

Quick Guide to Reading Images – Prof. Paul Miers|1

Introduction

Reading Images extends a linguistic paradigm developed by [Michael Halliday](#) to the semiotics of images:

Grammar goes beyond formal rules of correctness. It is a means of representing patterns of experience It enables human beings to build a mental picture of reality, to make sense of their experience of what goes on around them and inside them. (Halliday, [Introduction to Functional Grammar](#))

The key assumption is that natural language and many human sensory modalities share a common grammar. The five core chapters of the book are organized around the three components that Halliday posits for natural language grammar:

		Metafunctions		
linguistic		Ideational Processes & relations	Interpersonal tense / aspect / mood	Textual topic / focus
visual	narrative	Participants, vectors & circumstances		
	conceptual	classes & attributes		
	modality		coding orientations	
	interactions		spaces & points of view	
	composition			presentation & context

This framework assumes that grammars are a *social resource* used by members of a group to communicate with each other (p. 13):

- **“Communication requires that participants make their messages maximally understandable in a particular context.”**
- **“Representation requires that sign-makers choose forms for the expression of what they have in mind, forms which they see as most apt and plausible in the given context.”**



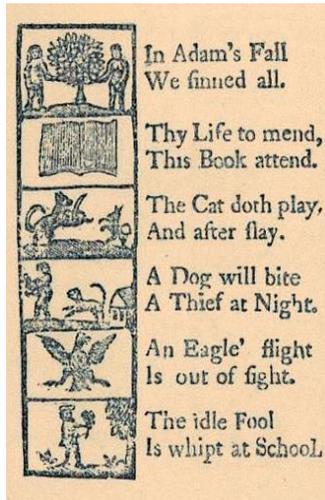
Kress and van Leeuwen are particularly concerned with culture specific aspects of “literacy” and how children learn the various modal grammars used in their social environment. They assume that signs are never completely arbitrary and always have a motivation tied to their context and purpose.

“Little Red Riding Hood” in [simplified Chinese](#) without [pinyin](#) or English

Quick Guide to Reading Images – Prof. Paul Miers|2

The Semiotic Landscape

"Landscape" is a concept borrowed from [evolutionary biology](#) which refers to all the possible states that a biological entity can occupy with respect to its local ecology. It assumes that semiotic entities vary and evolve across space and time.



[The New-England Primer](#),
*Improved; or, an Easy and
Pleasant Guide to the Art of
Reading. Adorned with cuts. To
which is added, the Catechism.*
Boston: Printed for Hall & Hiller,
1810.

The place of visual communication in a given society can only be understood in the context of, on the one hand, the range of forms or modes of public communication available in that society and, on the other hand, their uses and valuations. We refer to this as "the semiotic landscape." The metaphor is worth exploring a little, as is its etymology. The features of a landscape (a field, a wood, a clump of tress, a house, a group of buildings) only make sense in the context of their whole environment and of the history of its development ("wasteland" has meaning only in that context, as has "field" or "track;" "village" has meaning only as a group of building that is part of a history of ways of working the land). In the same way, particular features and modes of communication should be seen in the history of their development, and in the environment of all the other modes of communication which surround them. The use of the visual mode is not the same now as it was even fifty years ago in Western societies; it is not the same from one society to another; and it is not the same from one social group or institution to another (p. 35)

Kress and van Leeuwen assume language *has always* involved both auditory and visual signifiers (i.e. sounds and marks) They also assume our culture is undergoing a transition from an old to a new visual literacy. In the old literacy, speech as writing dominated all other forms of visual communication, and children in school were expected to transition from picture books to print literacy. The new literacy

is multimodal, composed of a mix of text, images, and sounds and often seen as a threat to the status of the old literacy. See, for example, Maryland's [Teaching and Learning: English Language Arts](#).

The key assumptions which follow from this evolutionary paradigm for social semiotics are (p. 41):

- a) Human societies use a variety of modes.
- b) Each mode has different potentials for forming subjectivities.
- c) Each mode has specific social valuation in specific social contexts.
- d) Different potentials for meaning may imply different potentials for the formation of subjectivities.
- e) Individuals use a range of representational modes, and thus have available a range of means of meaning making, each affecting the formation of their subjectivity.
- f) The different modes are not held discretely, separately, as strongly bounded autonomous domains in the brain or as autonomous communicational resources in culture; nor or they deployed discretely either in representation or communication.
- g) Affective aspects of human beings and practices are not discrete from other cognitive activity specific contexts, and therefore never separate or absent from representational or communicative behavior.
- h) Each mode has a continuously evolving history in which its semantic reach can expand or contract or move into different areas of social use as a result of the uses to which it is put.

Narrative Representations



There are two types of participants involved in every semiotic act: *interactive participants* and *involved participants*. The former are the participants in the act of communication – the participants who speak and listen or write and read, make images or view them, whereas the latter are the participants who constitute the subject matter of the communications; that is the people, places and things (including abstract 'things') represented in and by the speech or writing or image, the participants about whom or which we are speaking or writing or producing images.

(Kress, 48)

Narrative processes always have vectors. Vectors represent processes. Actors are the participants from which the vectors depart.

The "hallmark of a narrative visual 'proposition' is the presence of a vector" (59). According to Kress and Leeuwen, vectors are "formed by depicted elements that form an oblique line, often a quite strong, diagonal line" (59). These vectors may be formed from a number of objects, from bodies to limbs to tools. There are usually two participants in a narrative representation. The "**Actor**" is "the participant from whom or which the vector departs" (59). That is, the "**actor**" is the object in the image that produces the action on the other object in the image. The passive participant in the action process is called the "**goal**." The "**goal**" is "the participant at which the vector is directed" (74). Kress and Leeuwen distinguish different narrative processes by the types of vectors and the number and kinds of participants involved.

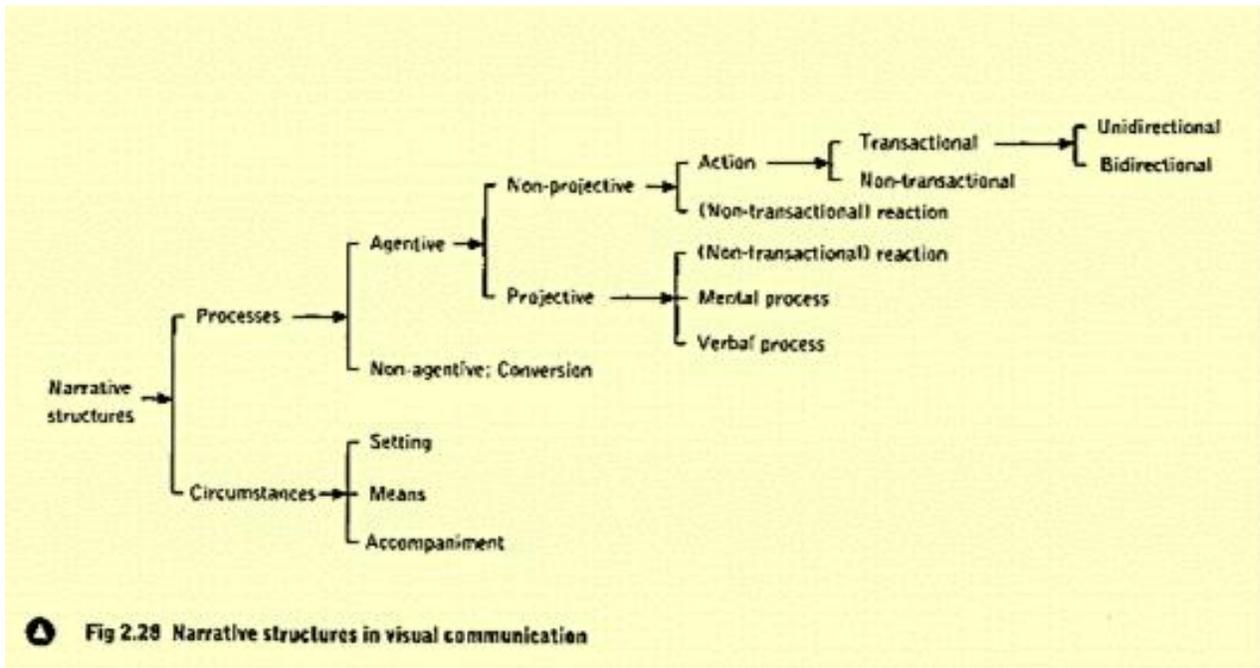
Action Processes

Those in which "the **Actor** is the participant from which the vector emanates...or forms the vector" (63). That is, in order to be considered an **Action Process**, the **Actor** must either have a vector coming out of it or actually form the vector. According to Kress and Leeuwen, when an image has only one participant, that participant can be considered the **Actor**. This is called a **non-transactional image**--there is no Goal in these images and vectors are not "done to or aimed at anyone or anything" (63). If there is only a vector and a Goal shown in the image, then Kress and Leeuwen call the action an **Event**. Lastly, when there are both an **Actor and a Goal** that are connected by a vector which stems from the **Actor**, then this process is called **Transactional**.

Reactional Processes

Those in which an eye line, such as a glance, by one or more of the participants, forms a vector that connects those participants. This process differs from the Action Process in that the vectors are formed solely by the gaze of the objects in the image. In this case the participant who does the looking is called a **Reactor** instead of an Actor, and the passive participant is called the **Phenomena, not the Goal** (67). The **Reactor** is the participant who forms the vector with his eyes; he does the looking. **Reactional Processes** can be **transactional and non-transactional** as well (68).

Realizations



REALIZATIONS

Unidirectional transactional action

A vector, formed by a (usually diagonal) depicted element, or an arrow, connects two participants, an Actor and a Goal.

Bidirectional transactional action

A vector, formed by a (usually diagonal) depicted element, or a double-headed arrow, connects two Interactors.

Non-transactional action

A vector, formed by a (usually diagonal) depicted element, or an arrow, emanates from a participant, the Actor, but does not point at any other participant.

Actor

The active participant in an action process is the participant from which the vector emanates or which is fused with the vector.

Goal

The passive participant in an action process is the participant at which the vector is directed.

Interactors

The participants in a transactional action process where the vector could be said to emanate from, *and* be directed at, both participants.

Transactional reaction

An eyeline vector connects two participants, a Reacter and a Phenomenon.

Non-transactional reaction

An eyeline vector emanates from a participant, the Reacter, but does not point at another participant.

Quick Guide to Reading Images – Prof. Paul Miers | 5

<i>Reacter</i>	The active participant in a reaction process is the participant whose look creates the eyeline.
<i>Phenomenon</i>	The passive participant in a (transactional) reaction is the participant at which the eyeline is directed; in other words, the participant which forms the object of the Reacter's look. The same term is used for the participant (verbal or non-verbal) enclosed by a 'thought bubble'.
<i>Conversion</i>	A process in which a participant, the Relay, is the Goal of one action and the Actor of another. This involves a change of state in the participant.
<i>Mental process</i>	A vector formed by a 'thought bubble' or a similar conventional device connects two participants, the Senser and the Phenomenon.
<i>Senser</i>	The participant from whom the 'thought bubble' vector emanates.
<i>Verbal process</i>	A vector formed by the arrow-like protrusion of a 'dialogue balloon' or similar device connects two participants, a Sayer and an Utterance.
<i>Sayer</i>	The participant in a verbal process from whom the 'dialogue balloon' emanates.
<i>Utterance</i>	The (verbal) participant enclosed in the 'dialogue balloon'.
<i>Setting</i>	The Setting of a process is recognizable because the participants in the foreground overlap and hence partially obscure it; because it is often drawn or painted in less detail, or, in the case of photography, has a softer focus; and because of contrasts in colour saturation and overall darkness or lightness between foreground and background.
<i>Means</i>	The Means of a process is formed by the tool with which the action is executed. It usually also forms the vector.
<i>Accompaniment</i>	An Accompaniment is a participant in a narrative structure which has no vectorial relation with other participants and cannot be interpreted as a Symbolic Attribute (see chapter 3).

Quick Guide to Reading Images – Prof. Paul Miers|6

Thematic relations in natural language grammar

Natural language grammar is traditionally said to be composed of three components: [syntax](#), [semantics](#), and [phonology](#). A fourth component, [pragmatics](#), must be added to account for information provided by context and expectation. Halliday's metafunctions involve all four of these components. His distinction between theme and rheme is usually given as a distinction between topic and focus. His notion of processes and relations roughly corresponds to what are known as [thematic relations](#):

- **Agent:** deliberately performs the action (e.g., Bill ate his soup quietly.).
- **Experiencer:** the entity that receives sensory or emotional input (e.g. Susan heard the song. I cried.).
- **Stimulus:** Entity that prompts sensory or emotional feeling - not deliberately (e.g. Kim detests sprouts).
- **Theme:** undergoes the action but does not change its state (e.g., We believe in one God. I have two children. I put the book on the table. He gave the gun to the police officer.) (Sometimes used interchangeably with patient.)
- **Patient:** undergoes the action and changes its state (e.g., The falling rocks crushed the car.). (Sometimes used interchangeably with theme.)
- **Instrument:** used to carry out the action (e.g., Jamie cut the ribbon with a pair of scissors.).
- **Force or Natural Cause:** mindlessly performs the action (e.g., An avalanche destroyed the ancient temple.).
- **Location:** where the action occurs (e.g., Johnny and Linda played carelessly in the park. I'll be at Julie's house studying for my test.).
- **Direction or Goal:** where the action is directed towards (e.g., The caravan continued on toward the distant oasis. He walked to school.).
- **Recipient:** a special kind of goal associated with verbs expressing a change in ownership, possession. (E.g., I sent John the letter. He gave the book to her.)
- **Source or Origin:** where the action originated (e.g., The rocket was launched from Central Command. She walked away from him.).
- **Time:** the time at which the action occurs (e.g., The rocket was launched yesterday.).
- **Beneficiary:** the entity for whose benefit the action occurs (e.g.. I baked Reggie a cake. He built a car for me. I fight for the king.).
- **Manner:** the way in which an action is carried out (e.g., With great urgency, Tabitha phoned 911.).
- **Purpose:** the reason for which an action is performed (e.g., Tabitha phoned 911 right away in order to get some help.).
- **Cause:** what caused the action to occur in the first place; not for what, rather because of what (e.g., Because Clyde was hungry, he ate the cake.).

There are no clear boundaries between these relations. For example, in "the hammer broke the Window," some linguists treat hammer as an agent, some others as instrument, while some others treat it as a special role different from these.

Conceptual Representations

Conceptual representations present the *participants* of the image in generalized categories: class, structure, or meaning.

Classification Processes

Participants are related in terms of a *taxonomy* (kind, type, classification) where one participant (or group of participants) act as *subordinates* in relation to another participant (or group of participants).

- Covert taxonomy
- Single-leveled overt taxonomy
- Multileveled overt taxonomy



Frida Kahlo, *My Grandparents, My Parents, and I (FamilyTree)* (1939)

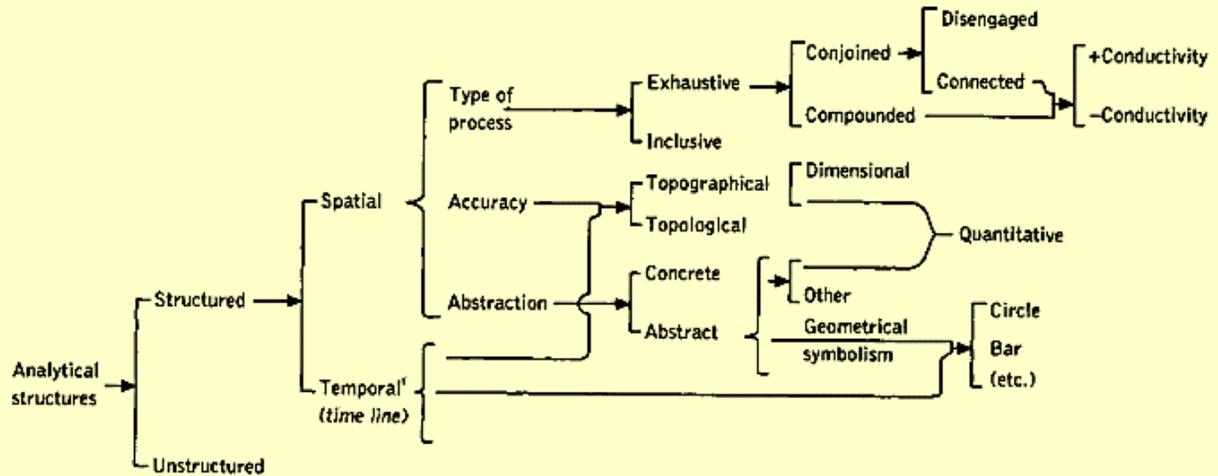


Fig 3.25 Analytic image structures

Quick Guide to Reading Images – Prof. Paul Miers | 8

Analytic Processes

Participants are related in a part-to-whole relationship where a *carrier* (the whole) is related to its *possessive attributes* (the parts).

- **Unstructured analytical processes:** exhibit the possessive attributes of the carrier, but not the carrier itself.
- **Temporal analytical processes:** the carrier represents "time" instead of an object and possessive attributes shown are points in time (i.e. a timeline).
- **Exhaustive/inclusive analytical processes:** representing all of the possessive attributes of a carrier (exhaustive) or representing only some of the possessive attributes of a carrier (inclusive).
- **Conjoined/compounded exhaustive structures:** possessive attributes are connected by a line that lacks direction or where the possessive attributes are separated, but it is clear how they fit together.
- **Topographical/topological processes:** possessive attributes are represented accurately within their spatial relationships within the carrier (topographical) or interconnections of possessive attributes are accurately depicted, but their scale is not accurate (topological).
- **Dimensional/quantitative topography:** the quantity or frequency of possessive attributes are drawn accurately to scale, but their physical dimensions are not.
- **Spatio/temporal analytic structures:** combine a structure where quantity, dimensions, relationships, etc. of possessive attributes are present, but they are depicted in accordance with their place in time.

Symbolic Processes

Symbolic Processes represent what a participant *means* or *is* (identity is established in the relationship between the *carrier* and the *symbolic attribute*).

Quick Guide to Reading Images – Prof. Paul Miers | 9

Modality

Reading Images applies the linguistic notion of mood / modality to visual representations.

In linguistics, grammatical mood (also mode) is a grammatical feature of verbs, used for signaling modality. That is, it is the use of verbal inflections that allow speakers to express their attitude toward what they are saying (e.g. a statement of fact, of desire, of command, etc.). The term is also used more broadly to describe the syntactic expression of modality, that is, the use of verb phrases that do not involve inflexion of the verb itself.



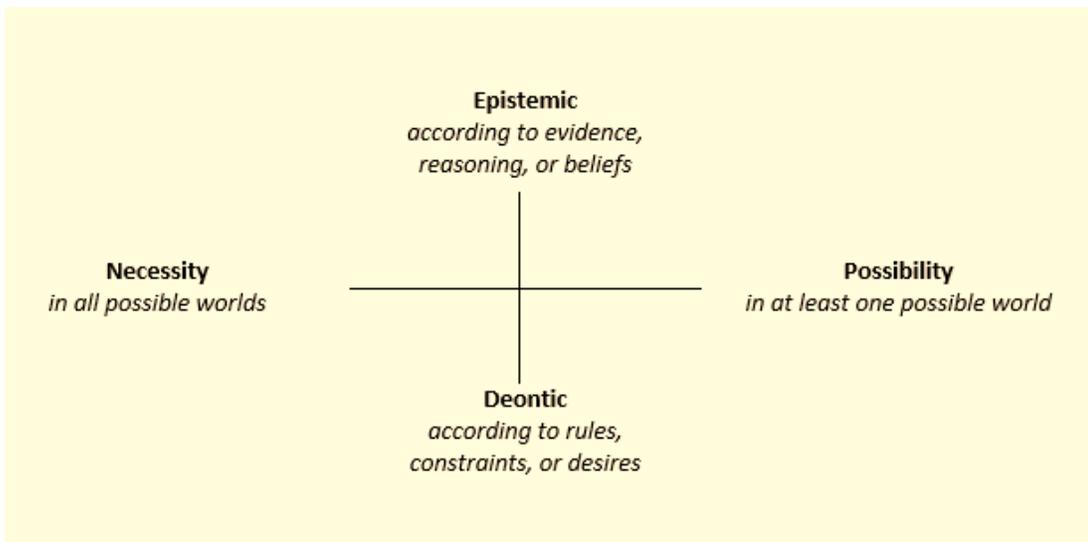
Caravaggio, *The Incredulity of Saint Thomas* (1601-1602)

Some examples of moods are indicative, interrogatory, imperative, emphatic, subjunctive, injunctive, optative, potential. These are all finite forms of the verb. Infinitives, gerunds, and participles, which are non-finite forms of the verb, are not considered to be examples of moods.

Epistemic vs. deontic modality

Epistemic modals are used to indicate the possibility or necessity of some piece of knowledge. In the epistemic use, modals can be interpreted as indicating inference or some other process of reasoning involved in coming to the conclusion stated in the sentence containing the modal ("Can I?").

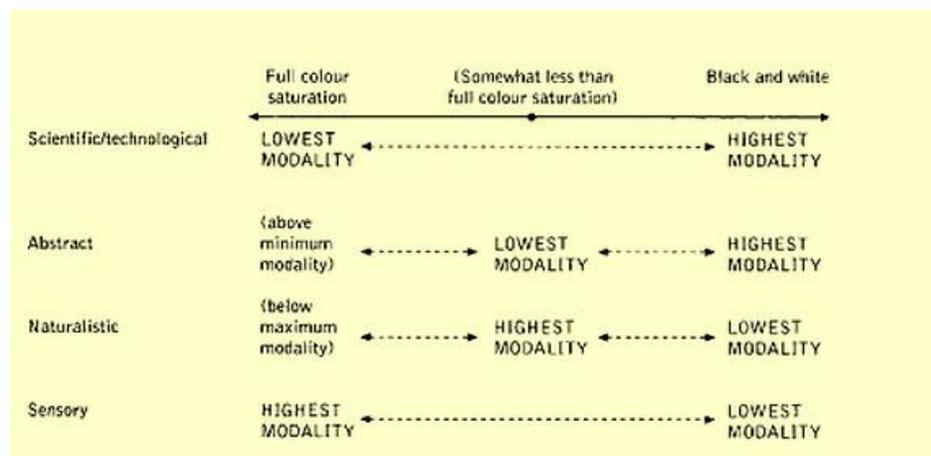
In contrast, deontic modality is concerned with possibility and necessity in terms of freedom to act (including ability, permission, and duty). English examples include She can go (ability), You may go (permission), You should go (obligation), and You must go (strong obligation) In English as in many other languages, some of the same words are used for deontic modality as for epistemic modality ("May I?").



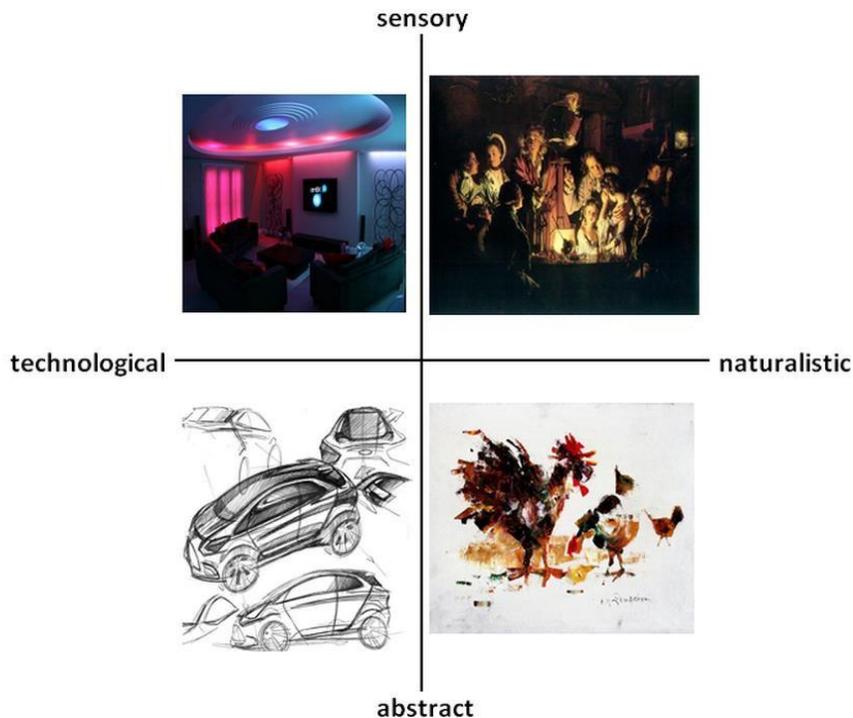
Quick Guide to Reading Images – Prof. Paul Miers|10

Coding orientations for visual modality

- (1) *Technological coding orientations*, which have, as their dominant principle, the 'effectiveness' of the visual representation as a 'blueprint'. Whenever colour, for example, is useless for the scientific or technological purpose of the image, it has, in this context, low modality.
- (2) *Sensory coding orientations*, which are used in contexts in which the pleasure principle is allowed to be the dominant: certain kinds of art, advertising, fashion, food photography, interior decoration, and so on. Here colour is a source of pleasure and affective meanings, and consequently it conveys high modality: vibrant reds, soothing blues, and so on – a whole psychology of colour has evolved to support this.
- (3) *Abstract coding orientations*, which are used by sociocultural elites – in 'high' art, in academic and scientific contexts, and so on. In such contexts modality is higher the more an image reduces the individual to the general, and the concrete to its essential qualities. The ability to produce and/or read texts grounded in this coding orientation is a mark of social distinction, of being an 'educated person' or a 'serious artist'.
- (4) The common sense *naturalistic coding orientation*, which remains, for the time being, the dominant one in our society. It is the one coding orientation all members of the culture share when they are being addressed as 'members of our culture', regardless of how much education or scientific-technological training they have received. Individuals with special education or group allegiance may draw on non-naturalistic coding orientations in certain contexts, but they are likely to revert to the naturalistic coding orientation when they are 'just being themselves'. They may, for example, use the abstract coding orientation when visiting a gallery, and the naturalistic coding orientation when watching television or reading a magazine. For those without such education, however, abstract and technological images will never have high modality and always remain 'unreal'. Today, however, naturalism is coming into crisis, as a result of new ways of thinking and new image technologies. In this context the role of some or all of the non-naturalistic coding orientations is likely to become of increasing importance.



Quick Guide to Reading Images – Prof. Paul Miers|11



Modality markers

Color Saturation a scale running from full color saturation to the absence of color, that is to black and white.

Color differentiation a scale running from a maximally diversified range of colors to monochrome.

Color modulation a scale running from fully modulated color, with for example, the use of many different shades of red, to plain, unmodulated color.

Contextualization

a scale running from the absence of background to the most fully articulated and detailed background.

Representation a scale running from maximum abstraction to maximum representation of pictorial detail.

Depth a scale running from the absence of depth to maximally deep perspective.

Illumination a scale running from the fullest representation of the play of light and shade to its absence.

Brightness

a scale running from a maximum number of different degrees of brightness to just two degrees: black and white, or dark grey and lighter grey, or two brightness values of the same color.

Position of the Viewer

The articulation and understanding of social meanings in images derives from the visual articulation of social meanings in face-to-face interaction, the spatial positions allocated to different kinds of social actors in interaction (whether they are seated or standing, side by side or facing each other frontally, etc.). In this sense the interactive dimension of images is the 'writing' of what is usually called 'non-verbal communication', a 'language' shared by producers and viewers alike.



Jan Vermeer, Lady Standing at the Virginal (1670-1672)

Demand – gaze at the viewer
Offer – absence of gaze at the viewer

Intimate / personal – close shot
Social – medium shot
Impersonal – long shot

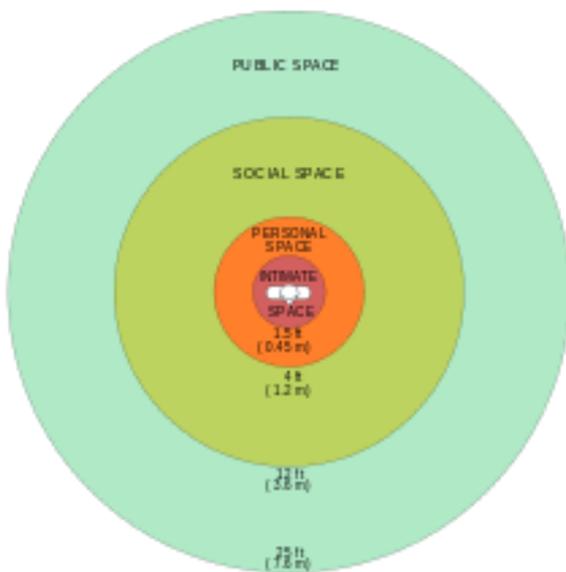
Involvement – frontal angle
Detachment – oblique angle

Viewer Power – high angle
Equality -eye level angle
Participant Power – low angle

contact

social distance

attitude



Proxemics is the study of human use of space and the effects that population density has on behavior, communication, and social interaction. Edward T. Hall, the cultural anthropologist who coined the term in 1963, defined proxemics as "the interrelated observations and theories of humans use of space as a specialized elaboration of culture".

- **Intimate distance** for embracing, touching, or whispering
- **Personal distance** for interactions among good friends or family
- **Social distance** for interactions among acquaintances
- **Public distance** used for public speaking

Composition

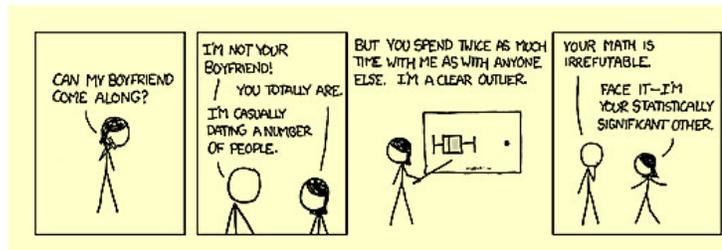
Composition relates the representational and interactive meanings of the image to each other through three interrelated systems:

- **Informational value.** The placement of elements (participants and syntagms that relate them to each other and to the viewer) endows them with the specific informational values attached to the various “zones” of the image:

- left and right; ○ top and bottom;
- center and margin.

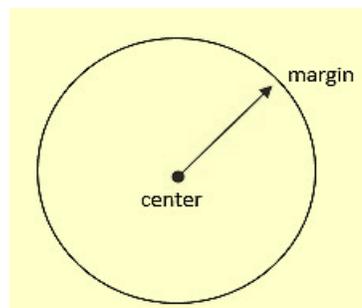


Traditional images in western culture are read from left to right without a signal or vector indicating the directional value. Images containing elements that flow top to bottom use an *ideal/real* relationship where the ideal is represented on top, with the real depiction shown underneath. The viewer experiences the ideal first, so the reading of that element is dominant.



As an image moves in meaning from left to right or top to bottom, the contrast between related images can be seen as *given and new*. The given image, read first, portrays something the viewer is already familiar with and has established as true. The new, the changed image, represents something that the viewer has learned or will come to conclude based on the inner meaning of what is represented in the shift. Advertisements often use text as the given and imagery as the new.

ideal given	ideal new
real given	real new





- **Salience.** The elements (participants as well as representational and interactive syntagms) are made to attract the viewer's attention to different degrees, as realized by such factors as placement in the foreground or background, relative size, contrasts in tonal value (or color), differences in sharpness, etc.

- **Framing.** The presence or absence of framing devices (realized by elements which create dividing lines, or by actual frame lines) disconnects or connects elements of the image, signifying that they belong or do not belong together in some sense.



[Edward Hicks, *The Peaceable Kingdom* \(1826\)](#)

Mode of reception



Public reception (movie theaters, billboards, city walls) allows for larger images that are often more simple to decode, as the scale is grand and the detail obvious.

Private reception (magazine ads, computer screens) are more intricate, as the viewer can see them up close and study them