Forgotten twins

Reason and visuality

Frederik Stjernfelt

Much has been said in recent years about imagology, about the increasing flood of pictures in media and internet, about computer manipulation liberating photos and moving pictures from reference, about images creating the world rather than forming part of it or giving access to it. This whole current has celebrated this allegedly dramatic event as an inevitable victory of pictures over the spoken word - if it has not mourned pessimistically over the dark destiny of public reason in the reign of irrational images.

This paper claims such analyses overlook basic structures in image production and use. Scenarios like the ones mentioned rest on a simplistic dualism contrasting images and words, or, in a more sophisticated terminology, the sign types of icons and symbols. Often, it adds the idea that images are sensuous, seductive, particular - as opposed to the cool, general rationality constrastively ascribed to words. But such a contrast not only rests on untenable dualisms projected onto the icon-symbol distinction, it also overlooks the basic fact that much if not most picture use takes place within the confines of propositions mixing the use of words and image in different ways. The typical occurrence of pictures, on the net, in tv, in papers and books, in urban environments etc., is not as single, mysterious images, but rather in intimate connection with text. Moreover, the dualist claim also overlooks that pictures are by no means, even if taken in isolation, alien to reason. Rather, images form a central tool of rationality - it is not only possible but indeed both widespread and normal to reason with pictures.

This property comes to the fore especially in artificially created pictures as diagrams in media, science, and politics - but it is, in turn, based on a more fundamental property in pictures as such, namely that they allow for experimental manipulation revealing their implicit properties.

The reason why reasoning with images has been overlooked and marginalized in academic discourse stems from the folk dualism of pictures and words - but also from certain intellectual traditions in recent philosophy of science. One is that reasoning is taken to be primarily or even exclusively
linguistic - often tied to the famous "linguistic turn" in philosophy. Assuming such ideas, images immediately become, by contrast, something alien to reasoning. The fact that many pictures may appear more ambiguous (or, rich in information) than some linguistic utterances has given rise to a parallel idea, that of images being something potentially misleading which should be controlled or even marginalized by a discourse articulated in ordinary or formal languages. Thus, in mathematics, a tradition beginning in mid-19 century to reach its peak with Bourbaki in the mid-20 century aimed at expelling diagrams completely from mathematical presentations - because such diagrams might mislead the reader because they invariably include some degree of particularity and thus an excess of information. Curiously, this hygienic idea developed at the same time as empirical diagrams, charts, maps, graphs, tables, schemas etc. blossomed across the sciences, from nature to culture - you only need to open any scientific periodical to assure yourself about that fact. In philosophy, the dominance of the above-mentioned "linguistic turn" in the 20 century and the ensuing hostility against pictures had a peculiarity which has seldom been thematized - it is one of the few influential ideas which cross over the divide between continental and analytic philosophy. To take the analytical tradition, pictures and images not only occupied a marginal position in the heydays of the linguistic turn - one of the basic characteristics of pictures, some degree of similarity between them and their object, was exposed to a famous attack by a leading proponent of the analytic tradition, Nelson Goodman. At the same time, presumed iconicity in signs was attacked by structuralists like Umberto Eco, assuming that all signs were conventional through and through, and that iconicity was rather an ideological surface seductively making some signs seem more "natural" than other signs.¹

What is overlooked by these anti-iconic currents is several important facts. One is that most linguistic expressions are no less ambiguous and vague than many images. Only formalized language in logic and sciences aim at a higher degree of clearness and disambiguation. But even such formalized languages do not eradicate intuition totally - even linear, symbolic logic preserves the spatial intuition of a line with discrete places occupied by tokens which may be exchanged and manipulated according to certain rules. So there is no such thing as language or symbolic formalisms completely deprived of spatial properties. Conversely, formal linear languages are far from the only means of disambiguation and clarity - images may be stylized and diagrammaticized in order to achieve exactly the same effect - as in the plethora of different image types like sketches, caricatures, public image signs, maps, diagrams, schemas, graphs, tables,
matrices, and much more. Such signs are different ways of simplifying expressions by leaving out certain aspects of the object depicted in favor of others which may even be exaggerated. These facts imply that the folk-ontological idea of an easy borderline between images and words, and between diagrams and formal symbol systems, can not be upheld.

On top of this comes the strangely overlooked even if ubiquitous fact of the co-occurrence of pictures and words. To take some examples: paintings with a title or legend, art exhibitions with catalogues, newspaper photographs with an accompanying caption as well as article, spoken tv-news accompanied by newsreel footage, feature films involving dialogue, narrator's voice-over, title sequence etc., comics featuring speech balloons and onomatopoetic words indicating noisy events etc., caricatures involving legends indicating the person or state-of-affairs caricatured, maps with geographical and other names indicated, books with illustrations, scientific papers with graphs summing up empirical results or theoretical models, mathematical articles with figures, websites constructed from intricate webs of images and texts, and much, much more. This omnipresence of word-image conglomerates is so pervasive that we should rather see the occurrence of pure text without any pictures - or the appearance of images not accompanied by any text - as rare, marginal special cases chosen for aiming at particular effects.

This pervasive word-image duplicity has its base in another often-overlooked fact - that of the central status of propositions in semiotics. The centrality of claiming something to be the case is easily ignored if focus is on the medium, the sign vehicle or the sensory mode of the representation discussed. Thus, propositions have mainly been studied by logic, which is probably why the perception has spread that propositions are something exclusively linguistic, confined to ordinary languages or the formalized languages invented to study the logical relations between propositions and their constituents. But propositions in the wild, as it were, are by no means confined to language. This is made evident in Charles Peirce, one of the central discoverers of modern logic, by his concept of "Dicisigns". This is his version of the notion of proposition, but the definition of it loosens it definitively from language - as well as from being a human privilege only. Peirce defines Dicisigns as signs which may convey truths (or falsehoods) - state that something is the case. The means for a Dicisign to state something to be the case is to connect to that something by two different relations simultaneously. The Dicisign at once indicates and describes its object. If the description fits the object indicated, the Dicisign makes a true claim, if not, the Dicisign is a false proposition. The indication part of the sign is an index,
making it possible for the receiver to locate the object of the sign; the descriptive part of it is an icon, making it possible for the receiver to imagine the property ascribed to it. This duplicity of Dicisigns is what has often been described as the Subject-Predicate structure of Propositions, and it may be realized in particular signs in many different ways. Both aspects, for instance, may be realized gesturally which is of course the case in sign languages for the deaf - but also in ordinary communication: if I point to some person and accompany this indication by pointing to my own temple with a rotating finger, I am making a Dicisign equivalent to the linguistic utterance "He's crazy". This gestural Dicisign has the same Subject-Predicate duplicity - composed of indicating plus describing - as its linguistic counterpart's duplicity of "He" and the unsaturated predicate "_ is crazy". A very widespread possibility is to let the descriptive task of the Dicisign be satisfied by an image. A painted portrait of Napoleon accompanied by a small text sign "Napoleon" constitutes a Dicisign claiming Napoleon looked like the person depicted. Here, of course, the title plays the subject role, indicating the reference object of the sign, while the painting plays the predicate role, describing that object. In some cases, a picture may, seemingly by itself, play the role of both S and P. If we recognize, e.g., president Obama in a photograph displaying him smoking a joint, one of the Dicisigns claimed by that photograph is that the president has indeed smoked pot. This interpretation of the photograph, however, is possible only given the fact that the recipient already knows the looks of the president and is able to recognize him - what Peirce calls "collateral knowledge" of the subject. Without such collateral knowledge, the viewer is unable to access the Dicisign mentioned and must content himself with the considerably weaker Dicisign that some black guy has smoked pot. But collateral knowledge of the type needed here depends upon some earlier indexical pointing out of the object. The observer must have become acquainted with the idea that a certain look is stably connected to the unique person indicated by the proper name of Obama and the role of US President of a certain term. So the picture's ability to constitute, by itself, full-fledged Dicisigns is dependent upon former indications of its subject.

The fact that a single such photograph involves an indefinite amount of Dicisigns ("a black man smokes something", "a well-dressed guy looks into the camera", "somebody took a photo of Obama as a young man" etc.) has often been taken to indicate the important vagueness of picture signs. This is indeed correct, but it should not go without adding that textual utterances making explicit one such Dicisign also involve other vaguenesses, albeit in other dimensions. The claim "Young Obama smoked pot" leaves
open when, how, how long, which clothes he wore, and, of course, whether he inhaled.

The invention of modern picture technologies (photo, tv, internet etc.) directly causally influenced by the object they depict adds the idea that certain picture representations are, in themselves, indexical and thus furnish the picture with its internal possibility of being indices and icons at the same time, thus constituting Dicisigns. Any photograph inevitably involves the vague Dicisign that "Something, somewhere, at some point of time, looked like this". In the absence of further collateral knowledge about the object of the Photograph, such a Dicisign admittedly says but little. Such indexical picture technologies, of course, are subject to forgeries, form Stalinist history rewriting to Photoshop. The fact that they may thus lie, however, does not imply they are not Dicisigns - quite on the contrary, being a Dicisign is the prerequisite of conveying truth as well as falsity. In some cases, the image's role in a Dicisign may be almost reversed as compared to the painting-with-title example. When we watch our TV reporter on the ground in Libya relating the recent development in the struggle between government and insurgents, the role of the picture of the journalist in the desert is primarily indexical: it is to insure us that we are actually gaining information from the relevant location itself. Here, the text read aloud by the reporter adds to the description of what takes place at the place indicated, thus largely (but not exclusively) falling on the descriptive side of the chain of Dicisigns presented by the news item. So many different sign combinations as regards the S and P aspects of Dicisigns are possible. As is well known, language itself involves means for both indication and description, the former taken care of by proper names, pronouns, time-and-place adverbs, context, etc., the latter supported by common nouns, verbs, adjectives, etc.

The subject-predicate divide of Dicisigns thus does not fit one-to-one with the empirical media facilities of text and picture; rather we can say that the duplicity of text and picture makes possible many different ways of satisfying the duplicity of Dicisigns. The important lesson to learn, however, is that images, and visual pictures in particular, play important roles in compound signs making truth claims possible. Pictures may be used truthfully to describe - and falsely to lie about - properties of all sorts of different objects, entities, events, states-of-affairs etc. A picture type like topographical maps are basically complex Dicisigns making a claim about the structure of a particular geographical domain. Map shapes provide the descriptive aspect of the Dicisigns, supported by different description conventions like that of the green-yellow-brown scale of landscape height
and the light blue-blue-deep blue-purple scale of sea depth. Longitude-latitudes, place names, scale indications etc. provide the indexical aspects of map Dicisigns, the two of them together constituting the map as one complex Dicisign. If such a Dicisign should be translated into linguistic propositions, of course, it would yield an indefinite range of propositional claims, such as "London is to the south of Glasgow", "Ireland is an island", "Longitude 0 crosses Greenwich", etc. GPS technology adds the indication of the location of the GPS itself (and its observer) on the map, as well as the putative route to selected goals, adding further explicit Dicisigns to the map (Dicisigns not explicit but which may be inferred from ordinary maps, cf. below). A more abstract diagram picture like a pie chart may depict the relation between the percentage of support to different parties by an electorate - each such party represented by one pie slice. Again, this textual-visual compound diagram will translate into a range of different simple Dicisigns if translated into linguistic statements only.iii Already on this level, visual signs are thus deeply involved in the representation of knowledge and in speech acts like instructions, imperatives, wishes, questions etc. utilizing such Dicisigns.

But visual signs as descriptive and reference-granting tools in Dicisigns only form the first step of reasoning with pictures. Visual signs also play important roles in reasoning as argumentative chains and complexes of Dicisigns. This is the issue of Diagrammatical Reasoning.

Images prototypically contribute to the descriptive aspects of Dicisigns. The possibility of diagrammatical reasoning lies in the fact that iconic signs are not necessarily simple. They may involve different degrees of complexity. Take again the example of topographical maps. They involve many different geographical locations in one complex sign; they involve several aspects of geographical properties, landscape types, shapes, heights, human additions to geography such as citites, railways, roads and much more. This has the important implication that most of the relations within such icons are implicit. And this implies, in turn, the possibility of extracting the information about such relations by means of manipulations of the icon: reasoning with diagrams. Such reasoning makes explicit icon information in the shape of Dicisigns. Such Dicisigns may be easily read off of the diagram: the fact that London is to the south of Glasgow is not explicitly stated in the diagram (unlike the fact that this large city in the south of England has the name of London), but it does not require complicated cognitive machinery to realize that fact. Not so with other relations in the diagram. Already the issue of the route between locations on the map is less
trivial. The direct linear connection between two points, of course, is a simple solution (although also dependent upon the map projection chosen) to the task of finding the route; more complications arise from the fact that linear traveling is rarely possible (see below). Given a particular route between two locations, however, to find the length of that route forms a non-trivial diagram experiment. It may be done, of course, with a quick approximate calculation with a ruler and the knowledge of the map's scale - multiply the map distance with the map scale in order to get the real landscape distance. The logically important issue here is that this distance was never explicit on the map, nor even during the construction process of the map - it is an implied information which may be made explicit in a Dicisign: the distance between London and Glasgow is 550 kilometers.

Actually, selecting the relevant road connection among many possible forms, in itself, a non-trivial problem: the intricacies of different road connections via different intermediary towns makes finding the shortest route a version of the well-known Traveling Salesman problem which is computationally NP-complete problem: it is well-known that given a number of cities and the task of finding the shortest route visiting each city once, the only algorithm known is the stepwise charting each single such route and comparing their length, a task which grows exponentially with the number of cities. This cumbersome experiment may be undertaken using map, ruler and the map's scale.

The topographical map example occurs to me as a both simple and convincing example of diagrammatical reasoning. But many simpler examples of such reasoning are undertaken automatically by our visual system - e.g. edge detection taking place already in the retina, in turn facilitating the parsing of the visual field into autonomous objects - implicitly facilitating Dicisigns like "There's an object to the left of me", making us turn the head in that direction. Stereoscopic vision immediately parses the visual field after distances, making possible perceptual judgments such as "this man is far away" (which is in no need of being linguistically expressed; we immediately realize, while seeing, that this man is far away. Cognitive neuroscience has long since established that the visual cortex in the occipital lobe sends out two different streams of information to further processing, the so-called Dorsal and Ventral streams, also nicknamed the "Where" and the "What" stream. The Dorsal stream informs motor cortex about the spatial whereabouts of the immediate surroundings, facilitating our easy and precise interaction with objects of the environment. Take Ventral stream, by contrast, facilitates categorization of the objects perceived - an
endeavor much more prone to error than the mere spatial charting of object locations, sizes, and shapes. The split between the Dorsal and Ventral stream exist in many higher animals and a daunting hypothesis is that the two of them seem to form the brain's adaptation to and realization of the S-P-structure of propositions (Hurford 2007): the Dorsal stream indicates the spatio-temporal place of an object, while the Ventral stream describes that object as belonging to some property-defined category. The Dorsal-Ventral split functions, then, as a sort of mental pair of thongs, making it possible for perception to realize simple pre-linguistic Dicisigns pertaining to single objects, as when we perceptually realize that "This thing is blue". Such a perceptual judgment, hence, functions as the conclusion of an (automatic) diagrammatical inference from visual material. Thus, the empiricist idea that perception is a simple process of impression which is only subsequently subjected to reasoning is plainly wrong. Perception already involves intricate logical inferences performed within the visual material. In that sense, simple seeing already involves reasoning.

The fact that reasoning is taking place in visual material all of the time is probably most conspicuously evident in cinema. The piecing together and drawing inferences from visually presented information is part and parcel of understanding any feature movie or tv-series. The piecing together of shots to form a scene, of scenes to form a sequence, is an inference process normally not requiring conscious attention to the reasoning. In scenes involving interacting human beings, reasoning on the basis of perceived gaze directions, objects perceived, persons perceived, eye contacts, accepted eye contacts, avoided eye contacts, perceived eye contacts etc. is crucial for inferring the intersubjective relations of a scene, the changing intentions of each character, etc. Of course, such reasoning also crucially relies upon dialogue, but in most cases both visual and auditory information goes together in the reasoning process establishing the understanding of a scene. The seamlessness of such inference facilitated by skilled mainstream visual products may be interrupted by jump shots, bringing together seemingly unrelated material in two neighboring shots - revealing that seamlessness is not a basic given, but rather the product of an ongoing reasoning process which may falter given the appearance of uninterpretable material (be it in bad productions or in aesthetically advanced art cinema). The seamlessness of ensuing match shots in cinema, however, is shaped so as to facilitate easy reasoning on part of the observer, thereby approaching the seamlessness of (most, not all) average visual reasoning in everyday life. Such visual-social intelligence embodied in the calculus of gaze-directions have been made the object of research, leading to the Bruner-Tomasello hypothesis of "joint
attention” as basic for the establishment of human common reasoning and culture. This only taken as an example of the centrality of visual reasoning already in everyday intersubjective social life.

On such bases, individually and socially, visual imagery may be made the object of more explicitly controlled, conscious experimenting. Watching a car approaching we may judge we can cross the road before it reaches us; watching a landscape painting, for instance, we may judge the distances between foreground, middleground, and background as a simple diagram experiment, watching a feature film, we may infer from two successive frames depicting the hero and villain, respectively, that they are on collision course in the same space. On the basis of such everyday pieces of visual reasoning, more complicated scientific experiments may be undertaken. One of Peirce's pet examples is the famous geometrical proofs of Euclid where the use of ruler and compass on a particular geometrical figure allows for general proofs of properties of whole classes of such figures. The solution of equations - from easy to very difficult cases - forms an algebraical example of diagrammatic reasoning. These examples are from the non-empirical sciences of geometry and algebra - but also experiments with idealized diagrams and models of empirical states-of-affairs may make possible inferences giving rise to the expression of surprising new hypothetical Dicisigns. This is, in general, an overlooked aspect of the formation of diagrams and models in the sciences - they not only synthesize amounts of data, propose general explanations, facilitate the comparison between a hypothesis and empirical findings - they also facilitate the rule-bound experiment with the diagram in order to chart new, unexpected implications of the model.

Psychologically, there are certain limits to the upper bounds of complexity of diagrams in the imagination. Few people, if any, can memorize a topographical map to use for their inner contemplation and experimentation, and even if certain autists are reported to be able to observe considerably larger imagined diagrams than average people, even their imagination has an upper bound. This is one basic reason for what makes externally stored visual representations so effective for many purposes, artistically, politically, scientifically, etc. Such representations transgress the complexity of controllable, imagined visual imagery in individuals and thus provide important tools for the "extended mind", relying upon externally stored information of many sorts for economically acting and thinking. Moreover, such externalized visual signs may be the co-product of several individuals who may collaborate on devising, developing, discussing and criticizing them, just as they may, in turn, serve as tools for even further
individuals. Thus, diagrammatic reasoning in the public sphere forms a central node for the development of science, politics, the arts, etc. Many scientific papers revolve around establishing the synthesis of empirical findings and theoretical assumptions in diagrams which formalize and communicate the central DiCisign chain of arguments of those papers, thereby communicating it to other scientists, challenging them to repeat the empirical experiments as well as the ideal thought experiment of the diagram presented.

On the basis of the notion of diagrammatic reasoning, Peirce proposed a daring generalization in several dimensions. First, he proposed that all mathematics consists of diagrammatical reasoning. This comes from the idea that mathematics in general concerns idealized hypotheses concerning forms of relation - and that it proceeds with experimenting with diagrams representing such idealized hypotheses. The possibility for mathematical to reach unshakable truths is thus bought at the price that such truths are possible for idealized worlds only, with few, well-controlled properties, leaving all other properties indeterminate (unlike the empirical world of which knowledge is always bound to be more or less approximate). Second, he proposed that all deductive reasoning is undertaken by diagrammatical reasoning. This implies that also pieces of deductive reasoning where no explicit diagram is present has an underlying diagrammatical structure. This also implies that ordinary deductive reasoning, in introspection, language, images, etc. is diagrammatical. And, taken together with the first generalization, it implies that both everyday reasoning and empirical scientific reasoning involves mathematical structures, even when not at all apparent.

Discussing diagrammatical reasoning in different contexts, these points always prove most difficult to get through to the audience. It may already be difficult to appreciate the idea that algebraical representations and formal languages should count as diagrams, in addition to more well-known core examples like maps and graphs. But it seems even more difficult to admit that linguistic structure and semantics share diagrammatical aspects (even if this may seem less strange in cognitive linguistics with a schematic view of semantics and construction grammar). Many people prefer to stick to their folk-semiotic spontaneous ideas where icons are signs which immediately look like something, where diagrams are visually explicit schemata only, where language is completely un-iconic and non-diagrammatical, where all reasoning takes place in language, etc. I tend to think Peirce's ingenious generalization opens up a completely new vista
which makes better sense by connecting issues kept far apart by such spontaneous folk semiotics: iconicity, visuality, ideality, reasoning, language, diagrams, ...

In particular, it is a vista which makes us realize the close and natural connection there is between visuality and reason.

Bibliography


- 1997 “Peirce’s Philosophical Conception of Sets”, in Houser et al. (eds.) 1997.


forthcoming "Theoric Transformations" and a New Classification of Abductive Inferences", in: Transactions of the Charles S. Peirce Society

Houser, Nathan, Don D. Roberts and James van Evra (eds.) 1997 Studies in the Logic of Charles Sanders Peirce, Bloomington: Indiana University Press


1975-87 Contributions to the Nation, I-IV (ed. K. Ketner & J.E. Cook), Lubbock TX: Texas Technological University Press.


(no year) “Logic, Considered as Semeiotic” [LCS], constructed from manuscript L 75 by Joseph Ransdell (1902) (http://members.door.net/arisbe/menu/library/bycsp/l75/ver1/l75v1-01.htm).

Manuscripts at the Houghton Library referred to by Ms. numbers in the Microfilm edition 1966 The Charles S. Peirce Papers, Microfilm Edition, Thirty Reels with Two Supplementary Reels Later Added. Cambridge: Harvard University Library Photographic Service. Manuscript numbers also refer to the Robin catalogue of the Mss.(Robin 1967). As to manuscript page numbers, "CSP" refers to Peirce's pagination (which is not
unanimous since several parallel drafts may belong to the same Ms.), while "ISP" refers to the pagination of the Mss. by the Institute for Studies in Pragmaticism, also used by the Peirce Edition Project in Indianapolis. When ISP numbers are not available, numbers refer to the sequence of the microfilm reel, e.g. "Reel 16, 1434-35" meaning the 1434th and 1435th page of reel 16.


- 1997 "Kant's Syntheticity Revised by Peirce", *Synthese* 113, 1-41
- 2000 The Iconic Logic of Peirce's Graphs, Camb. MA: MIT

Stjernfelt, F. 2007 Diagrammatology. An Investigation on the Borderlines of Phenomenology, Ontology, and Semiotics, Dordrecht: Springer Verlag
Stjernfelt, F. 2008 "Spatial Cognition Strategies in Map and Diagram Reasoning", in P.A.Brandt og C. Carstensen (eds.) *The Map is not the Territory*, Esbjerg, 167-73
Stjernfelt, F. forthcoming d) "Peirce's Notion of Diagram Experiment. Corollarial and Theorematical Experiments With Diagrams", in D. Wagner et al. (eds.) *Papers from the 33th International Wittgenstein Symposium in Kirchberg am Wechsel 2010*
Tiercelin, C. 2010 "Peirce on Mathematical Objects and Mathematical Objectivity", in Moore 2010, 81-122.
i Stjernfelt (2007) ch. 3 presents a more elaborated criticism of Goodman’s and Eco’s anti-iconicities.

ii See Stjernfelt (forthcoming a).

iii The potent visual-textual duplicity of the topographical map is prefigured in the close interaction between spatial cognition and language in spatial orientation vocabularies in languages. Egocentric, allocentric or geocentric strategies select different reference points for the bodily point-of-view reference - and they may combine in different ways in different languages. In any case, they provide for linguistic means for describing and reasoning about the spatial relations between objects and viewpoints in visual space, see Stjernfelt 2008.

iv Cf the cinematic notion of "eye matching" referring to the fitting together of e.g. a gaze direction in one shot and the object looked at in the next.

v In Stjernfelt (forthcoming d), I take Alfred Wegener’s experimental fitting together of Africa’s West coast and South America’s East coast as an example.

vi Peirce’s icon and diagram concepts are not confined to vision. Auditory or tactile, even to some extent olfactory diagrams are also possible, addressing spatial information - in other beings, diagrams with even more outlandish perceptual support may be imagined. But with the central role played by vision in human perception, diagrams are primarily - if far from only - connected to the mode of vision.