections that focus on the following icon attributes: angle, cuts, and camera movement. Each subsection begins with a formal, ad system definition of the attribute and its underlying dimensions and norms. An objective metric for gauging degrees of deviation from the dimension norm is also proposed.

Angle of Vision

The Ad System Specification of Angle. In both referent space and icon space, position relative to a focal object can be defined by three coordinates: the *x* (horizontal), *y* (vertical), and *z* (sagittal) axes (Zettl 1990). We might define the middle of our computer screen as the 0:0:0 point. If we move right 6 inches, down 3 ft., and back 8 ft., our relative position would be designated .5:−3:8.

Research on visual recognition suggests that not all points of origin are equal, that the position and orientation of the 0:0:0 point can be determined by a response latency value that may be a function of certain objective properties of the stimulus (Marr 1982). In other words, the 0:0:0 point can be defined as the viewpoint from which the object may be identified most quickly because its most important features are visible. Building on this idea and using the semiotic logic of transformational grammar, Marr (1982) argued that objects exist in visual memory as prototypes. The prototypical view of an object will usually be one that is at right angles to and, thus, features the object's principal axes. For example, the prototypical image of a human being would be the pose usually shown in a child's stick figure drawing where face, trunk, arms, and legs are all fully visible.

Other views are experienced as deviations from that prototypical view. To identify a person seen in profile, we must mentally rotate what we see using various contextual cues until the actual image matches the stored prototype for that person (Marr 1982). Identification occurs most quickly when the actual point of view closely approximates the prototypical view (Tarr and Pinker 1989). Response latency in recognizing a stimulus thus functions as an index of the degree to which a point of view deviates from the prototypical point of view that is stored in memory.

Response latencies may be predictable from objective properties of the stimulus. Thus, for most objects, the latency-minimizing zero point on the *x*-axis may be a position that bisects the object into two equal halves, for symmetry is parsimonious and minimizes cognitive demands (Kreitler and Kreitler 1972). The latency-minimizing zero point on the *y* axis would probably be the normal eye level view of a given object (Meyers-Levy and Peracchio 1992). On the *z*-axis, the zero point would be the closest position that makes visible the object's full length on its *x*- and *y*-axes (Marr 1982). Thus, the value of the prototypical position on the *z*-axis would be a function of the object's size. The larger the object, the further away the prototypical *z*-axis position would have to be.

The Human System Effects of Angle. The fact that response latencies change as point of view changes is a clear indication that the viewing angle affects required resources. The more the point of view deviates from the prototypical view, the higher will be the resources required for processing. The salience of an image should also be directly related to distance from the prototypical point of view since deviations from a norm are salient (Fiske and Taylor 1984). Salience then may result in the allocation of available resources to the image (Niepel et al. 1994). Thus, manipulations of the view angle may influence the degree to which available and required resources match.

> **P1a:** When RR < AR, deviations from a prototypical point of view (the *x*-axis position that bisects the object, the *y*-axis position at normal eye level, and the closest *z*-axis position that fully reveals the *x* and *y* axes) will enhance the persuasiveness of the image up to the point where RR = AR. Greater deviations from the prototypical point of view will result in a resource imbalance, RR > AR, and the persuasiveness of the image will be reduced.

Although there may be only one prototypical angle that minimizes required resources and available resources, there will normally be many alternative ways to achieve a given change in the level of required resources or available resources through the manipulation of angle. Further, an important moderator of the effect of increasing or reducing available and required resources through degree of angle deviation will be the overall complexity of the ad stimulus.

In ordinary experience, position relative to an object changes more frequently on the *x*- and *z*-axes than on the *y*-axis, so *y*-axis deviations from the norm may be especially salient. And research indeed indicates that *y*-axis changes in the point of view strongly influence affective responses. For example, Meyers-Levy and Peracchio (1992) created three versions of an ad that differed only in the *y*-axis angle (low, eye level, and high) from which the product was viewed. Attitude toward the advertised object was highest in the low angle condition, lower at eye level, and lowest in the high angle condition.

That study also found that the *y*-axis effect was strongest for respondents low in need for cognition, the group most likely to be low in available resources. Thus, it seems that the change in affect was mediated by resource matching. Affective responses may have been more extreme as angles diverged from the prototypical view because the salience of the execution directed available resources toward the stimulus. However, since the high and low angle ads were equally divergent from the prototype but produced opposite effects, a second factor must also be posited. Messaris (1992) supplies that additional factor in arguing that low and high angle shots have different effects because they resonate with different analogous real-life experiences: looking up at people (parents) who are larger and more powerful versus looking down at people (children) who are less powerful. The fact that low angle shots produce more positive affective responses than high angle shots may reflect the importance of social-psychological influences on ad responses. Messaris's hypothesis is consistent with the fact that the effects of low angle shots appear to be moderated by degree of trust in the depicted object. Viewing a trusted person from a low angle enhances positive affect whereas the same view of a distrusted person can create fear and a strong sense of threat (McCain, Chilberg, and Wakshlag 1977; Messaris 1992).

As the point of view of an object changes, so does its apparent height, length, and width. These changes in appearance are called foreshortening. Foreshortening is an important moderator of the effects of changes in camera position on required and available resources. The degree of foreshortening is a function of the relative magnitude of the axis on which a change in camera position occurs. Consider, for example, a typical bottle of salad dressing. If a point at the center of the front of the package is defined as 0:0:0 and this point is given an orientation at right angles to the surface of the package, 0:0:2 might be the approximate prototypical view of the package. The package is easy to recognize from this point of view because we are positioned at right angles to its two principal axes, the vertical *y*, which is the longest, and the horizontal *x*, which is the second longest. These axes are particularly important because they are the main features of the object's shape.

If the camera were repositioned to 2:0:0 (moved to a position at a right angle to the much shorter *z*-axis and on a line with the *x*-axis at a distance of 2 ft.), the package would be harder to identify. Its height would still be apparent and its depth would be more apparent, but its breadth would be hidden by foreshortening of the *x*-axis. Because one of its two principal axes would no longer be as visible, response latency in identifying the package should increase. And response latency should increase still more if the camera were positioned directly above the package at 0:2:0. In this position, the breadth and depth of the package would be apparent, but its longest axis, the *y*, would be foreshortened because the camera would be on a line with it.

Although the length of an axis is an important variable, the effects of repositioning the camera on any axis will be heavily moderated by the camera's position on the other two axes. Thus, if the camera is close to the focal object on any two axes (*x* and *y*, *x* and *z*, *y* and *z*), movement on the third axis will dramatically change the object's appearance through foreshortening. For example, the appearance of a bottle of salad dressing would be much more dramatically changed by moving the camera from position 0:0:.5 to 0:.5:.5 than it would be by moving the camera from 0:0:8 to 0:.5:8.

> **P1b:** The position of the camera on any two axes moderates the effects of camera repositioning on the third axis. If the camera is close to the focal object on two axes, even small changes in position on the third axis may dramatically change object appearance and have a large effect on required and available resources. If the camera is far from the focal object on either or both of the first two axes, repositioning on the third will have a relatively small effect on object appearance and required and available resources.

Icon Cuts

The Ad System Specification of Cuts. A camera cut occurs when there is an abrupt change from one shot to another. The values of this attribute may be formally defined by the presence or absence of spatial and temporal continuity. Most cuts occur when the viewpoint shifts from one set of angle coordinates, $x{:}y{:}z_1$, to a nonadjacent set of coordinates, $x{:}y{:}z_2$, without passing through intervening space. Taking referent space as the standard of comparison, it is apparent that these camera cuts create spatial ellipses or gaps in icon space. But cuts may also be created by temporal ellipses. Temporal cuts (which are frequently used in editing news interviews and are often called jump cuts) occur when the angle remains unchanged but the time shifts from time 1, t_1, to some nonadjacent time t_2 (e.g., by cutting out part of what the person said). Thus, with a temporal cut, there is no gap in icon space, just a gap in icon time, a temporal ellipsis, that makes icon time (the time used to depict the action in the icon) shorter than referent time (the actual temporal duration of the depicted action).

Figure 2 illustrates the set of possible cutting strategies. When there is neither a spatial nor a temporal ellipsis, an uncut shot occurs from angle $x{:}y{:}z$. More typical of advertising in the past, this execution is still used in some talking head commercials in which we view a spokesperson from a single point of view throughout the entire duration of the ad. A spatial cut occurs when there is a spatial but no temporal ellipsis (i.e., the shot shifts from $x{:}y{:}z_1$ at t_1 to $x{:}y{:}z_2$ at t_1). For example, we see a Nike-clad Michael Jordan rising toward the basket from one angle, then suddenly see him complete the dunk from another point of view. As mentioned above, a temporal cut occurs when there is a temporal but no spatial ellipsis (i.e., the shot shifts from $x{:}y{:}z_1$ at time t_1 to $x{:}y{:}z_1$ at time t_2). For example, against an unchanging background, a person is shown drinking a Coke as a young, then suddenly middle-aged, then suddenly old woman. A spatiotemporal cut occurs when there are both spatial and temporal ellipses (i.e., the shot shifts from $x{:}y{:}z_1$ at time t_1 to $x{:}y{:}z_2$ at time t_2). For example, a travel ad shows a tourist at the Sydney Opera House, then cuts to the Grand Canyon, having shifted in both space and time.

When a temporal or spatiotemporal camera cut occurs with its attendant temporal ellipsis, the relationship between icon time and referent time becomes, at least momentarily, ambiguous. The amount of icon time that passes is zero by definition, but referent time may take any positive or negative value. Because icon time is always zero, a single numerical parameter, a value measuring the duration of the deleted referent time, defines this aspect of the cut. This time value would be very large for the temporal cut ad mentioned above in which a woman went from being young to being middle aged during the cut. It would be much smaller for the travel ad. It would be zero for the Nike ad with the spatial cut since no time elapses between shots in the camera cut.

When a spatial or spatiotemporal camera cut occurs with its attendant spatial ellipsis, the relationship between icon space and referent space likewise becomes momentarily ambiguous. Distance actually traversed in the ad is zero but distance traveled in referent space may take any value between zero and infinity. The value of this space parameter would be small in the Nike ad where the shot shifts only to another vantage point around the basket. It would be much larger in the travel ad and zero in the temporal cut ad where time changes but the camera's relationship to the person depicted and the background does not.

FIGURE 2

TYPES OF ICON CUTS

Temporal Ellipsis	Spatial Ellipsis: Absent	Spatial Ellipsis: Present
Absent	**Uncut** $x{:}y{:}z$ --Least cognitively demanding --Least salient	**Spatial Cut** From: $x{:}y{:}z_1$ at t_1 To: $x{:}y{:}z_2$ at t_1 --Moderately demanding --Moderately salient
Present	**Temporal Cut** From: $x{:}y{:}z_1$ at t_1 To: $x{:}y{:}z_1$ at t_2 --Moderately demanding --Moderately salient	**Spatiotemporal Cut** From: $x{:}y{:}z_1$ at t_1 To: $x{:}y{:}z_2$ at t_2 --Most cognitively demanding --Most salient

Camera cuts are ubiquitous in television advertising, but the concept is applicable to print advertising as well. A camera cut exists in print whenever an ad contains more than one picture. Thus, the travel ad could be executed in print with separate still shots of the opera house and the Grand Canyon, a print example of a spatiotemporal cut.

The Human System Effects of Cuts. The discrete shots that cuts create are fragments of space/time torn from their referent space/referent time context. To understand the relationships among the different shots, the viewer must mentally map them back into referent space/referent time. Research indicates that this mapping task elevates required resources. Thus, when a cut occurs, reaction times on a secondary task slow, suggesting that processing the cut is cognitively demanding (Geiger and Reeves 1993). But cuts also evoke an orienting response, a reflexive allocation of additional attention (Lang 1990; Lang et al. 1993; Reeves et al. 1985). In other words, cuts both require additional processing resources and direct available resources to the stimulus. We can conclude that, as the number of cuts in an image increases, RR should increase and AR should first increase, then level off, and finally may even decrease.

P2a: When RR < AR, increases from zero in the number of cuts in an image will enhance the persuasiveness of the image up to the point where RR = AR. Additional cuts will result in a resource imbalance, RR > AR, and the persuasiveness of the image will be reduced.

The findings of research on cutting rates have been consistent with RMH predictions. For example, Heft and Blondal (1987) found that cutting rate influenced the strength of affective responses. Subjects were shown negative or positive film stimuli that contained either a large or small number of cuts. Positive responses to the positive stimulus and negative responses to the negative stimulus were both more extreme for the film with more cuts than for the film with fewer cuts. This intensification of the intended affective response may have resulted from resource matching that minimized counterargumentation and idiosyncratic thoughts.

Kraft (1986) also found that positive affect was higher for films with cuts than for those without them. But this main effect was moderated by the film's underlying level of complexity. Using films with lengths similar to commercials (35–39 sec.), Kraft found that cuts increased positive affect in simpler films showing just one activity (required resources presumably less than available resources) more than in more complex films showing four activities (required resources presumably more equal to available resources). In other words, as the RMH predicts, the less complex the film, the more likely it is to benefit from an increment in required resources through the addition of camera cuts (see fig. 1).

The effects of a cut may also be moderated by the degree of spatial and/or temporal displacement associated with the cut. When the change in spatial or temporal context from one shot to another is dramatic, image salience and required resources should substantially increase. When the shift is small, the effects on salience and required resources should be small. Thus, spatiotemporal cuts, which involve both spatial and temporal displacement, should produce larger changes in image salience and required resources than the one-factor temporal and spatial camera cuts. And salience and required resources should increase as the duration parameter for the temporal ellipsis deviates from zero toward positive or negative infinity and as the parameter measuring the magnitude of the displacement in the spatial ellipsis deviates from zero toward infinity.

> **P2b:** When changes during a cut occur in both spatial and temporal position and/or when the magnitude of the temporal or spatial displacements is large, the effects on required and available resources will be direct and proportional.

Camera Movement.

The Ad System Specification of Movement. It is more challenging to describe a dynamic system (a moving camera) than a static system (a stationary camera), but using the spatial coordinates *x*, *y*, and *z*, a temporal value, *t*, and a rate value, *r*, a precise description of camera movement is possible. To describe the movement of a camera from position $x{:}y{:}z_1$ to $x{:}y{:}z_2$, *x*, *y*, and *z* values must each be expressed as a function of *t* and *r* in the formula $x = rt + k$, where *x* is the terminal position of the camera on the *x*-axis, *r* is the rate of movement, *t* is the time spent moving, and *k* is the initial position on the *x*-axis. This calculation is nothing but the familiar formula distance = rate × time.

This *x* (and *y* and *z*) $= rt + k$ notation can be used to describe the full range of camera movements, which are generally referred to as the tongue or truck (*x*-axis), pedestal or boom (*y*-axis), dolly (*z*-axis), and crab or arc (*x* and *z* axes) shots. To take a simple example, a dolly shot in which the camera moves from 1:2:5 to 1:2:3 in 2 sec. could be denoted $x = 1$, $y = 2$, $z = -1t + 5$, where $0 \le t \le 2$. A more complicated 4-sec. arc shot in which the camera moved around the object from 1:2:4 to 5:2:−4 could be denoted $x = 1t + 1$, $y = 2$, $z = -.5t^2 + 4$, where $0 \le t \le 4$.

The Human System Effects of Movement. For any given scene, a moving camera will confront viewers with more information than a still camera because movement changes perspectives on a focal object and its background. As the pace of camera movement increases (i.e., as the *t* value falls), the flow of information and required resources generally increase. At the same time, movement may result in the allocation of available resources to the stimulus, which would facilitate resource matching (Kipper 1986).

> **P3a:** When RR < AR, camera movement will enhance the persuasiveness of the image up to the point where RR = AR. Further deviations from stasis will result in a resource imbalance, RR > AR, and the persuasiveness of the image will be reduced.

Like cuts, camera movement may draw attention by constantly shifting the viewer to a new point of view (Kipper 1986). But movement does not require viewers to infer the connections between one point of view and another. The continuous shot supplies all linkages between the different perspectives. So a moving camera may produce an orienting increment in attention that is countervailed by a smaller increase in required resources than is typical for cuts. Movement may, therefore, be especially useful for allocating available resources when the stimulus is already complex and would be made too cognitively demanding by the addition of cuts. And if the focus is on movement versus stasis, movement will facilitate resource matching with less complex, stasis with more complex stimuli.

In past research, movement has been treated as an undifferentiated manipulation that is simply present or absent (Kipper 1986; Reeves et al. 1985). But it is likely that the specific vector, the trajectory of the motion, will moderate its effects. For instance, movement on the *y*-axis may be more salient than movement on the *x*- and *z*-axes because vertical movement relative to an object (along the *y*-axis) is less common in ordinary experience than movement toward an object (along the *z*-axis) or parallel to an object (along the *x*-axis).

Attribute Interactions

The icon attributes defined in this study are likely to interact with each other and with individual difference and contextual factors, affecting persuasion differently across different levels of the interacting variables. For example, angle and camera movement are likely to interact. When a camera moves on a given vector, it passes through a series of camera angles. The sequence and degree of prototypicality of those angles is likely to moderate the effects of the camera movement on image salience, level of required resources, resource matching, and persuasion.

We might, for instance, anticipate different effects from an arc shot of a salad dressing package that begins at the deviant angle 0:2:0 (looking down on the package from above) and ends at the angle prototype 0:0:2 (an eye level view 2 ft. from the package) than from the same arc shot if it begins at the 0:0:2 prototype and ends above the package at the deviant 0:2:0 angle. Though the two shots would contain all the same camera angles, the differing angle sequences are likely to produce different cognitive effects. If the shot begins at the angle prototype, it may cue a gestalt into which the deviant overhead view may be easily mapped. The AR and RR would then increase gradually as the camera moves toward the deviant angle. But if the shot begins at the deviant angle, the viewer may struggle to make sense of the image until the camera moves to a position where the main axes that define the package gestalt become more apparent. Thus, at the outset, required resources may be higher for the shot if it begins at the deviant angle than if it begins with the prototype. Attention and available resources may also be temporarily high as viewers try to interpret (Niepel et al. 1994) or disambiguate (Peracchio and Meyers-Levy 1994) the deviant stimulus during movement toward the prototypical angle. A shot beginning with a prototypical view would miss this attention-grabbing opportunity.

The attributes defined in this study are very likely to interact with the semantic properties of the focal objects in the image. For example, the effects of cuts on required and available resources are likely to be very different when the focal objects are clearly linked by a shared semantic theme than when they are not. Thus, the required resources associated with a cut execution that showed an SUV in the Australian outback, then on the Alaskan tundra, then on an African savannah would probably be lower than those associated with an execution that showed the SUV in the outback, then in a parking garage, then inside a car repair shop. The first set of shots would be unified by an obviously shared semantic theme, wilderness. The second set of shots is not connected by any obvious shared themes.

DISCUSSION

Contributions to Theory

Our fundamental purpose in this article was to yoke visual semiotics to information processing because this combination of research traditions has special promise for filling a very important gap in persuasion research. Future advances in the understanding of visual persuasion will be facilitated by the development of an integrated visual rhetoric—a theoretical framework that combines detailed specifications of ad system attributes with equally well developed human system theories on how changes in visual attributes influence cognition and affect. Since the human system component of the visual rhetorical framework is currently much more fully developed than the ad system component, this article concentrates on the development of ad system theory and, in particular, on the development of theory at the most fundamental level of visual semiotic analysis. To complete our basic task of yoking together research traditions to form an integrated visual rhetoric, we also combine the ad system analysis of each attribute with a human system analysis of its effects on resource matching and persuasion.

Semiotics is often called the science of signs. The label is appropriate because semiotics, like the physical sciences, seeks to identify fundamental attributes whose nomothetic combination can parsimoniously explain the surface variability of observed reality. Thus, a successful semiotic analysis reduces apparent variety by showing how differences between signs are reducible to systematic transformations on one or several underlying semiotic dimensions. A semiotic analysis has value for social scientists in consumer research because it permits researchers to create stimulus manipulations that systematically sample a range of possible attribute values. Having systematically manipulated the properties of the sign, a researcher may empirically test generalizable and nomologically integrated theories on the effects that signs have on dependent variables of interest and on the psychological processes that produce those effects. In this article, the linkage between the ad system and the human system is crystallized in each section by research propositions. These propositions may give direction to future researchers, both because they identify key ad system attributes and because they specify related but, nevertheless, distinctive resource-matching consequences when various attribute values change.

One contribution of this article is that our analysis of spatiotemporal orientations and positions identifies attributes that exist at the lowest level of a layered sign system, a level analogous to phonetics in linguistics. The overall systems of verbal and pictorial communication are composed of various related but largely autonomous subsystems. In linguistics, the hierarchy of subsystems begins with phonetics and rises through syntax, morphology, and semantics, to stylistics and rhetoric and other still broader disciplines. Similar subdisciplines would exist in the semiotic iconology that would be part of a fully developed visual rhetoric. The disciplines at the bottom of the hierarchy are as important as higher level disciplines, for basic-level systems have properties of simplicity and physicality that can play an important role in the creation and explication of higher-order meanings.

Because they are comparatively simple, low-level semiotic systems may be fully specified and can then increase the precision and completeness of higher-order analyses. For example, advances in phonetics have greatly increased the rigor of the analysis of sound effects in poetic discourse.

The complete specification of object orientations may likewise make it possible to be more systematic in our analysis of higher-order effects such as the rhetorical figuration that McQuarrie and Mick (2003) have shown to be an important component of advertising.

Directions for Future Research

Visual Semantics. The greatest gains in the understanding of visual persuasion are likely to flow from the development of an adequate visual semantics, a task outside the scope of this article and not yet completed even in linguistics. Although the semantic system has not been fully described, its essential structure is clear. The meaning of words and images are defined at a deep structural level by a bundle of fundamental semantic attributes called sememes (Chafe 1970). Objects are semantically consistent or inconsistent to the degree that their sememes are compatible. A potential contribution of our taxonomy to future work on visual persuasion is that it defines a set of attributes that may prove to be semantically important.

Viewer Attributions. Although this article focused on the direct effects of our attributes on cognitive processing, these attributes may also indirectly affect attention and resource matching by influencing viewer attributions. Viewers may know or sense that their ability to process the stimulus thoroughly is greater for ads with prototypical angles than for those with odd angles and greater for ads with few or no cuts than for those with many cuts. Knowing too that the advertiser's selection of ad features is not random, they may attribute a rhetorical purpose to the creators of the ad based on the executions that have been selected. Specifically, they may assume that the discursive content of ads with unusual angles and many cuts will be limited because they know (and know that advertisers know) that it is hard to process discursive content when it is in those formats. If viewers allocate attention accordingly, it would sharpen the negative effects of combining a simple stimulus with a prototypical angle and no cuts or a complex stimulus with many unusual camera angles and lots of cuts. The attribution could cause viewers to devote extra attention to the undemanding uncut ad and thus exacerbate the mismatch between required and available resources. Conversely, viewers may likewise exacerbate the mismatch by devoting less attention to complex ads that are also heavily cut and full of challenging angles.

We suggested earlier that required resources rise as cutting rates rise because cuts entail the effortful mapping of icon space back onto referent space. However, nonrandom cuts may reduce required resources if the effects of higher-order attributional processes are factored in. While they fragment referent space, cuts also allow the creators of an ad to feature causal relationships. For example, a person might be shown eating spicy food, then immediately be shown suffering from a stomachache. The causal relationship between the eating and the stomachache could be sharply illuminated by the cut if viewers accurately impute a rhetorical intention—the signification of causation—to the juxtaposition of the shots. The causation might be unclear if the hour that intervenes between cause and effect were not excised through an ellipsis. The countervailing or interacting effects of spatiotemporal attributes on required resources and attributions should be explored in future research.

To the extent that the attributions of those who view ads and those who create them are consistent with each other, a visual language exists. It is therefore important to identify montage elements that can produce predictable attributions. But although linguistics has been our model for what must be and can be done to create a sound visual rhetoric, we should note an important point of divergence between linguistics and the semiology of iconic representation. When Eisenstein (1947) and other early developers of film theory and practice first discussed the norms of film montage, they believed that they were describing a sign system that was just as determinate as language. In other words, they anticipated that, like certain words, certain cuts and angles and other montage elements would produce specific and predictable responses in an audience. That expectation has not been borne out. The meanings of images have proven to be more underdetermined than those of words.

As a consequence of this important difference between linguistics and iconology, we should expect, as we seek to link the semiology of the ad system with responses in the human system, that there will be more variance in the responses to icons than there is in responses to words. Because they have a close relationship with experience that is unmediated by signs, icons may have special persuasive force. But lacking mechanisms for social agreement on their significance, the direction of that force may prove to be comparatively variable and yet clearly constrained by the formal properties of the icon.

[*David Glen Mick and Dawn Iacobucci served as editors and Wayne Hoyer served as associate editor for this article.*]

REFERENCES

Anand, Punam and Brian Sternthal (1989), "Strategies for Designing Persuasive Messages: Deductions from the Resource Matching Hypothesis," in *Cognitive and Affective Responses to Advertising*, ed. Patricia Cafferata and Alice Tybout, Lexington, MA: Lexington Books, 135–59.

——— (1990), "Ease of Message Processing as a Moderator of Repetition Effects in Advertising," *Journal of Marketing Research*, 27 (August), 345–53.

Berlyne, Daniel E. (1958), "The Influence of Complexity and Novelty in Visual Figures on Orienting Responses," *Journal of Experimental Psychology*, 55 (3), 289–96.

——— (1974), "Novelty, Complexity, and Interestingness," in *Studies in the New Experimental Aesthetics: Steps toward an Objective Psychology of Aesthetic Appreciation*, ed. Daniel E. Berlyne, New York: Wiley & Sons.

Cave, Kyle R. (1999), "The FeatureGate Model of Visual Selection," *Psychological Research*, 62 (2–3), 182–94.

Chafe, Wallace L. (1970), *Meaning and the Structure of Language*, Chicago: University of Chicago Press.

Childers, Terry L. and Michael J. Houston (1984), "Conditions for a Picture-Superiority Effect on Consumer Memory," *Journal of Consumer Research*, 11 (September), 643–54.

Chomsky, Noam (1957), *Syntactic Structures*, The Hague: Mouton.

Eisenstein, Sergei M. (1947), *The Film Sense*, New York: Harcourt & Brace.

Fiske, Susan T. and Shelley Taylor (1984), *Social Cognition*, New York: McGraw-Hill.

Fromkin, Victoria and Robert Rodman (1978), *An Introduction to Language*, New York: Holt, Rinehart, & Winston.

Geiger, Seth and Byron Reeves (1993), "The Effects of Scene Changes and Semantic Relatedness on Attention to Television," *Communication Research*, 20 (2), 155–75.

Heft, Harry and Ragnar Blondal (1987), "The Influence of Cutting Rate on the Evaluation of Affective Content of Film," *Empirical Studies of the Arts*, 5 (1), 1–14.

Johnston, William A. and Steven P. Heinz (1978), "Flexibility and Capacity Demands of Attention," *Journal of Experimental Psychology: General*, 107 (4), 420–35.

Kipper, Philip (1986), "Television Camera Movement as a Source of Perceptual Information," *Journal of Broadcasting and Electronic Media*, 30 (3), 295–307.

Kolln, Martha (1990), *Understanding English Grammar*, New York: Macmillan.

Kraft, Robert N. (1986), "The Role of Cutting in the Evaluation and Retention of Film," *Journal of Experimental Psychology: Learning, Memory and Cognition*, 12 (1), 155–63.

Kreitler, Hans and Shulamith Kreitler (1972), *Psychology of the Arts*, Durham, NC: Duke University Press.

Lang, Annie (1990), "Involuntary Attention and Physiological Arousal Evoked by Structural Features and Emotional Content in TV Commercials," *Communication Research*, 17 (3), 275–99.

Lang, Annie, Seth Geiger, Melody Strickwerda, and Janine Sumner (1993), "The Effects of Related and Unrelated Cuts on Viewers' Memory for Television: A Limited Capacity Theory of Television Viewing," *Communication Research*, 20 (1), 4–29.

MacInnis, Deborah J. and Bernard J. Jaworski (1989), "Information Processing from Advertisements: Toward an Integrative Framework," *Journal of Marketing*, 53 (October), 1–23.

Marr, David (1982), *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*, San Francisco: W. H. Freeman.

McCain, Thomas A., Joseph Chilberg, and Jacob Wakshlag (1977), "The Effect of Camera Angle on Source Credibility and Attraction," *Journal of Broadcasting*, 21 (1), 35–46.

McQuarrie, Edward F. and David Glen Mick (1996), "Figures of Rhetoric in Advertising Language," *Journal of Consumer Research*, 22 (March), 424–38.

——— (1999), "Visual Rhetoric in Advertising: Text-Interpretive, Experimental, and Reader-Response Analyses," *Journal of Consumer Research*, 26 (June), 37–54.

——— (2003), "The Contribution of Semiotic and Rhetorical Perspectives to the Explanation of Visual Persuasion in Advertising," in *Persuasive Imagery: A Consumer Response Perspective*, ed. Linda Scott and Rajeev Batra, Mahwah, NJ: Erlbaum, 191–222.

Messaris, Paul (1992), "Visual 'Manipulation': Visual Means of Affecting Responses to Images," *Communication*, 13 (2), 181–95.

Metz, Christian (1974), *Film Language: A Semiotics of the Cinema*, Chicago: University of Chicago Press.

Meyers-Levy, Joan and Laura A. Peracchio (1992), "Getting an Angle in Advertising: The Effect of Camera Angle on Product Evaluation," *Journal of Marketing Research*, 29 (November), 454–61.

Mick, David Glen (1986), "Consumer Research and Semiotics: Exploring the Morphology of Signs, Symbols, and Significance," *Journal of Consumer Research*, 13 (September), 196–213.

Miller, George A. (1956), "The Magic Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information," *Psychological Review*, 63 (1), 81–93.

Moates, Danny R. and Gary M. Schumacher (1980), *An Introduction to Cognitive Psychology*, Belmont, CA: Wadsworth.

Niepel, Michael, Udo Rudolph, Achim Shuetzwohl, and Wulf Uwe Meyer (1994), "Temporal Characteristics of the Surprise Reaction Induced by Schema-Discrepant Visual and Auditory Events," *Cognition and Emotion*, 8 (5), 433–52.

Nothdurft, Hans C. (1993), "Saliency Effects across Dimensions in Visual Search," *Vision Research*, 33 (5–6), 839–44

Pepper, Stephen C. (1970), *World Hypotheses*, Berkeley: University of California Press.

Peracchio, Laura A. and Joan Meyers-Levy (1994), "How Ambiguous Cropped Objects in Ad Photos Can Affect Product Evaluations," *Journal of Consumer Research*, 21 (June), 190–205.

——— (1997), "Evaluating Persuasion-Enhancing Techniques from a Resource-Matching Perspective," *Journal of Consumer Research*, 14 (September), 178–91.

Reeves, Byron, Esther Thorson, Michael Rothschild, Daniel McDonald, Judith Hirsch, and R. Goldstein (1985), "Attention to Television: Intrastimulus Effects of Movement and Scene Changes on Alpha Variation over Time," *International Journal of Neuroscience*, 27 (3), 241–255.

Saint-Martin, Ferdinand (1990), *The Semiotics of Visual Language*, Bloomington: Indiana University Press.

Scott, Linda M. (1994), "Images in Advertising: The Need for a Theory of Visual Rhetoric," *Journal of Consumer Research*, 21 (September), 252–73.

Snow, Charles P. (1959), *The Two Cultures and the Scientific Revolution*, New York: Cambridge University Press.

Tarr, Michael J. and Steven Pinker (1989), "Mental Rotation and Orientation-Dependence in Shape Recognition," *Cognitive Psychology*, 21 (2), 233–82.

Theeuwes, Jan (1992), "Perceptual Selectivity for Color and Form," *Perception and Psychophysics*, 51 (6), 599–606.

Turatto, Massimo and Giovanni Galfano (2001), "Attentional Capture by Color without Any Relevant Attentional Set," *Perception and Psychophysics*, 63 (2), 286–97.

Zettl, Herbert (1990), *Applied Media Aesthetics*, Belmont, CA: Wadsworth.