Roald Hoffmann and Pierre Laszlo

REPRESENTATION IN CHEMISTRY

Chemical structures are among the trademarks of our profession, as surely chemical as flasks, beakers and distillation columns. When someone sees one of us busily scribbling formulas or structures, he or she has no trouble identifying a chemist. Yet these familiar objects, which accompany our work from start to end, from the initial doodlings (Fig. I) to the final polished artwork in a publication (Fig. II), are deceptively simple. They raise interesting and difficult questions about representation. It is the intent of this article to reflect upon molecular graphics.

For the purpose of this discussion we will use the term "chemical structure" for the entire spectrum of representation that chemists normally use.[1] Thus we will include what are called for-

[1] For a general introduction to molecules see the beautiful book by P.W. Atkins, *Molecules*, Scientific American Library, New York, 1987.

https://doi.org/10.1177/039219218903714702 Published online by Cambridge University Press

mulas or molecular formulas, i.e. the listing of the elements in a molecule, with their correct ratios (for example water = H_2O, methane = CH_4, thalidomide = $C_{13}H_{10}N_2O_4$). But we will emphasize and focus on those representations so strikingly in the foreground in Figures I and II.

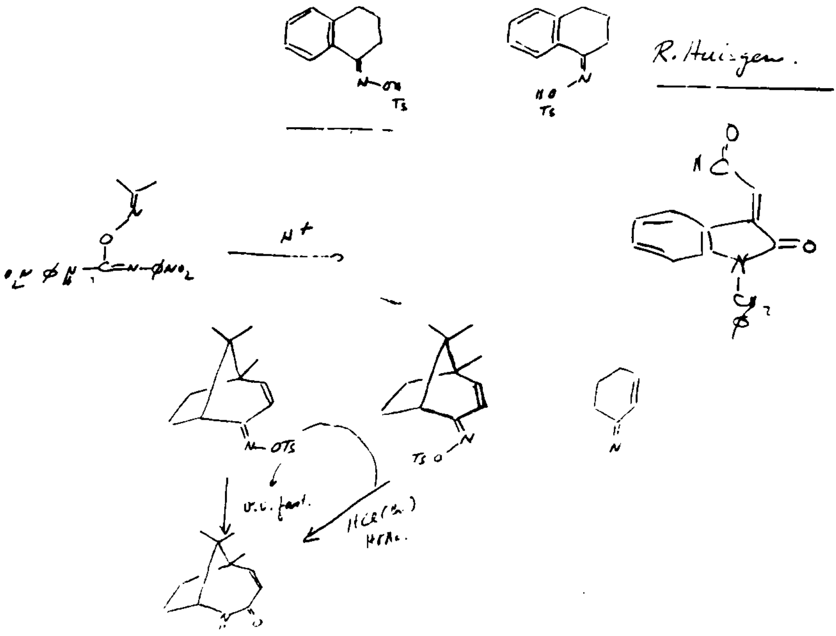
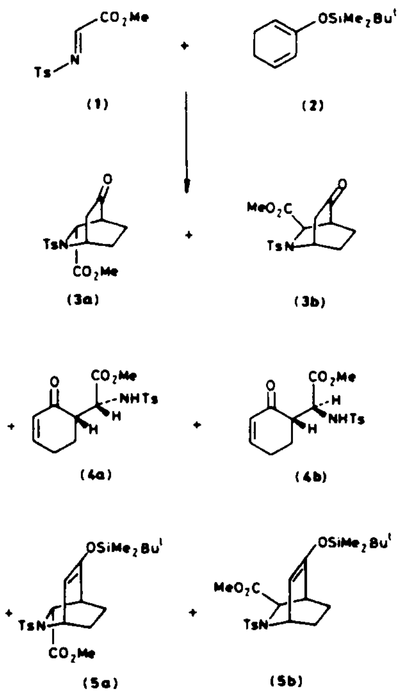


Figure I: A drawing by R.B. Woodward, ca. 1966, in the course of a discussion. R. Huisgen is a well-known German chemist.

These constitutional formulas embody the crucial connectivity of atoms in molecules, and hint at the three-dimensional arrangement of these atoms.

https://doi.org/10.1177/039219218903714702 Published online by Cambridge University Press



Ts = p-$MeC_6H_4SO_2$

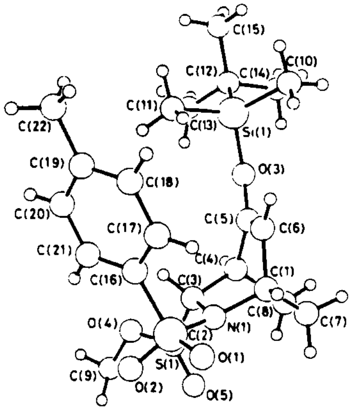


Figure 1. The structure of the *exo* silyl enol ether (**5a**).

The results of a systematic variation of solvent, tempera-
'e, and Lewis acid on the outcome of the reaction of (**1**) with
ı resemble the pattern reported for the trimethylsilyl
alogue of (**2**), but with several important differences.
'stly, the *exo/endo* ratio of bicyclic adducts derived from (**2**)
greater than that for the trimethylsilyl analogue, and
:ondly, there is a corresponding greater preference for (**4a**)
er (**4b**). Under certain low temperature conditions only
:lohexenone products (**4a,b**) and no bicyclic products were
tained. Hydrolysis of the hindered bicyclic silyl enol ether
ı) under relatively forcing conditions (AcOH–THF, 50 °C)
rely produced any retro-Michael product (**4b**). These
servations are consistent with (**3a**) and (**5a**) being derived
ım a [2 + 4] transition state in which the unfavourable steric interaction between the *N*-tosyl and t-butyldimethylsilyl groups is minimised,[5] while (**4a**) is derived from an open transition state in which the bulky silyl substituent avoids the steric clash with the ester group of the imine (**1**). Thus the proposed mechanism (c) in the previous paper,[1] involving a dual pathway explanation[6] for the origin of bicyclic and monocyclic products, seems to be supported in the present study.

We thank the S.E.R.C. for supporting this work, Merck, Sharp and Dohme, Harlow for a CASE award (A. B. T.), and Clare College for financial support (T. N. B.) and the award of the Denman Baynes Studentship (A. B. T.).

Received, 11th May 1988; Com. 8/01856A

References

1 T. N. Birkinshaw, A. B. Tabor, A. B. Holmes, P. Kaye, P. M. Mayne, and P. R. Raithby, preceding communication.
2 W. G. Dauben and R. A. Bunce, *J. Org. Chem.*, 1982, **47**, 5042.
3 F. Brisse, D. Thoraval, and T. H. Chan, *Can. J. Chem.*, 1986, **64**, 739.
4 A. J. Kirby, 'The Anomeric Effect and Related Stereoelectronic Effects at Oxygen,' Springer-Verlag, Berlin, 1983; K. L. Brown, L. Damm, J. D. Dunitz, A. Eschenmoser, R. Hobi, and C. Kratky, *Helv. Chim. Acta*, 1978, **61**, 3108; L. I. Kruse, C. W. DeBrosse, and C. H. Kruse, *J. Am. Chem. Soc.*, 1985, **107**, 5435 and references cited therein.
5 S. M. Weinreb and R. R. Staib, *Tetrahedron*, 1982, **38**, 3087.
6 S. J. Danishefsky, E. Larson, D. Askin, and N. Kato, *J. Am. Chem. Soc.*, 1985, **107**, 1246.

Figure II: A page (the second of two) from an article by T.N. Birkinshaw, A.B. Tabor, A.B. Holmes and P.R. Raithby, entitled "Imino Diels-Alder Reaction of 2-t-Butyldimethylsilyloxycyclohexadiene: Isolation of an Azabicyclo[2.2.2]octene Silyl Enol Ether Adduct", *J. Chem. Soc., Chem. Commun.*, 1601 (1988).

https://doi.org/10.1177/039219218903714702 Published online by Cambridge University Press

The Shapes of Molecules, and How they are Communicated

Shape matters in chemistry. Two molecules as subtly different from each other as a left hand is from a right, may have quite different physical, chemical and biological properties. Thus the mirror image of carvone, the main component of oil of spearmint, smells of caraway. The arrangement of atoms in space is not just a laboratory curiosity, it can be a matter of life or death. Thalidomide, a sedative of the early sixties, was responsible for thousands of fetal malformations. The pharmaceutical marketed was a mixture of left- and right-handed mirror image molecules. One form was teratogenic, causing malformation, its mirror image was not. Had this been known at the time, great anguish and human loss could have been prevented.[2]

Molecules are made up of atoms. But molecular structure is not just the identity of the atoms. Neither is it reducible to the interconnections of these elemental atomic building blocks in a molecule. At the operative level of modern chemistry, structure means the three-dimensional arrangement of atoms in space.[3] It is a graph, at the very least, a three-dimensional set of points connected by lines called bonds.

It is critical that chemists easily communicate this structural information among themselves. Via what's at hand, which are two-dimensional media—paper, a screen. The information is complex—many atoms, many bonds, a richness of geometrical structure. The information is at some important level inherently graphic—it is essentially a shape to be drawn. And now we come to the crux of the matter. The group of professionals to whom this visual, three-dimensional information is essential are not talented (any more, any less) at transmitting such information. Chemists are not selected, do not select themselves, for their profession on the basis of their artistic talents. Nor are they trained

[2] The story is more complicated. Under physiological conditions the harmless form transforms into a mixture of mirror images, one harmful, one not. See *Chemistry in Britain*, **25**, 259 (1989).

[3] The precise definition of molecular structure is still a subject of debate. See for instance R.D. Brown, *Chemistry in Britain*, **24**, 770, (1988) R.G. Woolley, *J. Amer. Chem. Soc.*, **100**, 1073 (1978) and the articles cited below by Mislow and Turro.

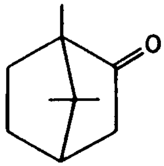
https://doi.org/10.1177/039219218903714702 Published online by Cambridge University Press

in basic art technique. The authors' ability to draw a face so that it looks like a face atrophied at age ten.

So how do they do it, how do we do it? With ease, almost without thinking, but, as we will see, with much more ambiguity than we, the chemists, think there is. The process is *representation*, a symbolic transformation of reality. It is both *graphic* and *linguistic*. It has a *historicity*. It is *artistic* and *scientific*. The representational process in chemistry is a shared code of this subculture.

Let us begin our look at the process by a look at the outcome. This was shown in Figure II, a typical page from a modern chemical article. The substancial amount of graphic content just stares one in the face. There are little pictures here. Lots of them. But the intelligent observer who is not a chemist is likely to be stymied. He finds himself in a situation analogous to that of Roland Barthes on his first visit to Japan, beautifully described in his *The Empire of Signs*.[4] What do these signs mean? We know that molecules are made of atoms, but what is one to make of a polygon such as structure **1**, here representing a white, waxy medicinal compound with a penetrating aroma, camphor? Only one familiar atomic symbol, O for oxygen, emerges.

1



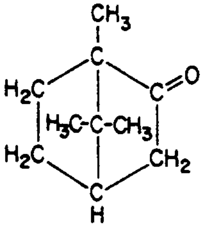
Well, it's a shorthand. Just as the military man gets tired of saying Commander in Chief, South Pacific Operations, and writes CICSPO, so the chemist tires of writing all those carbons and hydrogens, ubiquitous elements that they are, and draws the carbon skeleton. Every vertex that is not specifically labelled otherwise in structural representation **1** of camphor is carbon. Since the valence of carbon (the number of bonds it forms) is typically four, chemists privy to the code will know how many hydrogens

[4] R. Barthes, *L'Empire des Signes*, Geneva, Skira, 1980; *The Empire of Signs*, tr. R. Howard, New York, Hill and Wang, 1982.

https://doi.org/10.1177/039219218903714702 Published online by Cambridge University Press

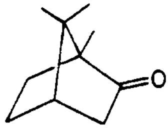
to put at each carbon. The polygon drawn above is in fact a graphic shorthand for structure **2**.

2



But is **2** the true structure of the molecule of camphor? Yes and no. At some level it is. At another level the chemist wants to see the three-dimensional picture, and so draws **3**. At still

3



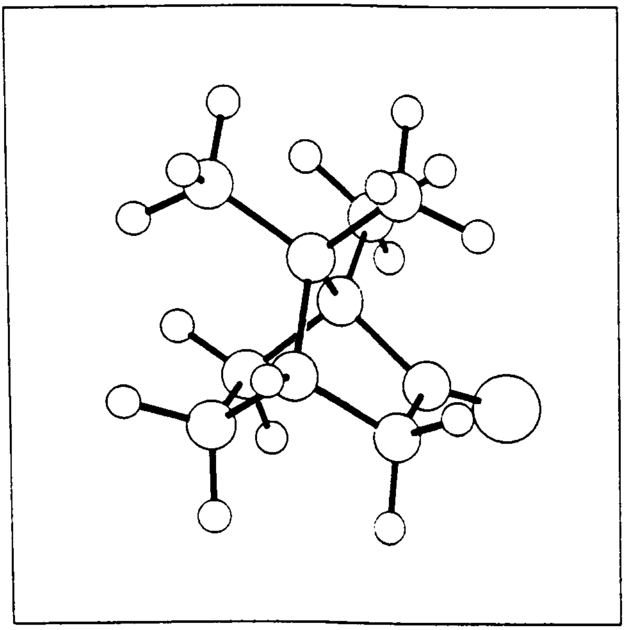
another level, he or she wants to see the "real" interatomic distances, i.e. the molecule drawn in its correct proportions. Such critical details are available, with a little money, a little work, by a technique called X-ray crystallography. And so we have a drawing **4**, likely to be produced by a computer.

This is a view of a so-called "ball-and-stick" model, perhaps the most familiar representation of a molecule in this century. The sizes of the balls representing the carbon, hydrogen and oxygen atoms are somewhat arbitrary. A more "realistic" representation of the volume that the atoms actually take up is given by the "space-filling" model **5**. Note that in **5** the positions of the atoms, better said of their nuclei, become obscured. And neither **4** nor **5** is portable. It cannot be sketched by a chemist in the 20 seconds that a slide typically remains on a screen in the rapid-fire presentation of the new and intriguing by a visiting lecturer.

The ascending (descending?) ladder of complexity in representation hardly stops here. Along comes the physical chemist to remind her organic colleagues that the atoms are not nailed down in space, but moving in near harmonic motion around those sites. The molecule vibrates; it doesn't have a static structure. Another

https://doi.org/10.1177/039219218903714702 Published online by Cambridge University Press

4

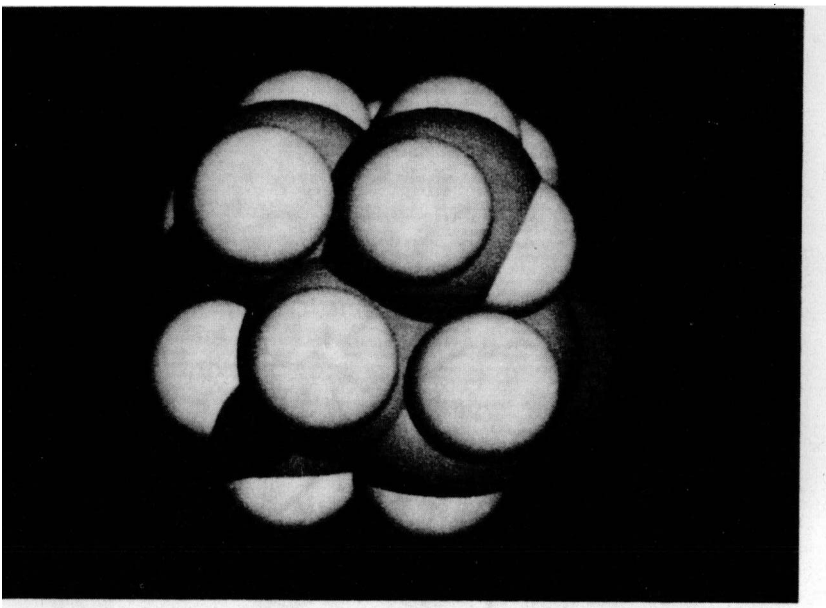


chemist comes and says: "You've just drawn the positions of the nuclei. But chemistry is in the electrons, you should draw out the chance of finding them at a certain place in space, the electronic distribution." As one tries to do in **6** and **7**.

We could go on. The literature of chemistry does. But let's stop and ask: Which of these representations **1** through **7** is 'right'? Which is the molecule? Well, all are, or none are. Or, to be serious—all of them are models, representations suitable for some purposes, not for others.[5] Sometimes just the name "camphor" will do. Sometimes the formula, $C_{10}H_{16}O$, suffices. Often it's

[5] For methodological discussions of how models are used in chemistry see C.J. Suckling, K.E. Suckling, and C. W. Suckling, *Chemistry Through Models*, Cambridge, Cambridge Univ. Press, 1978. C. Trindle, *Croat. Chim. Acta*, **57**, 1231 (1984); J. Tomasi, *J. Mol. Struct. (Theochem.)*, **179**, 273 (1988). And for the different meanings of "model" see N. Goodman, *Languages of Art*, 2nd ed., Indianapolis, Hacklett, 1976, p. 171.

https://doi.org/10.1177/039219218903714702 Published online by Cambridge University Press



5

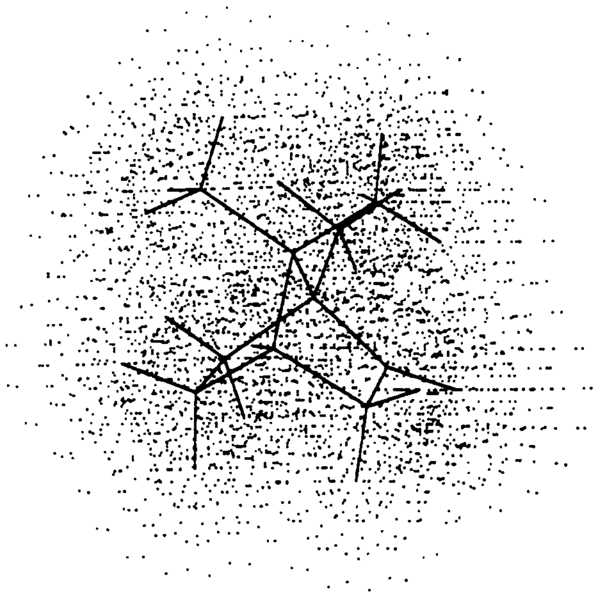
the structure that's desired, and something like **1** or **3** is fine. At other times one requires **4** or **5**, or even **6** or **7**.[6]

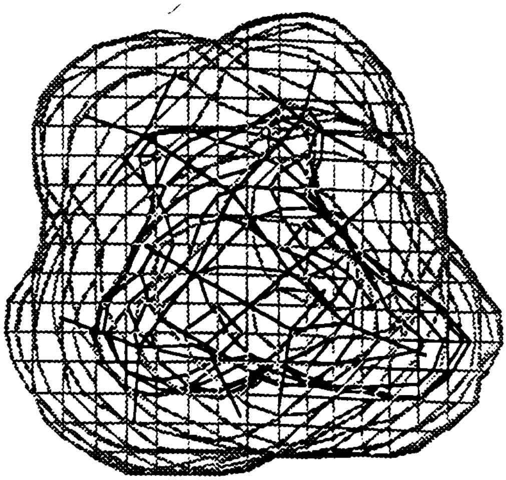
Let us fix on the *typical* level of presentation (of Figures 1 and 2), that of a polygon (**1**) or a three-dimensional idealization of it (**3**). But what *are* these curious constructions, these drawings, filling the pages of a scientific paper? We now ask the question from the point of view of an artist or draftsman. They're not isometric projections, certainly not photographs. Yet they're obviously attempts to represent in two dimensions a three-dimensional object for the purpose of communicating its essence to some remote reader.

The clues to three-dimensionality in these drawings are minimal. Some are conventional: here and there (one example in **3**) there

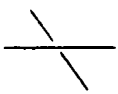
[6] That there are many ways to look at a molecule's structure is, of course, well known to the chemical community; we are not saying anything new here. See, for instance, G. Ourisson, *L'Actualité Chimique*, Jan.-Feb. 1986, p. 41.

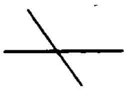
https://doi.org/10.1177/039219218903714702 Published online by Cambridge University Press

6 

7 

is a line "cut" to establish that some piece of the structure is in front of another, i.e. **8** instead of **9**.

8 

 9

https://doi.org/10.1177/039219218903714702 Published online by Cambridge University Press

Scattered about in the drawings of Figures I and II are sundry wedges and dashed lines. These are pieces of a visual code, simple in conception: a solid line is in the plane of the paper, a wedge in front, a dashed line in back. Thus **10** shows several views, all quite recognizable to chemists, of the tetrahedral methane molecule, CH_4. The tetrahedron is the single most important geometrical figure in chemistry.

10

Describing this notation may be enough to make these structures rise from the page for some people, but the neural networks that control representation are effectively etched in, for life, when one handles (in human hands, not in a computer) a ball-and-stick model of the molecule while looking at its picture.

A glance at the more complicated molecules of Fig. 2 shows that the wedge-dash convention is not applied consistently. Most compounds have more than a single plane of interest; what's behind one plane may be in front of another. So the convention is applied unsystematically, the author or lecturer choosing to emphasize the plane he or she thinks important. The result is a cubist perspective, a kind of Hockney photo-collage. The molecule is certainly seen, but may not be seen as the scientist thinks (in a dogmatic moment) that it is seen. It is represented as he chooses to see it, nicely superimposing a human illogic on top of an equally human logic.

Let us return again to the question of what these chemical structures represent, how they are drawn and read. Philostratus tells, much mythologized, the story of Apollonius of Tyana, a Pythagorean who lived around the time of Christ. In a dialogue with a disciple, Apollonius explores what painting is. It's done to make a likeness, to imitate. But what about cloud shapes in

https://doi.org/10.1177/039219218903714702 Published online by Cambridge University Press

the sky, read by us as horses or bulls? Are those also imitations? Apollonius and his disciple agree that these are but chanced configurations, that it is we who interpret those shapes, give them meaning. He continues "But does this not mean that the art of imitation is twofold? One aspect of it is the use of hands and mind in producing imitations, another aspect the producing of likenesses with the mind alone... I should say that those who look at works of painting and drawing must have the imitative faculty and that no one could understand the painted horse or bull unless he knew what such creatures are like".[7]

Knowing is not an unproblematical concept. How does that three-dimensional structure unfold in its full glory in the mind of a chemist? As we said, the direct images produced by contemporary techniques such as scanning tunnelling microscopy or electron microscopy (and these are not so "direct" on close examination) are few. Secondary knowledge, through X-ray crystallography, microwave spectroscopy or electron diffraction, is experienced by a small number of specialists. For most of us it is the real, physical handling of models that sets the stage—the analogy to seeing Apollonius' bull or horse in the first place. Or looking at many pictures of molecules drawn by others, assimilating thereby the set of conventions shared by chemists. It's much like art, and we will return to this below.

Chemical Representation as Language

Most, if not all, scientists make use of visual imagery for problem-solving, in order to sort out and organize information, to find analogies, to think.[8] But chemists are unusual among scientists

[7] F. Philostratus, *Life of Apollonius of Tyana*, Bk. II, Ch. 22, tr. F.C. Conybeare, New York, The Macmillan Company, 1912, I, 175-179. We owe this story to E.H. Gombrich, *Art and Illusion*, 2nd Ed., Bollingen Series, Princeton, Princeton University Press, 1961, p. 181-182.

[8] See R.R. Hoffman, "Some Implications of Metaphor for Philosophy and Psychology of Science", in *The Ubiquity of Metaphor*, R. Dirven and W. Paprotte, eds., Amsterdam, John Benjamin, 1985.

https://doi.org/10.1177/039219218903714702 Published online by Cambridge University Press

(but they share this with electronical engineers, architects, etc.) in having an iconic vernacular, that of the formulas.

A chemical formula is like a word. It purports to identify, to single out the chemical species it stands for. Chemical formulas embody the ancient dictum (going back, as far as teaching is concerned, to the Czech Comenius at the time of the Renaissance): "One thing, one word?"[9] Indeed, chemical systematics take great care to avoid ambiguous situations. One such is that of two compounds with the same composition, say HCN and HNC (H stands for hydrogen, C for carbon, and N for nitrogen). These two *isomers* (the term was coined by the Swede Berzelius in 1830) differ in the arrangement of the atoms: in the former, carbon lies in-between hydrogen and nitrogen; in the latter, nitrogen is the middle atom. Isomers may have dramatically different physical, chemical and biological properties (HCN is the notorious hydrogen cyanide). The language of chemistry, to a very large extent, is a nomenclature. Problems remain nevertheless in the perception of the entities of chemistry. Chemistry is a mature science. It has shed to a large extent its infantile habit of going no further than a phenomenological description of bulk properties, at the macroscopic level (that of sensory perceptions), accompanied by an apt denomination (such as "potash"—because the compound was first found, literally, in pot ash).

Chemistry has become a microscopic science. Explanations nowadays go routinely, paradigmatically from the microscopic scale to the observable scale: from the way the electrons are distributed in a dye molecule to its color; from the detailed shape of a molecule and of the electrostatic potential around it to its pharmacological activity; from bond energies to the tensile strength of a new polymer, such as Kevlar, sturdier than steel and very light.

Such a seminal characteristic of chemical sciences was noted already at the time of the Enlightenment. The entry for "chemistry" in the *Encyclopédie* by Diderot and d'Alembert points out that chemists (but not the Newtonian physicists of the time, who explained everything with central forces acting on material points)

[9] J.A. Comenius, *Orbis Sensualium Pictus*, Nuremberg, Michael Endter, 1658.

https://doi.org/10.1177/039219218903714702 Published online by Cambridge University Press

are wont to posit invisible and intangible entities or qualities to explain observations. They are forced to do so by the smallness of their particles. We are still under a similar constraint, even though "images" of some molecules and of some heavy atoms have become available very recently. Indeed we know our building blocks, molecules, much better. But it remains a long, long way from the molecular scale to the macroscopic world of the senses. We still have to represent molecules. And we tend to represent to ourselves atoms *as if* they were normal objects in our everyday experience: with a size, with a certain hardness or softness, with measurable attractions to other atoms or to electrons, and so on. This is a little naive, unavoidable, and endearing—not unlike a belief in angels in past centuries.

True, chemical formulas have been severed from subjective life experiences to a considerable extent, much more so than words such as "dog" or "automobile". Yet, despite such an excision, chemical structures retain a strong connection with sensory experiences; by contrast with such mundane words, they carry an essential *representational* component.

Let us elaborate some on this interesting paradox. As Ferdinand de Saussure pointed out, in one of his fundamental insights, the *word-thing* relationship is that of the *signifier* to the *signified*; and a key notion is that of the *arbitrariness* of the signifier.[10] When I say "dog" (or "*Hund*" or "*cachorro*"), each of these terms has been selected more or less randomly in the history of the language. It could have been a "snark" or a "livel" or a "rop" in English, who knows! The word "dog" has settled into its niche (forgive the pun) in the common language by way of the numerous cross-relations it bears with other words in the dictionary (such as "leash", "dogfood", "bitch" or "seeing eye", in the example chosen). We flesh out a word because we cannot help loading it with our private empirical experience, just as we do with atoms. But in actuality the word is a total fabrication. It could have been an entirely different choice.[11]

[10] J. Culler, *Ferdinand de Saussure*, Rev. Ed., Ithaca, N.Y., Cornell University Press, 1986.

[11] K. Mislow brought to our attention in this context the story Feynman tells of how his father taught him the difference between knowing the name of something

https://doi.org/10.1177/039219218903714702 Published online by Cambridge University Press

With this reminder, we return to this paradox: formulas forsake reality in a sense, and they aim to stand for reality in another. Indeed, chemical structures differ from words in the normal language because they combine symbolic and representational (iconic) values. Take the case of the molecule of natural gas, methane, drawn above in structure **10**. The various types of strokes in **10** indicate *both* connection of the carbon with the hydrogens via chemical bonds (the symbolic statement), and whether the bond buts out of the plane of the paper when wedge-shaped, or, conversely, whether the bond recedes from the viewer to the backside, if it is shown dashed (the representational or geometric statement). The chemical formula is trying to signify a lot with the utmost economy in graphics. It aims at portraying accurately the connectivity of atoms—the nearest neighbor relationships as stemming from chemical bonds—and the geometry.

Thus, we can state the second problem with chemical formulas: they are in-between *symbols* and *models*. This hybrid status is an uneasy one. These two poles pull formulas toward opposite and sometimes incompatible requirements.

As models, chemical formulas ought to be reliable and accurate representations of what one might term "molecular reality" (we shall return to this useful figment of the collective psyche of chemists). As symbols, they ought to be arbitrary, to a large extent. This is reflected, to give an all-too-familiar example, in the use of a capital O to represent an oxygen atom: behind this universal convention, there is an assumption of transferability (an oxygen atom *here* is very much like an oxygen atom *there*) which is not that easy to put on a firm theoretical basis. An oxygen atom free of a molecule is indistinguishable from any other oxygen atom. But inside a molecule, an oxygen atom bound to a single carbon (what a chemist would call a ketone or aldehyde) has some very different properties from one bound to two carbons (an ether).

Is the world of chemistry and chemical structures unambiguous, characterized by one drawing, one molecule? To some extent chemistry is such, because it is "a sign imprinted in matter",

and knowing something: R.P. Feynman, *What do You Care What Other People Think?*, New York, W.W. Norton, 1988, p.13, 14.

https://doi.org/10.1177/039219218903714702 Published online by Cambridge University Press

as J.-M. Lehn has called it.[12] The instructions for making aspirin work here as well as in Montevideo and Karaganda. The article reporting the synthesis of a new drug perhaps needs translation, but in another sense it doesn't, for it is understood around the world, it is infinitely paraphrasable. Chemical structures, chemical formulas are the signing tools of this language.

But in another way the graphic language of chemistry is quite ambiguous. We've seen clear evidence for this in the plurality of answers given for camphor to the simple question: "Draw for me the structure of the molecule". It could be argued that once drawn, no matter how drawn, the single molecule "camphor" enters the mind of the chemist. But as we will see later, how a molecule is thought about and subsequently manipulated in the material world is very much influenced by the way we carry it around in our minds.

A chemical formula is at once a metaphor, a model (in the sense of a technical diagram), and a theoretical construct. A chemical formula is part pure imagination, part inference. It is an attempt to depict the real by manipulation of symbols, just as language enables us to talk about the world and about ourselves by combining arbitrary utterances. The simile cannot be pushed too strongly. In a deep philosophical sense, calling something "acidic" and calling something else "red" are identical mental operations. Likewise, referring to "ethanol" or "reserpine" is akin to talking about "Rockefeller Center" or about the "Eiffel Tower".

"Acidic" and "red" are ill-defined concepts relating to everyday experiences. Conversely, "ethanol" and "Eiffel Tower" are defined unambiguously, but they need not be common-place objects; they are cultural objects in the widest sense.

If "ethanol" is in a similar mental category as "Eiffel Tower" the chemical formula of ethanol stands to it not unlike a dictionary definition for "Eiffel Tower": a chemical formula is a concise paraphrase, in a half-symbolical, half-iconic language, of some attributes of an object, so that the object be properly and unambiguously identified, *i.e.* differentiated from like objects (Eiffel Tower as distinguished from the Madeleine or the

[12] J.M. Lehn, *Traduire*, **116**, 62 (1983).

https://doi.org/10.1177/039219218903714702 Published online by Cambridge University Press

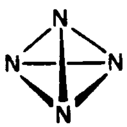
Centre Pompidou; ethanol as distinct from ethane or from acetic acid).

Language and chemical representation, besides their joint use of names, have other similarities. They share use of invariant elements. Many words and most chemical compounds are just that, compounds, put together by association of structural fragments. There is deep similarity between a word such as "one-upmanship" and a "chemical word" such as C_6H_5-CH_2-CO-OH. In the latter case, just as in the former, the structural fragments, known (in chemical language) from left to right as phenyl, methylene, carbonyl, hydroxy, are stable semes: to a first approximation they retain their basic meaning whatever the nature of the other modules they are connected with.

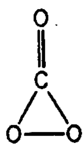
Chemistry is the science of change, of transformations. Every science starts with axioms about the integrity of certain of its objects. The mechanical engineer believes in the integrity of a steel girder. The cell biologist believes in the integrity of chloroplasts or mitochondria: he or she is convinced that these organelles are interchangeable, playing identical roles in one cell or in another.

For chemistry, it is crucial to make sense of change by constraining it to occur between well-defined states. Invariance and its equivalent, transferability, are basic assumptions to chemistry. At each level of understanding (or complexity), the lower units are set as invariant: starting with atoms, going on to the simple structural fragments such as the above (C_6H_5, CH_2, CO, OH), on to simple molecules, further on to chains or polymers, the helices of proteins and nucleic acids, and so forth.

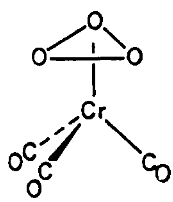
The concept of transferability goes back, beyond Dalton's atoms, to the Lavoisier revolution: *"Rien ne se perd, rien ne se crée"*. His use of the balance made Lavoisier discover the physi-



11



12



13

https://doi.org/10.1177/039219218903714702 Published online by Cambridge University Press

cal law of mass conservation: but the linguistic bent he had inherited from Condillac made him give a *linguistic* expression to it. Lavoisier founded modern chemical language on the explicit analogy to natural language. No wonder that chemical formulas, to this day, retain an important linguistic component.

It is possible to set up a formal relationship between chemistry and language; an initial step in this direction has already been taken by H.W. Whitlock, Jr.[13] In the terminology of Chomsky, a language is defined by a set of symbols (a vocabulary) from which strings (sentences) may be generated by a set of productions or transformations (rules for making changes).[14] The identification of symbols with chemical elements or those simple structural fragments (CH_2, CO, OH) we have alluded to above, and of productions with chemical reactions, is obvious. Whitlock interestingly shows that a certain problem in organic synthesis may be approached by analyzing it in the context of formal language.

The shared productivity of language, formal or natural, with chemistry applies also to the realm of what has not yet been said, or written, or synthesized. There exist rules, for instance such that we have the competence to pronounce a word that we have never heard: "a roor", "to roat", "the poot". Likewise, we can write utopian but plausible formulas for (so far) unknown chemical species (For example, **11-13**). This serves quite often as an inducement to try and prepare them.

These structures, waiting impatiently to be made, have for chemists the incongruous look, both attractive and shocking, of a novel object deemed by some as an impossibility: not unlike a first look at a laser printer in action, or a unirail train levitating above its magnetized track, a Stealth bomber, or a town-size space station...

[13] H.W. Whitlock, Jr., in "Computer-Assisted Organic Synthesis", ed. W. T. Wipke and W.J. Howe, American Chemical Society (ACS Symposium No. 61), Washington, 1977, p. 60. We thank K. Mislow for reminding us of this work.
[14] N. Chomsky in *Handbook of Mathematical Psychology*, ed. R.D. Luce, R.R. Bush and E. Galanter, Vol. 2, **323** (1963); N. Chomsky, *Cartesian Linguistics*, New York, Harper & Row, 1966.

https://doi.org/10.1177/039219218903714702 Published online by Cambridge University Press

Thus both languages, the natural and the chemical, witness an evolution of meaning, from so-called nonsense to highly significant statements. Even though it does not build upon known ("domesticated") words, and it invents instead new ("wild") words, the "Jabberwocky" poem by Lewis Carrol *makes sense* because of its impeccable syntax, which lets the imagination of the reader both be charmed by the word-play, and invest some of the words with meaning, by way of various associations. Chemistry has likewise its wild species (benzyne, **14**, tetrahedrane, **15**, or **11-13** above), besides, its more usual bottled samples. In fact most of the molecules of chemistry, wild dreams

14 **15**

or not, are invented, synthesized. They were not on earth before. This is similar to the state of words, where only onomatopeias and circumlocutions pre-existed their verbal invention.

Chemical representation is also language-like when dealing with the class of transformations known as chemical reactions. In an important sense, chemistry is the skillful study of symbolic transformations applied to graphic objects, the formulas. There are one-to-one correspondences between compounds and formulas. Likewise, it is possible to specify unequivocally the transform from one compound to another. If one thinks of chemical compounds as nodes in an infinite and multidimensional grid, then the connecting lines in such a network are the transforms. Chemistry has thus two facets, structural when the focus is static, on the points in the grid; and dynamic when what is examined is the interconversion along the edges in the network. There are strict rules about *rewriting* formulas to express transforms; not unlike rules of musical composition.

Chemical transforms are the analogs of action sentences in natural language. Just as action sentences carry subordinates as modifiers, to provide information about time, location, quanti-

https://doi.org/10.1177/039219218903714702 Published online by Cambridge University Press

ty, manner, so the chemical equation is wont to specify the solvent, the reaction temperature, the reaction time, the yield of product, etc...

In another sense, this parallel is too general. The product molecules in a chemical reaction contain the same atoms as the reactants, but reconnected in a different way. The relationship of subject and object is different—except in a category of sentences having a pronominal (or reflexive) verb. A sentence such as "Jane washes herself" or "John admires himself in the mirror" are somewhat like statements about chemical transforms. Funny that chemical equations come so close to psychological statements!

In Goethe's 1809 novel *Die Wahlverwandtschaften* ("Elective Affinities"), a by then outdated theory of chemical combination powers a work of fiction.[15] The actions and emotions of the characters of this work embody (and probe critically) the way some people thought molecules behave. One wishes one were able to point today to a similarly inspired literary text. Nevertheless, it is interesting to reflect on the deep morphological resemblance of chemistry and if not life, at least language. We do not have the space here to elaborate upon this statement: chemists take a kind of narrative approach to chemical structures. Two chemists will "read" a complex structure in similar manner.

It is thus our contention that chemical formulas are read according to conventional sequences. Recent research in neuropsychology teaches indeed that mental patterns do not spring up whole. They are built up gradually, a part at a time; and the parts, it is found, are visualized in roughly the same order as they are typically drawn. With respect to the two types of tasks (retrieval of archival shapes, and coordination of shapes into a mental image), the two brain hemispheres appear to play different roles.[16,17]

[15] J.W. Goethe, *Die Wahlverwandtschaften*; *Elective Affinities*, tr. J.A. Froude and R.D. Boylan, New York, F. Ungar Publ., 1962.
[16] See S.M. Kosslyn, *Science*, **240**, 1621 (1988).
[17] For an illuminating serie of articles on biological aspects of aesthetics see I. Rentscher, B. Herzberger and D. Epstein, Eds., *Beauty and the Brain*, Basel, Birkhäuser, 1988.

https://doi.org/10.1177/039219218903714702 Published online by Cambridge University Press

The opposite viewpoint is a necessary complement. As Verbrugge indicates, discussing the example of Kekulé's architectural formulas, his ring structure for benzene, and the oscillation of benzene between the two equivalent ring forms, "scientific understanding develops only when we are prepared to reshape our representation systems in fundamental ways".[18]

We submit that the combined pressures of (I) the learning, early on, of chemical nomenclature; (II) incessant on-the-job confrontations with formulas through seminars, the reading of publications, the handling of molecular models; and (III) the demands of communication with other chemists, have built this largely unconscious and stereotyped collective way of seeing. Probably art historians act the same; it is quite possible, even likely, that two specialists of Quattrocento painting will both scan a picture in much the same ways.

Let us now explore the opposite viewpoint. Because perception of chemical shapes is so stereotyped, conversely to be able to see a structure in a novel way can be extremely fruitful. Chemistry shares with poetry its notion of elegance, its mission so to say. This is to discover new relations between objects. Very often, the elegance of a key step in a series of chemical transformations is rooted in the perception of a non-obvious connection between parts of the molecular object. More generally, it would repay the student of psychological invention to take a close look at flowsheets, the sequence of molecules made and transformed, in synthesis of natural products. Indeed the synthetic elaboration of complex chemical structures offers fascinating glimpses into the creative process in the doing. In the hands of a master craftsman such as R.B. Woodward, structural fragments experience what are to the mind—that of the conceiver as well as that of the reader—genuine *Gestalt* shifts, from one part of the synthesis to another. By studying the Woodward syntheses in detail, one can just see him turning a molecule over in his mind, and seeing some of its elements from new angles. The changes of the structures in the course of a Woodward synthesis are metaphors of creativity. They exploit the plurality of meanings embodied

[18] R.R. Verbrugge, *Ann. N. Y. Acad. Sci.*, **433**, 167 (1984).

https://doi.org/10.1177/039219218903714702 Published online by Cambridge University Press

in a chemical structure. We submit that it may be difficult to move closer to the creative imagination in action.[19]

BUT IS IT ART...?

Art or the reaction to it, the aesthetic response, has never been easy to define. There are so many forms of pleasing human creation, so many constructed objects or patterns evoking emotional reactions... Cognizant of the complexity and venerable history of aesthetics, let us hazard a definition.[20,21] While it is one contestable in all of its parts, perhaps it touches on most of the qualities of what we have chosen to call art. Then we will examine representation in chemistry as it measures up against this definition.

Let us call Art those symbolic acts or creations of human beings which aspire to the extraction from the complex realm of Nature, or the equally involved world of the emotions, of some aspect of the essence of these worlds. Art functions by communication of a symbol, meant to convey information and/or evoke an emotional response.

The essential components of the aesthetic system are (a) the creator—painter, composer, photographer, writer, dancer, (b) the audience—both that perceived in the creator's mind and the real one, the viewers, (c) the set of symbols by which communication takes place—the watercolor, sound waves and images evolving

[19] For a presentation of the marvels of organic synthesis, see N. Anand, J.S. Bindra and S. Ranganathan, *Art in Organic Synthesis*, 2nd Ed., New York, John Wiley and Sons, 1988; E.J. Corey and X.M. Cheng, *The Logic of Chemical Synthesis*, New York, Wiley & Sons, 1989. Ms. Crystal Woodward has written a psychobiography of her father. We have been privileged to read this yet unpublished manuscript. Some of its perceptive insights bear on the topics we discuss. One of us (P.L.) refers to her analogy of molecules and molecular models with the *transitional objects* of D.W. Winnicott in a book, *La parole des choses*, soon to be published.

[20] To get a feeling for the complexity of definitions and the range of opinions in this field see M.C. Beardsley, *Aesthetics*, 2nd ed., Indianapolis, Hacklett, 1981.

[21] Related to this section are a series of articles one of us has written on "Molecular Beauty". This is a kind of anthropological study of the objects chemists admitted as possessing aesthetic value: (R.H.) *Amer. Sci.*, **76**, 389, 604 (1988); **77**, 177, 330 (1989).

https://doi.org/10.1177/039219218903714702 Published online by Cambridge University Press

in time, a text, (d) the act of communication itself—to an audience that is present (watching a dance, listening to a cantata) or absent (reading a novel).

If this definition sounds too rational, too "scientific", devoid of the gut emotional response we should like good art to hit us with, so be it. Nelson Goodman argues persuasively that "in aesthetic experience the *emotions function cognitively*", that feeling is knowing. He goes on to examine the usual attributes of art (that the aesthetic is directed to no practical end, that it gives immediate satisfaction, that in art inquiry is a means of obtaining satisfaction... all of these existing as marks of art mainly by contrasting them with an opposite attributed to science), and he finds them wanting. He says "...the difference between art and science is not that between feeling and fact, intuition and inference, delight and deliberation, synthesis and analysis, sensation and cerebration, concreteness and abstraction, passion and action, mediacy and immediacy, or truth and beauty, but rather a difference in domination of certain specific characteristics of symbols".[22]

To return to the question heading this section: are chemical structures art? It seems clear that they possess all the components of the "aesthetic system". Structural formulas are symbols created by one chemist (or several) to communicate information to others. The drawing of the structure of camphor is certainly a symbolic motion, a communication of an essence—the arrangement in space of the atoms of this molecule. Some might call it just a sketch, an information reducing stratagem by someone not able or willing to compute and show others the all-important electron density around the nuclei. *That* electron density is the real molecule; the structural formula—well, that's "just a poor representation".

The chemical structure is an artistic construct because it is a transformation of a model of reality (note the second-hand if not n-th hand relationship to the real) for the purpose of communication. Neither chemistry texts nor anatomy books are much illustrated with photographs. There are some photographs, to be sure. But by and large a photograph (and we certainly don't wish

[22] N. Goodman, *Languages of Art*, 2nd. Ed., Indianapolis, Hackett, 1976, 264.

https://doi.org/10.1177/039219218903714702 Published online by Cambridge University Press

to imply that photography is mere representation; it is far from that...) contains too much detail. What one wants to communicate is the essence needed for the moment. One wants to teach, to evoke a response. Drawings, with their artistically selected detail, are much better for that purpose.[23]

The symbolic nature of the chemical formula, the fact that chemists know that a hexagon stands for a ring of carbon atoms that in turn is much, much smaller, the implicit knowledge that hexagon is *not* an enlarged photograph of the ring, all that symbolic distancing of course enhances the metaphorical nature of the chemical discourse. Structures are not what they stand for; they stand for what they are not.

But is it art...? Having argued that chemical representations, such as structural formulas, share all the symptoms of art, let us take the opposite tack, at least for a while.

Just as it is impossible to ignore the artist and his audience, the mental set of both, so it is impossible to put out of mind the context of a picture or a scientific illustration. And chemical structures, in particular, are often really part of the text. Oh, they may have artistic value or expressive power on their own. One could mount a good art exhibit around them. But their function, their organizing referent, is the text.

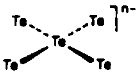
By way of illustration, Fig. III shows the beginning of a paper written by one of us.[24] Note four little drawings in the first two paragraphs. They are typically floating in mid-air, typically using the wedge-dash-line notation. The structures are numbered boldface, and referred to specifically in the text. In fact they are part of the text, and that they are referenced to (by their boldface numbers) even in the middle of a sentence, confirms this role. In one chain (labelled **2** in Figure III) the units of five telluriums are related by a so-called screw-axis, whereas in another chain (labelled **3** in Figure III) they are simply translated, one relative to the other. All that could have been said, one supposes, in words. But it was easier to draw a picture. So a structure was born.

[23] See in this context, the article by C. Rose-Innes, *New Scientist*, 7 Jan. 1988, p. 42.

[24] J. Bernstein and R. Hoffman, *Inorg. Chem.*, **24**, 4100 (1985).

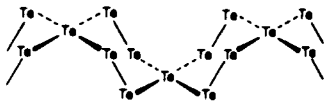
https://doi.org/10.1177/039219218903714702 Published online by Cambridge University Press

The recent literature contains a number of examples of a square-planar tellurium structural unit which may formally be defined as Te_5^{n-} (**1**). In some cases such as Rb_2Te_5[1] and Cs_2Te_5[2]

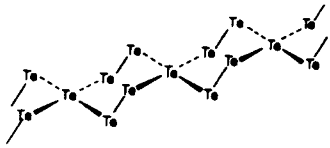


1

the stoichiometry clearly defines the charge on the unit as 2−. In these instances the unit is the basic building block of a one-dimensional anionic chain. In Cs_2Te_5 the Te_5^{n-} units are screw axis related (**2**) while in Rb_2Te_5 the Te_5^{n-} units are related by translation (**3**).



2



3

For most of the other cases in which the Te_5^{n-} unit appears, the attribution of a particular charge to the unit is more ambiguous. For instance, on K_2SnTe_5,[3] the unit again appears embedded in a one-dimensional chain (**4**) but the Te_5^{n-} units alternate

4

with tetrahedral Sn in making up the chain. Formally, at least, the square-planar units could still be considered Te_5^{2-} if the tin is Sn(0), but the Te–Sn bond length of 2.74 Å is almost identical with the sum of the tetrahedral covalent radii (2.72 Å) given by Pauling.[4] This suggests a formal oxidation state of 4+ for the Sn and a net charge of 6− on the Te_5^{n-} unit.

Figure III: The beginning of a paper, "Hypervalent Tellurium in One-dimensional Extended Structures Containing Te_5^{n-} Units", J. Bernstein and R. Hoffmann, *Inorg. Chem.*, 24, 4100 (1985).

https://doi.org/10.1177/039219218903714702 Published online by Cambridge University Press

Actually, the motives of the representation are not so simple here. The authors use the pictures not only to save space. They are also plumbing strategies to capture their audience. These tellurium compounds are not obviously interesting to everyone in chemistry. Specialization is a plague of any field of scholarship. The geometry of the tellurium structures are difficult to see. Since in science, as everywhere else, what we do not understand we are afraid of, what we do not understand we find uninteresting, the authors are using the visual appeal (and density of explanatory power) of a drawing to inform, to pull in, to attract, to seduce. Still another motive: it is the "style" of one of us to decorate his papers with such drawings. He is establishing a visual signature.

There is still another argument against assigning full artistic value to a chemical structure. It is not, to use Goodman's terminology, "replete".[25] Not every stroke in the representation matters; the lines could be a different color, the molecule readily recognizable as what it is even if drawn from a somewhat different viewpoint. This is in contrast to a Goya etching, which, were similar changes made in it, would be another Goya etching, a different work of art, or at the very least a different "état".[26] To put it in another way, the chemical structure addresses the inherent paraphrasability of scientific knowledge (paraphrase as one of the (few) differentiations of art and science has been persuasively forwarded by G. Stent[27]). It is the same molecule, intended to be the same, perceived to be the same. But what if the slightly different representation enters the unusual mind, prompting an experiment untried before? The icon powers the iconoclast.

Perhaps another way to approach the artistic content of chemical drawings is to think of their relationship to various visual art genres. For instance we see a similarity between chemical structures and what has been called, not without controversy, primitive art. Tribal art, be it of Australian aborigines or of Eskimos, often appears schematic to us, deficient in perspective. That's our

[25] N. Goodman, *op. cit.*, p. 229, 230.

[26] This point was made to R.H. by H. Pardee.

[27] G. Stent, *Engineering and Science* (Calif. Inst. of Technol.), September 1985, p. 9; Nobel Symposium, Royal Swedish Academy of Sciences, 1986, private communication.

https://doi.org/10.1177/039219218903714702 Published online by Cambridge University Press

problem, for to the native group which shares the culture that informs that art, the representation may be highly accurate and perceptive. So if it is important to show that someone owns six sheep, the sheep will be posed so that they are distinct, and clearly seen as sheep. A deity will be represented by its symbolic attributes so that we cannot confuse it with any other. So it is with chemical drawings—their perspective may be inadequate, their representation artistically unsophisticated. But they tell a concise story to the chemical reader. Like primitive images or sculptures, chemical drawings will usually distort a view if the viewers' ability to clearly *classify* an object is enhanced by that distortion.

The iconic representation of camphor (see drawing **4**) is simplified and distorted (compare with **5** or **6**) so as to allow us to identify the molecule, to trigger a connection in the mind. The Aivilik drawing (Figure IV[28]) of seal hunting through the ice rearranges space (top view of sled, side view of hunter, ice, and seal) and time (the seal approaching, the harpoon already thrown) so as to render visible, in all its richness, the hunt.

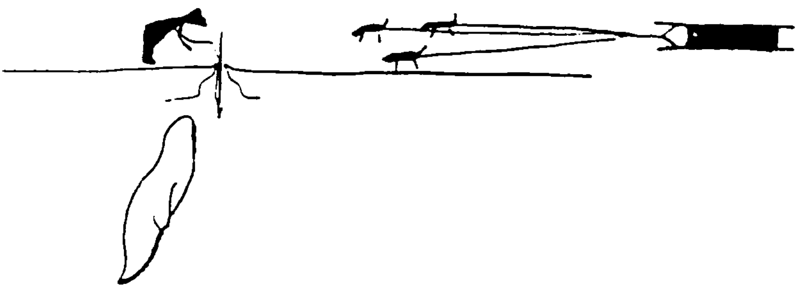


Figure IV: An Aivilik drawing of a seal hunt. From E. Carpenter, *Eskimo Realities*, Holt, Rinehart and Winston, New York, 1973, p. 177.

Something similar occurs in the pictorial language of chemists. If a piece of a molecule, some functionality such as a CHO, aldehyde, group is essential, even if it is hidden behind another part of the molecule, it will be brought forward without much regard

[28] For an insightful, beautifully presented account of Eskimo ways of seeing and representing, see E. Carpenter, *Eskimo Realities*, New York, Holt, Rinehart and Winston, 1973.

https://doi.org/10.1177/039219218903714702 Published online by Cambridge University Press

to faithfulness of representation. There is still another anthropological point of contact. In some cultures, knowing the true name of an object or a person forms a special bond, even gives power. So it is in chemistry. Knowing the "name" of a compound, which means its structure, gives the chemist tremendous power over the molecule. A range of its properties, its behavior, are implied by that structure.

G. Ourisson, in a thought-provoking article that deals with many of the same issues we have discussed, identified the chemical structure as an ideogram or pictogram, a symbol that represents an idea or object directly (this point was made also by R. Etiemble). He makes the analogy to Chinese characters, and perhaps one could also do so productively to some among the Egyptian hieroglyphs. Like the character, the chemical sign enters the conscience of a chemist directly. All its meanings are attached, and the chemist manipulates that little picture mentally in a multiplicity of ways. The chemical structure implies not only a molecule but its physical, chemical, even biological properties.

The drawing of chemical structures also has a kinship with caricature and comic strips. If one examines a successful cartoon or schematic book illustration closely one finds that a wide spectrum of emotion—grief, terrible anger, ecstasy—is communicated in just a few strokes of a pen. Think of the Dr. Seuss books, Jean de Brunhoff's Babar, Hergé's Tintin books, Tove Jansson's

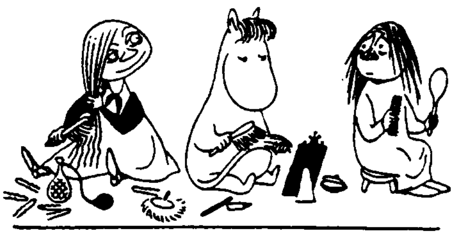


Figure V: An illustration from Tove Jansson's *Moominsummer Madness*, Henry Z. Walck, New York, 1955, p. 53.

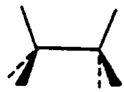
https://doi.org/10.1177/039219218903714702 Published online by Cambridge University Press

Moomintrolls (see Fig. V[29]), or Walt Disney's numerous cartoon characters. And not just for children. Gombrich discusses the effect of caricature perceptively, arguing that we ..."accept the grotesque and simplified partly because its lack of elaboration guarantees the absence of contradictory clues".[30] In examining a work of art we look at the information in it, and unconsciously for relationships. It is not the absolute flux of light entering our eye from a painted white spot that makes us see it as brightly lit, it is its differentiation from neighboring patches of paint.

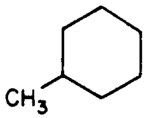
The act of viewing is collaborative (between painter and viewer) and forgiving. We always create space, in our minds, when we see a two-dimensional representation. And we elaborate the information, interpolating our experience to fill out what is omitted. At least providing there are no contradictory clues, no signals to tell us that we are wrong. Caricatures or cartoons (or if those are not "serious" enough, take a Goya or Picasso etching) work by providing the appropriate minimal information.

So do chemical structures. The chemist's mind is very tolerant. It will accept both **16** and **17** as representation of ethane. It will read the substituent attached at lower left to the six-membered ring in **18** through **21** as a methyl, or CH_3 group.

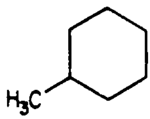
16



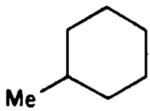
17



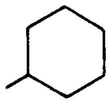
18



19



20



21

[29] T. Jansson, *Moominsummer Madness*, New York, Henry Z. Walck, 1955, p. 53.

[30] E.H. Gombrich, *op. cit.*, Chap. 10; E. Kris and E.H. Gombrich in E. Kris, *Psychoanalytic Explorations in Art*, New York, International Universities Press, 1952, Ch. 7.

https://doi.org/10.1177/039219218903714702 Published online by Cambridge University Press

It will fill in the missing three dimensional background required to make these molecules come to life; it will geometrize the floating world of these little symbols.

In any case, the drawing does not exist by itself, but is an integral part of the text. It is as if the chemist invented a new language, part text, part picture, part the tactile sense in model-building. These explain the ability of a chemist to reconstruct a molecule in his or her mind from such minimal information. We are pushed back to the logic of language. And in another way, to viewing the combined text-structure complex as an art form.

R. Hoffmann and P. Laszlo
(Cornell University, Ithaca, N.Y.
and
École polytechnique, Palaiseau)

Acknowledgment.

We have benefited much from the critical comments of Noel Carrol, Gilles-Gaston Granger, Kurt Mislow, Guy Ourisson, Hearne Pardee, Vivian Torrence. The assistance of Dennis Underwood, Donald Boyd and Laura Linke is also gratefully acknowledged, as are the photographs of Stephen R. Singer, the drawings of Jane Jorgensen and Elisabeth Fields, and the typing of Joyce Barrows.

https://doi.org/10.1177/039219218903714702 Published online by Cambridge University Press