How to cite: *Angew. Chem. Int. Ed.* **2020**, *59*, 12590–12610
International Edition: doi.org/10.1002/anie.201902527
German Edition: doi.org/10.1002/ange.201902527

Quantum Chemistry

Simulation vs. Understanding: A Tension, in Quantum Chemistry and Beyond. Part A. Stage Setting

Roald Hoffmann and Jean-Paul Malrieu**

Keywords: quantum chemistry · simulation · theory · understanding

Abstract: *We begin our tripartite Essay with a triangle of understanding, theory and simulation. Sketching the intimate tie between explanation and teaching, we also point to the emotional impact of understanding. As we trace the development of theory in chemistry, Dirac's characterization of what is known and what is needed for theoretical chemistry comes up, as does the role of prediction, and Thom's phrase "To predict is not to explain." We give a typology of models, and then describe, no doubt inadequately, machine learning and neural networks.*

In the second part, we leave philosophy, beginning by describing Roald's being beaten by simulation. This leads us to artificial intelligence (AI), Searle's Chinese room, and Strevens' account of what a go-playing program knows. Back to our terrain—we ask "Quantum Chemistry, † ca. 2020?" Then move to examples of AI affecting social matters, ranging from trivial to scary. We argue that moral decisions are hardly to be left to a computer.

At this point, we try to pull the reader up, giving the opposing view of an optimistic, limitless future a voice. But we don't do justice to that view—how could we? We return to questioning the ascetic dimension of scientists, their romance with black boxes.

Onward: In the 3rd part of this Essay, we work our way up from pessimism. We trace (another triangle!) the special interests of experimentalists, who want the theory we love, and reliable numbers as well. We detail in our own science instances where theory gave us real joy. Two more examples-on magnetic coupling in inorganic diradicals, and the way to think about alkali metal halides, show us the way to integrate simulation with theory. Back and forth is how it should be—between painfully-obtained, intriguing numbers, begging for interpretation, in turn requiring new concepts, new models, new theoretically grounded tools of computation. Through such iterations understanding is formed.

As our tripartite Essay ends, we outline a future of consilience, with a role both for fact-seekers, and searchers for understanding. Chemistry's streak of creation provides in that conjoined future a passage to art and to perceiving, as we argue we must, the sacred in science.

Extended Table of Contents

---

[*] R. Hoffmann
Dept. of Chemistry and Chemical Biology, Cornell University
Ithaca NY 14850 (USA)
E-mail: rh34@cornell.edu

J.-P. Malrieu
Laboratoire de Chimie et Physique Quantiques, Université de Toulouse 3, 118 route de Narbonne, 31062 Toulouse (France)
E-mail: malrieu@irsamc.ups-tlse.fr

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/anie.201902527.

A15. From Theoretical to Computational Chemistry. *How chemistry fell into simulation, with theory suffering along the way.*
A16. Calculemus.
A17. The New Wave. *Machine Learning and Artificial Neural Networks. An imperfect introduction to the current implementations of artificial intelligence in theoretical chemistry.*
A18. Explainable AI. *The field is moving on; perhaps there will be a meeting ground.*

“*Most scientists believe that their inquiry is aimed at more than predictive power. While scientists want to know how a system will behave in the future, they also want to an explanation of why it behaves the way it does.*”[1] – Michael Weisberg

“*Understanding science without explanation? Impossible, or so I will argue—in the case of science, at least.*”[2] – Michael Strevens

“*Give me insight, not numbers.*”[3] – Charles A. Coulson


Photo: Santiago Alvarez

Roald Hoffmann was born in a part of Poland that is now Ukraine in 1937. The US was good to him, as to many immigrants, and he became in time a theoretical chemist. He has taught several generations of chemists how one could productively use molecular orbitals in thinking about organic, inorganic and solid state chemistry. With time, he also built his own land between chemistry, poetry and philosophy. Relevant to this paper, one way to see Roald's involvement with computers is that the science he did was entirely dependent on those marvelous tools. And yet he spent all his efforts, over fifty years, in a way fighting computers, transforming the multitude of numbers they produced into chemical explanations.

Jean-Paul Malrieu was born in 1939, son of a couple of philosophers. He went through the Ecole Normale Supérieure in Paris and started his research in the Pullmans' laboratory. He moved to Toulouse in 1974, where he gathered an important Quantum Chemistry group. His targets are both methodological, developing original techniques to treat the electron correlation problem (with a particular focus on magnetism), and interpretative, since he considers that the production of rationalizations, models and even metaphors is as important as reaching accurate numbers. Jean-Paul values deduction and loves translations from one language of Quantum Chemistry to another, for instance between Molecular Orbitals and Valence Bond Theory. He draws, and his social concerns have led him to write several non-scientific essays.

“*…the more accurate the calculations became, the more the concepts tended to vanish into thin air.*”[4] – R. S. Mulliken

“*It is only because we accept the risk of error that we can reap new discoveries.*”[5] – René Thom

At times it feels like we are under a huge wave, falling on us with terrific speed and strength. All around us, in chemistry and physics, our fields, but also in every aspect of our lives, simulation is growing. Better and more extensive calculations in theoretical chemistry are the least of it—computer programs are trying to speak to us, computer programs are trading in the stocks in our retirement accounts, without knowing much about the stocks they are selling and buying. Is danger to us lurking in this? Or is this not only unavoidable, but good, in some absolute sense?


Figure 1. Hokusai, “Oshiokuri hato tsūsen no zu” We have not chosen Hokusai's canonical woodblock print “The Great Wave off Kanagawa,” made by the artist some 25 years later. https://en.wikipedia.org/wiki/The_Great_Wave_off_Kanagawa.

Let us be more explicit about the questions that preoccupy us in this tripartite Essay: To what extent will the new technologies of computation and of simulation change our scientific practices? And more provocatively: Will their predictive efficiency, the better computation of observables the new technologies allow, push to obsolescence the traditional deductive paradigms, the crafting of theories, the historical pride of scientific knowledge? Will accurate numbers and constant referral to computers, replace derivations and story-telling?

We shall focus on our discipline, quantum chemistry, in describing the quandaries we face, at a level that we think many chemists will be able to share. From informal debates and echoes it is clear that similar concerns have also surfaced in the Physics near to us—in solid state physics, physics of fluids, statistical physics. And in other domains of science.

More widely, the addiction our society has fallen into (by design and chance) for social media, and the incredibly attractive realizations of artificial intelligence that wash in around us—these have certainly opened a societal discussion, across the world, on artificial intelligence (AI). So the wave breaks over us, and not just in theoretical chemistry. We will certainly broach the social concerns that surface as consequences, how could we not do that? But to be illustrative, and to speak from what we know, of both the risks of misuses and the potential of fruitful applications, our examples will be taken from the land between Physics and Chemistry to which we have dedicated our efforts for over half a century.

A1. We All Come from Somewhere

The authors' prejudices

As we approach our subject, we have to recognize our intellectual predispositions. We think the main one, revealed in the first paragraph of our Essay, and the quotations that preface it, is our skepticism about this brave new world.[6] That skepticism derives from several sources—the world we grew up of civilization crumbling in World War II, the failure of most *isms* that moved us. We are not crying; the smile of our grandchildren remains. Then, simply, there's our age. We come from a period when society attributed a higher value to knowledge than to efficiency, at least in the formation of its intellectual elites. We do recognize that experience tends by and large to breed conservatism (there are exceptions, for sure, take Noam Chomsky). Both of us calculated away merrily when we were young—a friend gave Roald the nickname "Numbers". But with time the weight of perceived, if not hyped, computational optimism began to weigh on us. Both of us underwent conversion experiences giving value and precedence to understanding. Of which more anon.

Readers who know our respective scientific work may wonder what gathers the two of us (coauthors for the first time!) in writing this Essay. We are in the same profession, for sure—a discipline of the last 100 years that stands between Physics and Chemistry,[7] and may be called theoretical physical Chemistry, or chemical Physics, (sometimes molecular Physics). Yet we have different "profiles," recognizable to fellow practitioners. RH comes more from Chemistry, always listening to the questions posed by molecules and their makers. Sometimes he says "I'm a chemist hiding as a theoretical chemist." He loves models and translations (geopolitics and history put him through a few languages). JPM is closer to Physics; he publishes most of his articles in journals pertaining to Chemical Physics. Many of his efforts are methodological, in search of more refined approximations. But he systematically considers complexity as something he has to go through and fight, to produce schemes of interpretations.

Both authors are reflective. This is not meant as a conceit, nor pretension. It simply states that we are willing to take the time to think about why we do what we do. And… we are not afraid of making fools of ourselves by writing of our untutored ideas. It is on this ground the authors meet.

We can't be where we don't want to be—the two of us value, deeply and emotionally, understanding and explanation. To put it simply, we value theory. And simulation, at least the caricature of simulation we describe, gives us problems. So we put our prejudices up front. Yet we believe we have something to say, something to be concerned about. We have come back from the *isms*, but remain socially concerned. Our passion for understanding is untouched. But, as you will see, throughout this discussion we seek what we see as a necessary consilience between simulation and understanding.

In beginning, a word of apology. In this long Essay, we want to discuss nontrivial matters of philosophy and risks to society. The words that enter such discussions are perforce more complicated than those of scientific English. We also want to enter the dialogue on understanding and artificial intelligence that is sweeping the world. We believe, strongly so, that chemistry and quantum chemistry have something of substance to contribute to that conversation. We want to bring chemists into that dialogue.

We've got a problem. What language, what mode of expression shall we use if we want chemists to enter the cultural discussion? And if we want the intellectual world outside to understand and value our experience with understanding and simulation? Shall it be the "high" English (or French) of philosophical monographs? Or the equations of coupled-cluster theory? Do we need to explain "memes" or a reference to Ingmar Bergman? You will see us struggle with this middle ground of appropriate language. Far from cultural aristocrats, or just writing to please ourselves, we want you to enter with us in that cultural/philosophical conversation. The journey is worth it.

A2. The First Triangle: Theory, Understanding, Simulation

Excursus on words across languages. The central role of "light"

We will structure our argument around a triangle, that of understanding, theory, and simulation (Figure 2). All around that triangle is the chemical (and physical, and biological) universe. We do not belittle the macro- and microcosm of the experiment around; Chemistry for sure, and even Physics,

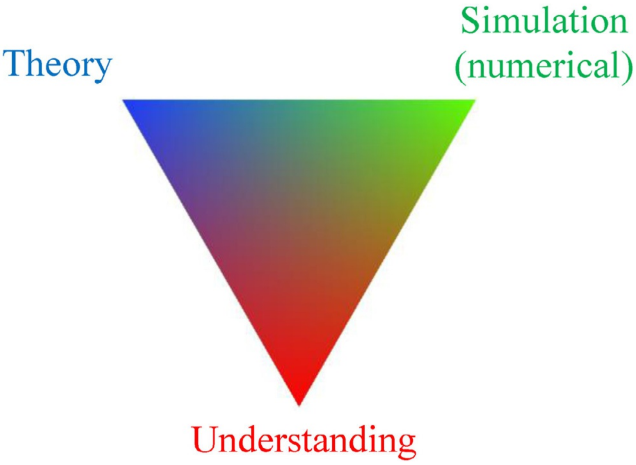


Figure 2. The first triangle: Theory, Simulation (numerical), Understanding.

remain predominantly experimental sciences. The points of this triangle and experiment, of course, are in constant dialogue. The data, even if reliable, are mute—one needs ideas and words to make sense of observed properties, to interpret them, to distinguish enigmas from the routine, to design the next experiment. We will return, rest assured, to what experiment does to the theory–simulation interface. But in our initial approach we remain in the conceptual realm of language and philosophy.

The concepts we care about, the ones that will trouble us, are complex in meaning, hardly as simple as "a 1957 250 GT Ferrari Berlinetta". Table 1 shows them in a selection of European languages. From one language to another, the words are supposed to denote the same concept. But never strictly coincide, because in each language the shades of meaning associated with the words add to our understanding. As systematically illustrated in the *Dictionnaire Européen des Philosophies*,[8] it is useful to identify the roots of the terms.

Table 1: The basic concepts in a variety of European languages.

understanding	explanation explication elucidation	Knowledge Scots: wit and ken
Verständnis	Erklärung	Wissen Kenntnis
compréhension	explication éclaircissement	savoir
(as a verb) entender comprender	explicación elucidación	Conocimiento

Compare the words for "understanding" or "to understand" in different European languages. In all the Romance languages, there are typically two words relating to understanding, for instance in Spanish "entender" refers essentially to an immediate perception, while "comprender" carries the logical connotation. There are also interesting nuances in surrounding words. So the Italian "capire" refers to taking, which is also an immediate action (and connects to the French word "saisir"). The French "comprendre" may be understood as "take together", which suggests a construction, but it may mean as well "take with oneself", or, more likely, integrate, assimilate a new element in the representation of the world by the subject. Does the English verb (to understand) suggest the identification of what is fundamental, that we stand at, and contemplate, the root of the fact or phenomenon?

The perceived ambiguity of these words, actually of any words in any language, is not a signal to retreat into the seemingly safe world of science. Rather it testifies to the complexity of life and of human beings as they try to communicate and create. Note the Latinate root of "light" in many of the European words around explanation. We desire enlightenment.

Let us elaborate the import of these words for the soul—understanding, theory, simulation.

A3. Understanding, Explanation, Knowledge

Our definition of these words, and the role of teaching, storytelling, rhetoric and performance

Understanding is often tacit, a state of the mind. It is usually qualitative, though it may have quantitative aspects, and these (the quantifiers) may be strongly reinforcing or denying. An example is Descartes' and Newton's explanation[9,10] of the rainbow as arising from internal reflection of light in water droplets, as startling and clear as it was over three hundred years ago. That the droplets must be opposite to the sun, that reflection, refraction, dispersion matter—these constitute the qualitative aspect of the explanation. The 42° maximum angle of the rainbow above the horizon is the quantitative one (see Figure 3 for an indication of how the explanation goes.) Quantitative reinforces qualitative. And simulation is around the corner.

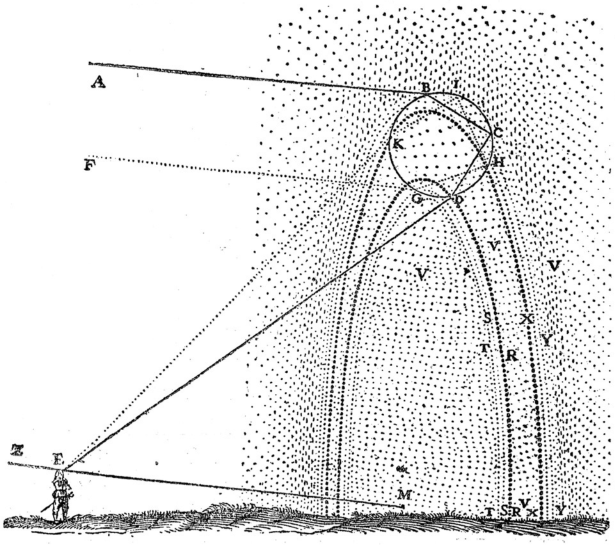


Figure 3. Schematic of the explanation for the rainbow, from René Descartes, Discours de la méthode (1637). (Figure courtesy of https://en.wikipedia.org/wiki/Rainbow).

Understanding is universally satisfying. Therein lurks danger, for incorrect, seeming understanding may also be satisfying. Witness disbelief in evolution or climate change for too many Americans.[11]

Explanation is inherently more pedagogic, rhetorical, storytelling, and performative than understanding.

1) The pedagogic aspect refers to the fact that it is well nigh impossible to separate explaining to oneself from explaining to others. The impetus to do the latter—to teach—"activates" understanding in us. We have more to say about this in a section that follows.
2) The sense of rhetorical here is the positive one, of using spoken or written language to persuade, inform, and effect change in human beings. From the time of Plato and Aristotle a tension lies under rhetoric, in that the ethical value of what sways a human being can be in some

absolute way just. Or people can be convinced to do wrong, to hurt others. The sense in which we think of explanation being rhetorical is that it is spoken or written, by another human being. So one is led to agree or disagree, and out of that dialogue understanding forms.

3. Human beings tell stories, naturally and easily. Yes, some are "just so" stories, teleological pseudoexplanations. More useful to science are hypotheses, basically alternative narratives. We can come up with a few ourselves, though given human nature it is likely that we delimit the hypotheses we pose as we pose them. As convinced as we are likely to be that each of us has the true explanation, it is good for understanding that we listen to other people telling stories.
4) We have crossed to the performative aspect of understanding in the process of formation. To paraphrase Jacques Derrida, writing is the message that abandons.[12] Inside our minds is an inchoate potpourri of pieces of understanding, shaped by our teachers, parents, the books we have read. When you physically voice (or write) an argument, the context of justification emerges—all the ways to support your explanation flood into your mind. You voice them, write them, perform them. Be they reasonable or not.

As scientists, we perhaps believe that we have a clear definition and experience of what understanding means. But we all know the ambiguity of this word: it designates both a feeling and a process. The feeling is this sudden impression that something which had been obscure two minutes before is now clear. The process of understanding lies in our ability to follow a chain of implications, a causal trail. We may follow this trail in two opposite directions, either upward, starting from hypotheses and deriving consequences, or downward, starting from a fact we have to explain and finding its causes.

Notice that we sometimes have the impression of having understood, while in fact our deduction is either logically false or introduces erroneous facts or assumptions in its construction. In the scientist's practice, understanding is both a subjective motivation and a pleasure. Yet it is—in principle—submitted to the verification, of its reliability if an observable is involved, and of the rigor of deductive explanation, if the understanding is theoretical. The ultimate test, making sense to others, is, given human nature, almost always preferred. Again, here is the importance of teaching. And tentative understanding is paired in the real world with experimental reality. We will have much more to say about this pairing, and prediction as well.

Understanding/explanation both have strong elements of causality $A \rightarrow B$, $B \rightarrow C$ built in. However, especially in modern times, in both the humanities/arts and in science, chance (the unpredictable, the aleatory) plays an essential role, and is viewed positively. It is intriguing to see how chance enters into understanding in the sciences, be it through entropy or chaos theory.

Knowledge encompasses both understanding and explanation. Yet it also includes awareness of many facts that have not been assimilated into a framework of being understood or explained, many sensory inputs. The German words *Wissen* and *Kenntnis*, the Spanish *entender* and *comprender* also probe the difference between the more logical and intuitive ways of processing the world around us.

We have not yet broached Wisdom. And let's not even start on Truth. The point is that we are not qualified to teach or even discuss knowledgably what has been achieved in the millennia of philosophy. Nevertheless, we sail ahead—perhaps something will be gained if we do not wander far from common sense, yet inject into the discussion the special, hard-won knowledge of chemistry that is ours.[13]

To return to our main topic: We see understanding as more toward the passive end of the spectrum, explanation of the active one. Interesting—if understanding tends to be passive, it is best formed in an active, very social mode, by teaching, listening, talking. And both—understanding, explaining—are certainly welcomed by the mind. More than that, as we will argue, they are a source of joy.

Figure 4. The imperfection in the frame is important. The image, from Andy Dean Photography, is reproduced by permission.

A4. Teaching

Why understanding and teaching are so close

Teaching plays a critical role in the formation of understanding. And learning, at least that aspect of it known as machine learning, will figure prominently below. So it makes sense to explore at least some of the richness of these activities.[14]

People have tried to move beyond the teacher–student or master–apprentice paradigm for transmitting knowledge and mastery. Yet non-biological transmission of our culture defines the human condition, and it seems to us that some variant of this mode of knowledgeable—uneducated interaction will persist. Think only of the instructional context, the clarification (there is light again!) of a seeming complexity, making things plain.

Usually we think that involvement in research makes for a better teacher. Here is an argument for the reverse, focusing on the special role of teaching in enhancing understanding/explanation, for the teacher. While understanding is usually a tacit, contemplative quality, to attain it, it clearly helps to exercise the facility. Socratic dialogue serves, but closer to home is what many of us are paid for, and this is to teach. Teaching, if it goes beyond training in the use of recipes, forces you to explain, thereby voicing what may be tentative and muddled in your brain. In the process, you get nearly instant feedback. We learn to interpret the nonverbal signals—which may be overt, like falling asleep—by which young human beings inform us of whether we have awakened the mental facilities in them. Or not.

Teaching leads you to construct simple narratives. Another way to say this less positively is to say that both teacher and students have entered into a nonverbal contract to make the world simple. Too simple. But there is a very positive side to this process. A story is woven. The teaching process tests your narratives of explanation quickly. Not only are they understood or not, but students are also likely to apply your models/explanations in ways you would not have thought of. So the power and generality of understanding are immediately tested.

What we write is not new; from Seneca the Younger's "*Homines dum docent discunt*" to modern times, this idea – that teaching enhances learning – has been rediscovered, applied, and tested.[15] An interesting point has been made to us by Alexander Frank (private communication), that teaching serves in another distinctive way; Frank posits "that being able to explain a theory, model, or system is a strong differentiating agent between "plain" knowledge, and "deeper" understanding." This makes sense.

Let us finally quote a Feynman story, as told by David and Judith Goodstein:

"Feynman was a truly great teacher. He prided himself on being able to devise ways to explain even the most profound ideas to beginning students. Once, I said to him, 'Dick, explain to me, so that I can understand it, why spin one-half particles obey Fermi-Dirac statistics.' Sizing up his audience perfectly, Feynman said, 'I'll prepare a freshman lecture on it.' But he came back a few days later to say, 'I couldn't do it. I couldn't reduce it to the freshman level. That means we don't really understand it.'"[16]

Feynman was optimistic, but the philosophy behind his conclusion makes sense.

A5. Learning

The essential feature of cultural evolution

Teaching and learning are near mirrors, even as they are distinct. There is much more that one can say about learning, hardly a passive action in which the student's head is filled with "facts." Thomas Aquinas was closer to what happens in learning—the latent powers of reasoning, always there, are awakened in the student.[17] Lucky is the awakener, as well.

In time we will explore "machine learning," so it is useful to think about the relations between understanding and learning. Understanding a phenomenon or a theory enables one to verbally formulate an explanation, to retrace the deductive trail followed in the demonstration. In other words, to teach what one has understood. And to learn. One has understood "why" things are as they are. The caveat is that such understanding is certainly provisional, and just might be wrong.

But one may as well teach (and learn) something different, the "how to," a procedure to accomplish something real, without knowing why the procedure works. In the process, there is verbal transmission of knowledge, an empirical recipe. Our practical life makes broad use of such mastery, let us call it training. Even further from understanding are non-verbalizable learnings—the way we learnt to walk, learnt the language we speak, all the basic recognition processes which allow us to move autonomously through a day. There was a time we did not know these, and we learnt them—through now obscure connections of our neuronal system. Much effort is expanded by psychologists to reconstruct these childhood learning processes. Machine learning and neural networks, as modern tools of simulation, stand at various places in this broad spectrum between understanding and empirical mastery. The expression "training" will recur in our description of these.

When we get to machine learning, it will also be interesting to reflect on how much less common the expression "machine teaching" is. Indeed there is a use for computers in patient, student-calibrated instruction. But learning is so much more—it is the basis of cultural evolution (in distinction to biological evolution); it's what enables ways of knowing and doing to be passed on, enabling the rapid development of human beings. For better or worse.

A6. Revelation

The emotional quality of understanding

There is a special quality to understanding, an emotional surge when it is attained. By bridging cognition and emotion, that impact forms an important spiritual link between science and art. That quality is sufficiently important to us so that we will return to exemplifying it below, in Part C of our Essay.

In understanding there is logic. And there is mystic truth, the revelation of some deep enlightenment, the "fire" of Blaise Pascal's night, sewn by him into the lining of the coat he wore to the end of life:

FIRE
GOD of Abraham, GOD of Isaac, GOD of Jacob -
not of the philosophers and of the learned.
Certitude. Certitude. Feeling. Joy. Peace.[18]

This is pure feeling, the words for a verbal transcription of which are not to be found.

Yet a by far more rational and logically confirmed intuition may take the same form. Let us quote Andrew Wiles speaking of the resolution of the last difficulty he faced on the route to the solution of Fermat's theorem:

"I was sitting at my desk one Monday morning, September 19, examining the Kolyvagin–Flach method. It wasn't that I believed I could make it work, but I thought that at least I could explain why it didn't work. I thought I was clutching at straws,

but I wanted to reassure myself. Suddenly, totally unexpectedly, I had this incredible revelation. I realized that although the Kolyvagin–Flach method wasn't working completely, but it was all I needed to make my original Iwasawa theory work. I realized that I had enough from the Kolygin–Flach method to make my original approach to the problem from three years' earlier work. So out of the ashes of Kolyvagin–Flach seemed to rise the true answer to the problem… It was so indescribably beautiful; it was so simple and so elegant."[19]

The aesthetic component of the pleasure of understanding seems different from the pleasure of mastery of a problem. As a reviewer has commented, there is an opening here for an fMRI (functional magnetic resonance imaging) study, to see if similar areas of the brain "light up." The specific aesthetics of logical deduction deserve analysis, to which we will return one day. So do the intuitive jumps in that logic. On the way to understanding there may lie a pleasure of simplification, of return to the most direct path, to strict necessity. *En route* one may encounter the emergence of similarities, isomorphisms, sometimes (Maxwell's revolution) the fusion of different constructions, born from seemingly remote questions, into a unique paradigm. The pleasure may be combinatorial, imagining cross effects at the intersection of domains perceived. It is not an accident that the joys of understanding in science are akin to those in artistic production.

Understanding pleasures the mind.

A7. Two Operational Definitions of Understanding

Roald's pragmatic definitions, in the context of chemical theory

Shall we come back to earth? One of us (RH) has found utility in two ways of describing or reaching scientific understanding.

In the first approach, understanding means knowing the mix of physical mechanisms (there may be more than one) behind an observable, and making an order of magnitude estimate of contribution of each. That sounds fancier than it is. Imagine for instance that you want to have a dry towel for tomorrow. Which of the following will expedite the process: wringing out the towel after washing, hanging it on a shady or sunny spot of your yard, where the wind blows, or in a sheltered space? Will a flashlight shined on the towel, or your breath on it speed it up? Sounds silly, right?

What goes into describing the energy of interaction and orientation of two molecules a certain distance apart is not that different. We might talk of charges attracting or repelling each other each other (electrostatics), or a purely quantum-mechanical requirement, so-called "antisymmetrization" of the electrons tending to share some spatial regions, or orbital interactions, or of the coordinated yet random motions of the electrons (technically called dispersion forces[20]). More technical, each with meanings that could be made more precise (but often aren't), those are just descriptors of contributing factors that a theoretical chemist might invoke.

The implication of this "method", if one accepts it, is that understanding is inherently qualitative. That you must have prior understanding, perturb the system, and reason out what happens, if you really understand it. What changes when you hang out the laundry when the temperature is −10 °C (Ithaca, not Toulouse). What changes if you bring the two molecules from 6 Å to 2 Å apart? Do you go back to the computer? If you need to do that, who understands? We are ahead of ourselves in these questions.

The second definition Roald proposes is that understanding means being able to predict the result of reliable computation qualitatively prior to making the calculation. Yes, this is not entirely consistent, as it assumes a perfectly reliable calculation. Actually, what he advocates is an iterative process:

1) Predict the result of the calculation qualitatively, before the computation is carried out. Or, as John Wheeler recommended in a provocative manner "Never make a calculation until know the answer."[21]
2) If the result of the computation is qualitatively right, never stop, try a variation. By way of example: what happens to the ionization potential of an ethylene if you replace an H by an NH_2 (a typical organic chemistry substitution)?
3) If the result given by the computer disagrees with your qualitative prediction, after you eliminate your mistakes (á la "garbage in, garbage out"), then cogitate, the way Rex Stout had Nero Wolfe do.[22] You need to find the physical effect which you have omitted or supposed to be of lesser importance. There is a necessary stop, while you perform an analysis which correctly handles both effects, the one you supposed to be decisive, and other ones. In time you will resolve the contradiction, gaining more understanding in the process. The explanation, when it comes, will seem so obvious that you could kick yourself in the behind for not having seen it.

Friends are there to get you out of a rut in thinking. But never stop, let the chain of human-machine interactions remain unbroken.

Here's Charles Coulson:

"So much for the business of trying to rival Nature and get the numerical answer absolutely right. As I have said, this is only one part of the role of theoretical chemistry. Often, without such accuracy and with only just sufficient numerical agreement to satisfy ourselves that we have really got the right model, we can acquire a profound and deeply satisfying insight into some particular phenomenon."[23]

A8. Theory

Dictionary definitions. And why people believe theories

An American dictionary definition has theory as "a coherent group of propositions used as principles of explanation for a class of phenomena."[24] Like all definitions, this one is deeply circular—it usually takes only 4 links in a chain of definitions to come to a word being defined in terms of the word that is being defined. No matter, we live perfectly well in a world of quasicyclical definitions.[25] And the operative words in this definition are "phenomena" and "explanation."

There are several ways to think about theories. First, theories are the way human beings "organize" observables. We pass by problems with what constitutes reality (elsewhere one of us has thought about why chemists are such convinced realists[26]), and the insistently fallible nature of our senses, and their extension, our scientific instruments. Theories give answers to the "Why?" question. That the answer is likely ephemeral, bothers no one.

A second definition of theory is that it is a logical, internally consistent construction of statements. Or, as James Bogen puts it, theories are "collections of sentences, propositions, statements or beliefs, etc., and their logical consequences.[27]" Note no reference to experiment, to confrontation with some reality. Theories of this kind can be quite speculative, without any possibility of checking their adequacy to the real world we live in. At least not yet—science has given us repeated examples of seemingly esoteric mathematical structures that end up being useful, witness non-Euclidian geometry and Penrose tilings.

Note that up front in any definition of theory is that it is an explanation. Some would say that the words are synonymous; we don't think so. For instance, a certain level of explanation is explicit in History, the great one or your personal history. There (if accurate, or honest) it provides a chain of causes, usually without invoking a theory. But theory and explanation are awfully close in the way they function in science. And, as we will see, it is through theory that simulation is linked to understanding.

As a reviewer reminded us, a fact of life with theories is that many are incomplete (can a theory ever be complete?), and many are just plain wrong. Even though these theories might have served smart, good scientists for the longest time. We have in chemistry our striking story of phlogiston. Evelyn Fox Keller gives some good examples for biology in her excellent book, "Making Sense of Life."[28] That many theories have been historically wrong has led to the "pessimistic induction" hypothesis, that current theories in science are likely to be wrong. Not without dispute.[29,30]

The theory of theories simplistically begins by positing that people accept theories because you understand more with them (rather than with alternatives). So theories are believed because they explain more facts more economically, and at their edges blend in understandable ways with existing theories that explain remainders of the universe (there is the "coherent" in the dictionary definition.) *Pace* grand unification, theories are piecewise signposts in a complex universe.

There are other reasons people believe theories,[31] but surely the most convincing belief-inducing stratagem of the theorist is to make a risky prediction. Here "risky" signifies a prediction such that if you polled the experts in the field, 90% offhand would doubt the veracity of the prediction. The general theory of relativity's explanation of the anomalous precession of the perihelion of Mercury just about one hundred years ago, is a good example.[32]

A9. Mathematics and Computing, Physics and Quantum Chemistry

A little on mathematics in modern times, more on physics, especially when it cannot be done deductively

Mathematics, which may be considered as the very heart of theoretical practices, proceeds primarily through deduction, at least in the final presentation of a work. It does so even if the theorems proven, or the categories established, have been guessed or suggested from analogy and intuition, according to a non-deductive process. Number theory is a hard core subject of mathematics, but to produce general statements, Alain Connes recalled recently this expression of Évariste Galois (who died in a duel at 20), one must *"jump in with both feet on calculations."*[33] This sentence illustrates, among other things, the status of computations, at most a springboard.

Remembering that numerical computation appears in our triangle, it is worth mentioning that informatics has come to play some new roles in modern mathematics. Some proofs rely on the systematic examination of a finite but exceedingly large number of cases, too numerous to be exhaustively treated by "manual" approaches (an example is the 4-color theorem). Some purists judge that such computer-assisted proofs, even if they may be called deductive, do not satisfy the full requirement of a mathematical demonstration. Another inroad of computing in mathematics came from the exploration of some iterative systems. So "strange attractors" appeared as mathematical subjects in what could be called numerical games on computers.[34] In this case the computer is a source of new formalizable mathematical problems. Even the most abstract scientific practice now is tempted (or rescued) by computation…

In recent times computers and programs have advanced to the stage where (in limited areas) they are able to come up with problems (theorems, conjectures, not numbers) that human mathematicians deem interesting. Computers have apparently learned to derive demonstrations. One must, however, notice that these impressive achievements involve well-defined sets (numbers and graphs); the production of new concepts seems to remain the privilege of human brains.[35]

Where shall we place quantum chemistry in the description of theoretical practice? Let's begin with Physics, certainly the most deductive Science of Nature, since it is the terrace on which theoretical Chemistry builds its specific constructions. In its most formal version, Physics loves to be formulated from a finite set of concepts and hypotheses, entering into formal constructions. These constructions must be as consistent as possible, and the body of concepts and fundamental laws as compact as possible. Physics tries to obey an esthetic of economy, going to the conclusion by the most direct way, with the minimum of hypotheses.

But Physics can never (or rarely) be a pure theoretical construction, for the deduction may introduce well-established phenomena, and fundamental constants, which have to be accepted as basic building stones. The number of *ad hoc* assumptions in a physical theory, for instance the Bardeen–

Cooper–Schrieffer (BCS) conceptualization of superconductivity, accepted by the community, is substantial. Well-established results which we invoke in an explanation may or may not have received a fundamental justification. Yet they will enter into the deductive process as do the fundamental laws, as pieces of a logical explanation. When quantum chemists invoke the aromaticity of 6-membered carbon rings, for instance, they make use of such intermediate building blocks (rings, resonance stabilization, delocalization) in their rationalization.

Theoretical Physics tries to make use of the fundamental laws only. But for most of the systems to which a consistent theory (say for instance Quantum Mechanics) may in principle be applied, one is compelled to leave the deductive track. For there is no exact, closed form solution of the fundamental equations for complex systems. Then one adopts one of two procedures (or combines them):

- the first one consists in simplifying the problem, making crude approximations which make possible an analytic derivation, offering the possibility to formulate theorems and to identify qualitative phenomena, the validity of which is restricted by the simplifications one has accepted.
- the second approach consists of accepting the smallest number of simplifications and treating numerically, on computers, the general equations applied to specific problems, for instance to atoms, molecules or solids. General statements are not expected from these numerical explorations, as accurate as they may be.

Theory in chemistry is hardly only quantum chemistry, to be defined roughly as seeking the solutions of Schrödinger's equation or its relativistic variants. It also includes the theories behind the remarkable spectroscopies available to us, the beautiful formalisms of statistical mechanics, all the interactions of radiation and matter that have served us so well to probe structure and dynamics, and…. much more. In progressing, Quantum Chemistry makes use of both recipes just mentioned. A devil's bargain here. And can one get out of that?

A10. A Quintessential Anxiety in Doing Theory

As long as we define ourselves as theoreticians (not as efficient predictors), Theory stands above us. It looms, a dominating Figure which we can never put out of mind. The question of the deductive rigor of the way we have established our statements is always present in the mind of some of us; JPM belongs to this family of anxious scientists, preoccupied by the distance between what they are forced to do by their simplifications and what the exact solution would be. As a methodologist among many, he devotes much systematic effort to the production of new tools. Which is not Theory, since it does not touch the fundamental principles, but which may be called theoretical (or formal) development.

As explained later, the satisfying new tools of artificial intelligence (if not simulation) should provide more accurate prediction. But interpretation? The pleasure that the methodologist may find in the conception of such tools somehow equilibrates the obsessive anxiety JPM feels facing the persisting gap between what he says and what would be an exact solution of the fundamental equations. The libido of the methodologist is of course a thermodynamic machine with cold and hot pieces.

RH is less anguished.

It is worth mentioning that Quantum Chemists have contributed significantly to the understanding of the logical structure of the quantum many-body problem, and that their main partners in these developments were, surprisingly, Nuclear Physicists. The two groups work in completely distinct ranges of spatial scales and energies, indeed they may appear very far from each other. And yet, the fields have established an unexpected but fruitful dialogue.

A11. Dirac's Dictum

Two ways to interpret a 90-year-old insight. An introduction to the jargon of computational and quantum chemistry

Quantum Chemistry exemplifies clearly what chemists are forced to do. Here Dirac (Paul Adrien Maurice Dirac) has been often quoted:

"The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble. It therefore becomes desirable that approximate practical methods of applying quantum mechanics should be developed, which can lead to an explanation of the main features of complex atomic systems without too much computation."[36]

This quotation is just as often abbreviated to the first phrase, and as such unfairly seen as an example of the arrogance of physicists. The full statement is much more astute, telling us that Dirac foresaw the difficulties. With which we live.

Much has happened in our lifetimes; we, all of us, have witnessed the incredible growth of information technologies. A revolution of the magnitude of the European reinvention of printed books has taken place, is taking place, all around us. The change is truly transformative, affecting every aspect of our lives, far away from quantum chemistry.[*]

In our profession, theoretical chemistry, there has been a two-fold development—first the remarkable introduction of density functional theory and calculations (DFT) by Kohn and co-workers, and Parr and Pople. There are two aspects of DFT that are interesting from a general point of view: 1. It was not expected that the method would, could work. That the density, if we knew it exactly, had all of chemistry in it. 2. Nor was it expected that the auxiliary functions used to express the density, which initially were denied a reality akin

---

[*] We highlight in grey those sections throughout this paper which use more than the normal dose of quantum chemistry jargon. We need the technical language, we feel, but we are painfully aware of the barrier to understanding that technical jargon may create.

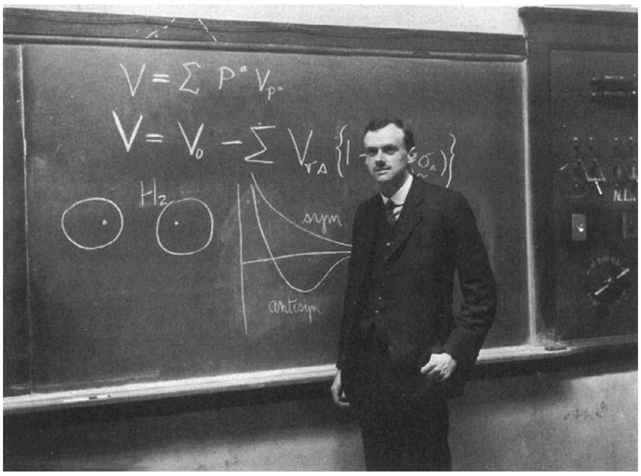


Figure 5. P. A. M. Dirac (1902–1984). Note the presence of the hydrogen molecule and its potential energy curves on the blackboard. Also of interest, if not just a chanced drawing, is (as a reviewer noted) the positions of the nuclei in the orbital or electron density. AIP Emilio Segrè Visual Archives. Gift of Mrs. Mark Zemansky.

to that of orbitals (by then familiar, almost tangible), in fact would prove to have a like approximation to reality.

The second line of research, which has been evolving steadily with growth of computer power, from Dirac's time, through James and Coolidge, to Boys and Pople, is the growth—technical term coming—in the sheer number of configurations that could be examined in what we will call "wave-function" methods.

A *configuration* is just a way to distribute the electrons among the molecular orbitals of a molecule. In many instances the electronic structure of a molecule (and therefore its properties) can be reasonably described with just one configuration and we then talk of *monoconfigurational* problems. But often we need to employ more than one configuration in what are then called *multiconfigurational* problems.[37] In the latter case, *configuration interaction* is one time-honored way of improving the wave function (Figure 6). It has been augmented in modern times by ingenious so-called "coupled cluster" methodologies, and clever perturbation theoretic approaches.[38] Later on, progress came from explicit consideration of the so-called "Coulomb cusp", the discontinuity of the derivative of the wave function when two electrons of different spins approach each other.[39]

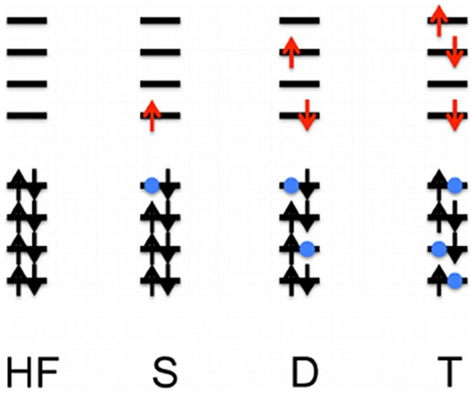


Figure 6. A schematic of configuration interaction at work. HF stands for the ground state "Hartree–Fock" configuration in the description of the electronic structure of a molecule. S, D, and T stand for configurations that might mix into the ground state, related to the reference configuration by S = Single excitation, D = Double, T = Triple. Figure reproduced with permission from "Truncated configuration interaction expansions as solvers for correlated quantum impurity models and dynamical mean-field theory": *Phys. Rev. B* **2012**, *86*, 165128.

You will notice that in describing the various approaches we avoid the descriptors "ab initio" and "first principles" so beloved by practitioners in the field. Despite the legitimate attention one must give to the theoretical foundations of a problem, one may consider that these adjectival phrases have as their main function to confer imagined pseudo-sophistication on the people who use them, and to make themselves feel superior to others. One should not serve natural human weaknesses in this way (politicians do enough)—the various methodologies often have massive assumptions and parametrized functions internalized in the programs.

It should be noted that computational chemistry also consistently and justifiably makes assumptions even when it thinks it is not making any. The Born–Oppenheimer approximation for molecules is one such simplifying approximation; projecting the Schrödinger equation onto a finite basis set (a set of functions in which the solutions are approximately expanded) is another. The choice of the basis set (the Hilbert space) is based on a reasonable empiricism—the use of atom-centered orbitals is logically based on the strength of the local nuclear attraction, the precise basis sets are obtained from preliminary accurate calculations on the atoms.

Much more can be said about the practice of our trade, and we will return to it through some case studies which show just how good its results often are. And how circumscribed the understanding reached from good quantum chemical calculations can be. We must do better.

And, to approach the battleground, shall we call that practice, that of computational quantum chemistry done as well as humanly possible, shall we call it simulation? In a way it is, a quantum mechanical simulation of reality. For the moment we prefer to label as quantum chemical "numerism," an abiding interest in numbers.

A12. Prediction, Linking Understanding to Simulation

The special role of prediction, and the bridge it forms to simulation. "To predict is not to explain."[40]

Prediction is the way the conceptual passage between understanding and simulation is shaped. So let us discuss prediction.

There are certainly different types of predictions. One might be magical, say, based on the feeding patterns of chickens.[41] Another may be based on experience, the observed recurrence of signs and events, their correlation, legitimately even a link which one cannot rationalize. Or a prediction may be grounded on a logical (or at least discursive) representation of causality, offering a possibility of foretelling what might occur in a well-defined new situation or process.

The logical prediction is, of course, closely related to understanding, in that it consists of applying a logical machinery to a specific case or a new problem. It will in general confirm, sometimes (and this is even more exciting) invalidate, the strength of the accepted logical construction.

But danger lurks. The logical acquires a psychological tinge: there is certainty in the pleasure of efficiency, of possessing a tool which enables us to anticipate the result of a new experiment, a new design. The efficiency of wielding your intellectual tool leads you to a seductive feeling of mastery. Heady, and dangerous stuff, this. You have now left aesthetic pleasure for the pleasure of power (on things, and eventually on human beings, either rivals or subjects). Prediction offers the operational counterpart of contemplative understanding, the "I know how" instead of the "I know why". And of course we may face a problem if we do not know why the "know how" works.

This is the danger—that what started out from Understanding and Theory becomes—in stages or suddenly—focused on numerical agreement, on prediction.

But… "Prédire n'est pas expliquer/ To predict is not to explain." (Figure 7) One should pay attention to this strong and provocative statement of René Thom, author of catastrophe theory. The IT revolution has indeed changed the terms of the confrontation between prediction and understanding.

In the non-pejorative sense, simulation is a relatively modern notion, which is why the word is similar in modern languages. We take simulation (about which we will have much more to say) to mean the construction of a variety of models that reproduce reality as closely as possible. Without being that reality, though that division may be showing signs of fraying (see the motion pictures *The Matrix*,[42] *Ex Machina*[43]).

"Reproduce". To a scientist, the word automatically invokes criteria of similarity, and the notions of accuracy and precision which we try, with moderate success, to teach our students. So one thread of argument leads us to numerical simulation, one face of the practice, and the one theoretical chemists most often desire or offer. There are other kinds of simulation and perhaps the connection is best made through "modeling", a concept that has a longer history. In fact, it is possible to see all the thought activities we have discussed as the construction of mental models and their confrontation with each other and with human reason.

A13. Models

Michael Weisberg's typologies

> *"Few terms are used in popular and scientific discourse more promiscuously than "model". A model is something to be admired or emulated, a pattern, a case in point, a type, a prototype, a specimen, a mock-up, a mathematical description—almost anything from a naked blonde to a quadratic equation—and may bear to what it models almost any relation of symbolization."*[44] *– Nelson Goodman*

Michael Weisberg has written an excellent book on modeling.[1] (Figure 8) He begins by distinguishing three kinds of models: concrete (physical constructions to some scale), mathematical, and computational. All contain at their core "an interpreted structure that is used to represent a real or imagined phenomenon."

Idealization is a core activity of modelers. There are several reasons for proceeding in this way. Weisberg distin-

Figure 7. The cover of René Thom's book. Reproduced by permission from the publisher, Éditions Flammarion, Paris.

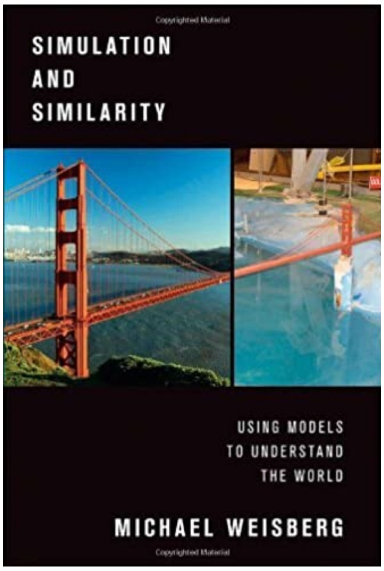


Figure 8. Michael Weisberg's Simulation and Similarity: Using Models to Understand the World. Reproduced by permission from the publishers, Oxford Univ. Press.

guishes three such: The first he terms Galilean—with the goal of simplifying, justified pragmatically. The simplest molecular orbital or valence bond approaches are of this type. The second goal is minimalist, so as to keep only the core factors which give rise to a phenomenon. Two of the great models of chemistry and physics, the Hückel and Ising models, are superb examples. The third category is that of "Multiple-Models", a practice of building related but incompatible models, each of which makes distinct claims about the nature and causal structures giving rise to a phenomenon." We think we don't have these in chemistry, but we sure do—hydrogen bonding is a fine example (is it due to electrostatics or the interaction of lone pairs with σ* orbitals?). So is "halogen bonding". So is our understanding of VSEPR (Valence Shell Electron Pair Repulsion model).[45]

It is also possible, Weisberg says, to have modeling without a specific target. He notes first generalized modeling; aimed at understanding classes of phenomena, not individual observations. He places "How possibly" models in here, but perhaps they are better in next category. A second category he calls hypothetical modeling: to construct models of nonexistent phenomena. Here chemical examples abound, for instance Albert Eschenmoser's incredible construction of an experimental nucleic acid universe based on six-membered rings (five carbons and one oxygen, hexoses) as an alternative to the ribose one (five-membered rings, four carbons and an oxygen) of RNA.[46] And RH's fun with thinking up molecules with carbon in square planar environment, instead of the typical tetrahedral one.[47] The third category is targetless modeling, just playing games.

Weisberg's book is entitled "Simulation and Similarity: Using Models to Understand the World." So we can see that similarity, the model-world relationship, is important to him. Of course, it is to us too—to modelers and their believers, or sceptics. The many ways of idealization lead to many models. Sometimes Roald tells his students that if there are three people in this world, they are guaranteed to find three ways to describe the same thing. Which does not mean that the various interpretation are equally grounded and robust, as was illustrated in a previous section on the example of molecular interactions.

Seriously, what criteria shall we use for judging similarity between models and their targets? In evaluating models Weisberg points to the following, *inter alia*:

1. Completeness—inclusion and fidelity.
2. Simplicity; good old Ockham's razor, whose shave Roald would avoid.[48]
3. "Maxout"—"the theorist should maximize the precision and accuracy of the model's output". This sanctions "the use of black-box models, the sort that have amazing predictive power, but for unknown reasons." i.e., a caricature of simulation.

And this is where one of the quotations that heads this Essay comes from, more fully:

"At first blush, it may seem unscientific to adopt an ideal that values predictive power over everything else. Most scientists believe that their inquiry is aimed at more than predictive power. While scientists want to know how a system will behave in the future, they also want to an explanation of why it behaves the way it does. MAXOUT ensures that we will generate models which are useful for predicting future states of the target system, but gives no guarantee that the models will be useful for explaining the behavior of the system."

This sounds like a lot of contemporary computational chemistry.

Weisberg has no inclination to prioritize Maxout, or "just" simulation; he is with us, as the quotation at the beginning of this paper makes clear. He describes what we would call a multivalent approach, including a place for qualitative comparison, a richness in the structures to be compared, and a place for disagreements.

In another place, one of us has, with less sophistication, mulled over the reasons why people buy a certain theory. Or a certain model. RH begins with the theory of theories account mentioned above, stressing the role of risky predictions.[31] He then goes on to point to other factors: 1. Aesthetic appeal (which to physical scientists sadly often means simplicity); 2. The quality of the story; 3. Portability—a theory that others can use is much, much more likely to be accepted than one for whose implementation one must return to the originators; 4. Productivity—that theories suggest experiments. In Roald's breezy account, one can find many analogies to Weisberg's exposition.

Let us return to René Thom for a parting word on models: *"The ultimate goal of science is not to accumulate empirical data indiscriminately, but to organize these data into more or less formalized structures that categorize and explain them. For this purpose, we must have ideas "a priori" on the way things occur, we must have models. Until now, the construction of models in Science has been above all a question of chance, of the "lucky guess". But the moment will come when the construction of models itself will become, if not a science, at least an art."*[49] – René Thom

A14. Simulation, of the Old Kind

First approach to defining different kinds of simulation—analogic, phenomenological, approximating. Numerism

One thing we have to get out of the way, for it colors to a certain degree our attitudes toward simulation. In every European language there is a pejorative sense of simulation, as a quality that hides a difference or a distance between something (behavior, belief, identity…) which is considered as socially desirable (or the truth) and some kind of imitation. It could be that we owe this negative feeling to Plato, with his low view of artists.[50] It is the hiding or obscuring of the distance that provokes moral condemnation. So, for instance, in the visual arts or theatre, the ability to explicitly simulate real objects (in painting) or characters (on scene), and to do so well, is certainly not considered negatively.

Though abstract artists do not agree. Some view their art as "pure" art, unchained from naturalistic simulation. Here is Kasimir Malevich, in his 1915 booklet "From Cubism to Futurism to Suprematism: The New Realism in Painting," a foundational tract of abstraction:

"And only a cowardly consciousness and meagre creative powers in an artist are deceived by this fraud and base their art on the forms of nature, afraid of losing the foundation on which the savage and the academy have based their art.

To reproduce beloved objects and little corners of nature is just like a thief being enraptured by his legs in irons…

The artist can be a creator only when the forms in his picture have nothing in common with nature."[51]

Put aside the negative tinge to simulating; we are interested, positively, in its manifest utility in science. Let us start with a short typology of simulations. Analogical simulation has not been extensively used in science, though Weisberg tells the striking story of one such model, and wind tunnels have proven utility.[1] If one knows that similar equations rule the behavior of two systems, one which is too large or too microscopic to be handled, the other one being accessible on the human scale, one may exploit this similarity to construct a device mimicking the first system. For instance, this is possible when one knows that the micro or macro system obeys hydrodynamic Navier-Stokes, or more generally, a set of coupled linear equations, which might be translated into the conductance across a network of resistances. Progress in electronics was so rapid in the middle of the last century that the digital computation approach, much more flexible, led electrical engineers and circuit designers to abandon the analog way. As will be briefly discussed later, one may wonder whether neural networks cannot be considered as analogical (the recognition processes working as our brains do).

A second class of simulations, common in our trade, are "phenomenological models." These are simple models, the theoretical foundations of which are not (yet) clear, but which correctly reproduce some properties of sets of molecules. For instance, chemists and physicists wanted in the 20th century to model the magnetic properties of molecules in which there were two or more "magnetic centers." Simple energy expressions, models essentially, called spin Hamiltonians (sometimes Heisenberg Hamiltonians) were first proposed for these systems, quite intuitively. Before their relationship to the underlying exact expression for the energy was clarified.

How quickly were these phenomenological Hamiltonians adopted! Their parameters became tangible; they could be tuned. Something else is at work here (aside from appealing simplicity, an aesthetic criterion for theory adoption). This is the phenomenon of reification; the model is so useful, so widely used, that people begin to think the model represents directly reality. We will return to a detailed consideration of the spin Hamiltonians in part C.

But the theoreticians had to clarify the link between the phenomenological "Hamiltonians" and the exact ones. It was not easy; but it is only then that they could be used with perfect mental comfort.

Some modern machine-learning predictive tools may be considered as phenomenological. For instance, if a computer program is given ("learns" from) a huge set of measurements or computational results of thousands of molecules, and establishes somehow complex connections between geometric features of these molecules (separation of atoms, intramolecular N-atom patterns) and the property under discussion, the resulting predictive tool might be considered as purely phenomenological. The nature of the connection found by the computer may be indiscernible to its operators, yet no less powerful. It is not likely that one will ever rationalize the so-obtained structure–property relationships from first principles, as was possible for spin Hamiltonians. Or to put it another way, it is unlikely that the machine learning program will spit out the suggestion of a Heisenberg spin Hamiltonian for analysis of magnetic data. Or will it? As we will discuss, the translation of the process of AI methodologies into human language and ideas, is an active field. Perhaps the future of AI.

The third class of numerical simulations concerns the application, on a computer, system by system, of the correct equations the quantum mechanical system is supposed to obey. It can't be done exactly, but well-defined approximations can get you close. This is the meat and potatoes of our community of quantum chemists, and of the profession we have spawned, of computational chemists.

The triangle with which we began this paper ignores the essential setting of theory in the sciences, surrounded by and intertwining with experiment. We will return in part C to the imperatives that experiment imposes. Here we recognize that many of the demands addressed by experimentalists to quantum chemists are quantitative: "what may be the value of this reaction barrier, or the dipole moment of this transient compound?" To the experimentalist, explanations or support from computations—both of these—of a rather unexpected number indirectly extracted from complex experiments, are welcome. We are in an intricate, amusing game—and to complicate it, happily enough experimenters do not allow theoreticians the monopoly of interpretation.

And as for predictions, the closer one gets to potential utility the more valued these become. In our utilitarian modern world, quantum chemists think they have to justify their existence (and wages); the demand from experimentalists for numbers is considered as providing this justification. Amusingly, some experimentalists justify their studies by the fact that they are or should be of interest to theoreticians.

The quantitative aspect of what some experimentalists need (as well as necessary calibrations of theory), has inclined quantum chemists to concentrate on the provision of accurate numbers. This requires the acquisition of "know how" knowledge—levels of CI, basis sets, functionals—valued specialist knowledge, indeed, highly citable and cited. A larger section of the community applies the available codes to specific problems. Both groups tend to reinforce a perspective that concentrates on getting the numbers right, let's call it a "numerist" attitude. There isn't much place for understanding in this partition of the community.

The emphasis on numbers within the trade is a curious mirroring of a simplistic attitude outside science, in which a facility of expressing oneself in a numerical way becomes associated with scientific authority.

As theory and computations in chemistry get better, there remains an important, secure place, even if a small one, for highest quality calculation (along the lines of the Lamb shift correction for hydrogen). Once one understands things pretty much correctly, one needs the exact agreement, or the lack of it, to guide one toward important little things that are missing. But often in chemistry, it's the "big understanding" we miss, such as "Is this compound going to have a lifetime of half an hour at room T? Is the triplet the ground state of this molecule by 1 cm^{-1} or 100?

We are not criticizing others as much as ourselves—we have been there, avoiding the labor of shaping understanding. Returning to an explanation seems qualitative and risky for many of us. It implies a return from numbers to words and/or images, with an attendant loss of security for anxious minds.

However, the person who understands the physics at work may feel rather secure with the steps he performs on returning from a heavy computation to a model, when he or she produces and handles and understands (for instance) a valence Hamiltonian with an effective on-site repulsion four-times smaller than the value given by the exact Hamiltonian in the simplified space where the model was formulated.

A15. From Theoretical to Computational Chemistry

How chemistry fell into simulation, with theory suffering along the way

The door to simulation opened naturally in our profession, without the least creak. It did so in stages, marked by the advent of electric calculators (today we can buy a serviceable cluster for what RH paid in 1963 for the marvel that could take square roots), followed by IBM-604s, still larger computers, then swinging to clusters of PC processing units. Along the way the characteristic noise of card-punches, the dry swish of accordion-folded paper, is fixed in the ears of the old-timers. Noises that came and went…

How much of our discipline, formerly called Theoretical Chemistry, has become computational! Announcements of open academic positions routinely specify "Computational Chemist", not even "Computational or Theoretical Chemist." New highly ranked (by the fashionable impact factor) journals have appeared, which give a prominent place to computation over theory – *Journal of Chemical Theory and Computation*, *Journal of Computational Chemistry.*

Indeed, today, general journals, such as the respected *Journal of the American Chemical Society,* publish very few purely theoretical (in contrast to computational) manuscripts.[52] It was not always that way. One of us (JPM) remembers that in 1984 he published in this journal a theoretical and computational paper, making use of Feynman diagrams,[53] and infinite summations of series, which has received 270 citations (significantly more than the mean number of citations for this journal). 3 years before he had proposed in the same journal a treatment of conjugated hydrocarbons reducing their electrons to their spins (150 citations). While it is considered very much appropriate, if not recommended, by popular chemistry journals to have a computational section in a paper, judging by what is published formal developments in theory are presumed not to merit the attention of a sufficiently broad audience. It seems that theory (in opposition to computation) needs to be "redirected" to specialized journals.

Figure 9. Drawing by Jean-Paul Malrieu.

The computational sobriquet abounds in the title of papers. And Editors require "state-of-the-art computation." Which means the most accurate, the most expensive. Lazy refereeing comments reject a paper because it does not reach the referee's image of best computational standards, simply saying that the basis set or the number of determinants are too small. Sometimes this sounds like a demand to use the heaviest hammer to kill a fly, when the answer could have been reasonably obtained from simpler models, taking into account the leading effects governing the property under investigation.[54] Here also fades the ancient impetus to reach a conclusion from the smallest number of arguments. Perhaps even "arguments" are not desired—the implicit demand is for numerical certainty, of course assumed to be insured by the use of the latest technologies.

The "Computational Details" sections of papers acquire the feeling of being written by a machine (and editors' programs to detect plagiarism probably red-flag these sections). The whole idea of the computational details section of a paper being conceived of as being purely factual and implicitly positive is worth thinking about.

Qualitative modeling in chemistry papers has by and large lost its place in these papers.[55] Though the careful reader will note how in a computational chemistry paper, sometimes following the advanced calculations reported, there often is in the discussion part of the same paper an orbital or physical reasoning argument "tacked on." We use the characterization because one gets that feeling—not quite pin-the-tail-on-the-donkey, but a perception that the bridging, logical argument relating good numbers with qualitative theory is missing. Actually, we are happy to see these stratagems, because they testify to the pervasive pedagogical imperative, which is much more difficult for an intellectual profession to shake than thinking. People want to understand, they want an explanation.

Yes, we bemoan that our discipline has largely shifted from theory to computation. With that shift, there is a loss of ambition—that understanding is possible. And, yes, there is increased technical utility in providing reliable numbers regarding definite properties of specific molecules. One question we shall address in the following concerns the thought mutations which may be a consequence of addiction to the newly available technical tools.

There is way out of this pessimism, formed by thinking of another dictum of which a reviewer of this paper reminded us, that of Ernie Davidson "Getting the right answer for the right reason."[56] This feels right, and certainly supports the importance of theory. And of the need for high level computations/simulations—the cases where we may be certain that we get the right answer for the right reasons without numerical tests are rare in our discipline. More importantly, often (witness in Part C a journey of this kind by JPM) the best calculations are intimately intertwined with the formation of real understanding.

A16. Calculemus

That computation is a panacea, a way out of a messy world, is an old, seductive, idea. Gottfried Wilhelm Leibniz, one of the greatest of philosophers (and through his independent invention of the calculus, a heroic Figure for theoreticians of chemistry and physics) wrote at 19 his "De Arte Combinatoria." Over the next decade he expanded his thinking to a *characteristica universalis,* an algebraization of thought. Here is his optimistic characterization of what his system would accomplish:

"[…] if controversies were to arise, there would be no more need of disputation between two philosophers than between two calculators. For it would suffice for them to take their pencils in their hands and to sit down at the abacus, and say to each other (and if they so wish also to a friend called to help): Let us calculate."[57]

Leibniz had more than abacus in mind; he also invented a simple calculating machine, shown in Figure 10.

Figure 10. A replica of the Stepped Reckoner, Instrumentum Arithmeticum of Leibniz (original is in the *Gottfried Wilhelm Leibniz Bibliothek, Niedersächsische Landesbibliothek, Hannover, image reproduced by permission*).

A17. The New Wave: Machine Learning and Artificial Neural Networks

An imperfect introduction to the current implementations of artificial intelligence in theoretical chemistry

In recent years there have appeared efficient and moderately reliable predictive tools which no longer rely on quantum mechanical theory. And which may replace our heavy quantum chemical techniques. The first type of non-quantum predictive tool makes use of machine-learning. Here is a simple description of the process, placing it squarely in the evolution of science:

"One of the fundamental goals of science is the development of theories that can be used to make accurate predictions. Predictive theories are generated through the scientific method. Here, existing knowledge is used to formulate hypotheses, which are then used to make predictions that can be empirically tested, with the goal of identifying the hypotheses that make the most accurate predictions. The scientific method can be expressed mathematically by considering predictive theory as a function f that maps a set of input data x to a predicted outcome y. The function may be relatively simple, as in Newton's laws of motion, or it may be complex, as in models that predict the weather based on meteorological observations. The collection of known input(x) and output (y) values, called training data, may be generated through observations or controlled experiments. The goal of the scientist is to use such training data, as well as any other prior knowledge, to identify a function that is able to predict the output value for a new set of input data accurately. The process of identifying such a function from a set of known x and y values is called supervised learning."[58]

One starts from a set of molecules and of their properties (experimentally or computationally obtained). And instead of cogitating, or calculating quantum-mechanically, one perfects a program that asks a computer to find the inherently complex, often topology-based (nature of the atoms, distances, connections between them) relationships to observables. The program then acts on what the machine finds, without our intervention, driven only by the aim of better agreement with the training data. There is room for substantial ingenuity in devising the atomic or molecular indicators to be correlated by program. The nature of the "training set" for molecules or people[59] is very important.

Applied to a new molecular architecture, one outside the "training set", it is likely that the so-calculated properties will be reasonable. One might be more skeptical about the prediction of intrinsically delocalized properties, such as ionization potentials and excitation energies. But it turns out that these also can be correlated well.

Note in the above description of the methodology the appeal to the scientific method, and the identification of that process with the making of numerical predictions, the success or failure of which are the sole criterion of success. It goes without saying that we disagree with this caricature of hundreds of years of Theory seeking more than numbers; we will return to the subject.

In part B of this Essay we will give some specific examples of machine learning studies. Here, we mention one recent set of applications of machine learning, and this is to the derivation of many-body potentials.[60] These enter chemistry in a variety of ways as it is, serving to expedite and enable calculations in simulations of actual physical collisions in a gas. The potentials are basically "look-up" linear combinations of simple polynomials times powers of, say, the distance between two atoms in a molecule. In most cases, the shape of the function rests (or pretends to rest) on theoretical considerations in the way we have thought of theory. The new, precipitous step away from theory that one takes in machine-learned potentials is that—as far as we understand—one does not assume any form of the above-mentioned function. But fits the function from some complete set, without "prejudice" as to the underlying physics.

The intriguing (and sobering) outcome is that such potentials may be used in simulations that in fact yield realistic properties of complex systems. Should we care? And if the outcomes be reliable, would one blame the designer and purveyor of the methodology from doubting that some simple physical explanation, yet to be found might, just might, underlie the hard-won result?

In describing the use of machine learning for deriving molecular potentials we have sinned, exaggerating the degree of empiricism out there. In fact, most uses of machine learning in this field are more intelligent—they begin with a space of functions or functional forms derived on the basis of previous experience and physical intuition.[61] Then machine learning is used to save computational labor, to find the best function in that space. And, perhaps coupled with further artificial intelligence algorithms, suggests a new synthetic strategy.[62,63] If it saves the work of a graduate student, or gets to a physical prediction, we are all in favor of it. Though we'd still like to know what was learned, what can be taught, from that effective design.

A second type of non-quantum predictive tool comes from the application of artificial neural networks to problems of chemistry and physics. This fascinating tool [Figure 11 is a schematic diagram of the process.] seems different from the previously described machine-learning procedures. A neural network is established from trial and error connections, and a simple processing algorithm, driven only by seeking the greatest agreement between the data set and the answer. Wikipedia's characterization of artificial neural networks as "vaguely inspired by the biological neural networks that constitute animal brains." is on the mark. The physical structure of the neural network (the connections and inten-

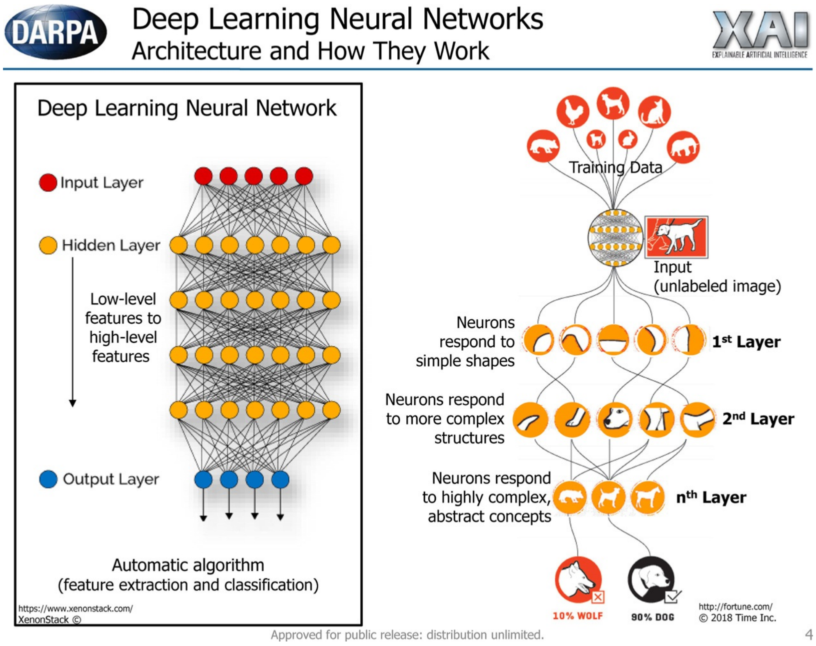


Figure 11. A schematic drawing of a deep learning neural network at left, and (equally schematic) the way it processes the picture of a dog. From a presentation by David Gunning on Explainable Artificial Intelligence, DARPA.[65]

sities of outputs) depends on the order in which the data have been provided for its construction. Or, more generally, on the underlying programmed architecture of the nodes and their connectivity. The neural network does not tell us how it is established (just as we do not know how we recognize faces or sentences), nor the way the recognition has been reached.[64]

You will notice that while we have not defined "Artificial Intelligence," we have in a fact slid down its seductive slope. We would guide the reader to the spirited discussion of Floridi of the history and future of the idea.[66] The original (1955) definition of McCarthy, Minsky, Rochester, and Shannon was that *"the artificial intelligence problem is taken to be that of making a machine behave in ways that would be called intelligent if a human were so behaving."*[67] Floridi prefers to

"conceptualise AI as a growing resource of interactive, autonomous, and often self-learning (in the machine learning sense) agency, that can deal with tasks that would otherwise require human intelligence and intervention to be performed successfully. This is part of the ethical challenge posed by AI, because artificial agents are "sufficiently informed, 'smart', autonomous and able to perform morally relevant actions independently of the humans who created them.""[68]

We will return to the ethical challenge of AI in practice in this world.

A18. Explainable AI

The field is moving on; perhaps there will be a meeting ground

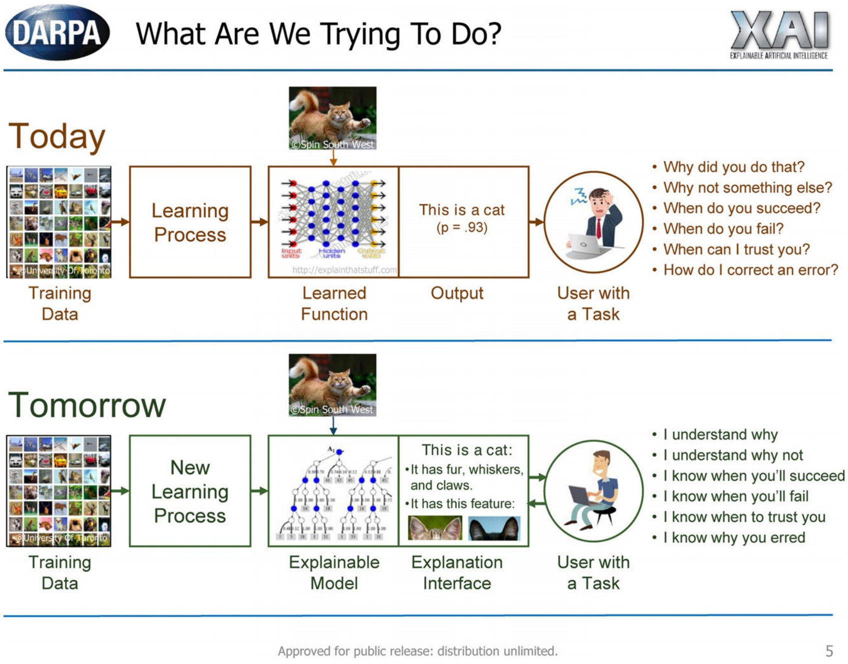
Part of the field is moving on to crafting the programs to tell us how an AI implementation (machine learning or neural networks) learns, how it does what it does. To cite a statement often attributed to E. Wigner, "It is nice to know that the computer understands the problem. But I would like to understand it, too".[69] No considerations yet of moral implications of uses of AI yet (to come in part B), but just learning how AI simulates is an outstanding problem, an area of research of the AI community.[70]

By way of example, DARPA, the US Defense Advanced Research Projects Agency, has had for several years a program (sponsoring university research) on Explainable Artificial Intelligence (XAI). Here is David Gunning's exposition of the program, with an accompanying very informative illustration, Figure 12. It is couched in "Defense Dept." language, but clearly the goal is more general:

Dramatic success in machine learning has led to a torrent of Artificial Intelligence (AI) applications. Continued advances promise to produce autonomous systems that will perceive, learn, decide, and act on their own. However, the effectiveness of these systems is limited by the machine's current inability to explain their decisions and actions to human users (Figure 1). The Department of Defense (DoD) is facing challenges that demand more intelligent, autonomous, and symbiotic systems. Explainable AI-especially explainable machine learning-will be essential if future warfighters are to understand, appropriately trust, and effectively manage an emerging generation of artificially intelligent machine partners.

The Explainable AI (XAI) program aims to create a suite of machine learning techniques that:

- Produce more explainable models, while maintaining a high level of learning performance (prediction accuracy); and
- Enable human users to understand, appropriately trust, and effectively manage the emerging generation of artificially intelligent partners.[71]



DARPA What Are We Trying To Do?

XAI EXPLAINABLE ARTIFICIAL INTELLIGENCE

Today

- Why did you do that?
- Why not something else?
- When do you succeed?
- When do you fail?
- When can I trust you?
- How do I correct an error?

Tomorrow

- I understand why
- I understand why not
- I know when you'll succeed
- I know when you'll fail
- I know when to trust you
- I know why you erred

Approved for public release: distribution unlimited.

5

Figure 12. A conception of present and intended state of AI. From a presentation by David Gunning on Explainable Artificial Intelligence, DARPA. Drawing reproduced from ref. [72].

Been Kim, at Google's "Brain" division, is a leading researcher in this direction. She describes her work directly as "*I am interested in designing high-performance machine learning methods that make sense to **humans**.*" and more technically as "*Her research focuses on improving interpretability in machine learning by building interpretability method for already-trained models or building inherently interpretable models.*"[72] There is an excellent description of her work in an interview by John Pavlus.[73]

Kyndi, a Silicon Valley start-up, is working to identify concepts, not just numbers. It says "*AI cannot be a "black box," as it so often is today. Explainable AI™ means that our software's reasoning is apparent to the user, and that the system can explain its rationale.*" It continues, "*The Kyndi AI Platform uses machine learning to streamline regulated business processes and offer auditable AI systems for enterprises and government. Kyndi's product exists because Deep Learning is a "black box" and cannot be used in regulated industries where organizations are required to explain the reasons for any decision.*"[74,75] One is grateful that the regulation is there, in driving research and development, as once California's automotive exhaust requirement drove the development of catalytic converters.

The work of one researcher in the field, Leman Akoglu, is aptly described as "Explainable AI: What Happens Inside the Black Box."[76]

Researchers close to our field are very well aware of the need for moving beyond maximized correlation to understanding. An example may be found In the recent work of McCloskey, Taly, Monti, Brenner and Colwell, whose concern is with the discovery of small molecules binding to proteins, loosely termed drug discovery. They write:

> *Deep neural networks have achieved state-of-the-art accuracy at classifying molecules with respect to whether they bind to specific protein targets. A key breakthrough would occur if these models could reveal the fragment pharmacophores that are causally involved in binding. Extracting chemical details of binding from the networks could enable scientific discoveries about the mechanisms of drug actions. However, doing so requires shining light into the black box that is the trained neural network model, a task that has proved difficult across many domains.*[77]

The authors provide some guideposts for moving ahead. We will return to this work in a later section.

What Part B Holds

The foregoing part of our tripartite Essay could be seen as an attempt to establish a kind of topology of neighboring, partially overlapping, but definitely non-coincident concepts—theory, understanding, modeling, and simulation. We have no pretentions to sophistication in epistemology, in a way the underlying philosophical discipline. Consider what we have done as setting the stage in pretty plain language for a discipline-moored discussion—we try to understand in a reflective way what people mean by the concepts decorating the points of the triangle in chemical theory.

Trying hard to be neutral, we have laid out, in purely descriptive manner, a spectrum of simulation tools. Their status regarding their relation to theory is very different. Still, all claim, with numerical proofs, to be efficient predictors. And before the last few sections, we wrote of "old-fashioned" theory, explanation, and understanding, manifest in its ties with teaching.

It is time now to confront these two worlds. The second part of the Essay will have less philosophy and definitions in it, jumping quickly into confrontation, on several fronts. A sketch of Roald's being whipped by simulation goes on to what is happening to quantum chemistry today. Outside our field, we do not resist a tour through the multitude of ways, from trivial to dangerous, in which AI has penetrated every aspect of society. We will try to repair the imbalance of our argument, and give the other side a chance to make their optimistic case. But can we do so effectively? You will have to judge.

And as we go on we heed the observation of a thoughtful reader, Santiago Alvarez, who reminded us not be categorical in our typology: "As you say, there is a spectrum. Are there differential positions in this spectrum for theorize, understand, predict, simulate, reproduce, model, idealize?"

Acknowledgements

We are grateful to the many colleagues and friends who have given us their criticism, comments, and supplied us with literature and drawings. Among there are Santiago Alvarez, Alán Aspuru-Guzik, Bo Chen, Odile Eisenstein, Alexander Frank, Hartmut Frank, N. Guihéry, Johannes Hachmann, J.-L. Heully, Gail Holst-Warhaft, Paul B. Kantor, Frank Neese, J.-M. Lévy-Leblond, Dasari L. V. K. Prasad, Grant Rotskoff, Zellman Warhaft, Michael Weisberg, and the reviewers of this paper.

Conflict of interest

The authors declare no conflict of interest.

---

[1] M. Weisberg, *Simulation and Similarity: Using Models to Understand the World*, Oxford University Press, Oxford, **2013**.

[2] "No Understanding without Explanation": M. Strevens, *Stud. Hist. Philos. Sci.* **2013**, *44*, 510–515.

[3] C. A. Coulson, Proceedings of the R. A. Welch Foundation Conferences on Chemical Research XVI; R. A. Welch Foundation: Houston, TX, **1973**; p. 117; "The purpose of computing is insight, not numbers" is the motto of Richard Hamming's important 1962 book, R. W. Hamming, *Numerical Methods for Scientists and Engineers*, McGraw-Hill, New York, **1962**.

[4] "Molecular Scientists and Molecular Science: Some Reminiscences": R. S. Mulliken, *J. Chem. Phys.* **1965**, *43*, S2.

[5] "*C'est seulement parce qu'on accepte le risque de l'erreur qu'on peut récolter de nouvelles découvertes.*" R. Thom, É. Noël, Predire n'est pas expliquer," Flammarion, **2009**.

[6] Aldous Huxley's dystopian novel with this title appeared 106 years ago. It remains, for good reasons, on many lists of the best English novels of all time.

[7] When we refer to a field of some historical import and cohesion, we will capitalize it. Yes, this is to show respect to the field and its practitioners.

[8] *Dictionnaire Européen des Philosophies* (Ed.: B. Cassin), Le Seuil, Paris, **2004**.

[9] R. L. Lee, A. B. Fraser, *The rainbow bridge: rainbows in art, myth, and science*, Penn State Press, College Station, **2001**.

[10] There was precedent for the explanation, as always. Theodoric of Freiberg and Roger Bacon played important roles here; the History section of the Wikipedia article: https://en.wikipedia.org/wiki/Rainbow is a good introduction. It was Newton who was responsible for explaining the colors.

[11] "We acknowledge opinions that deserve to be heard, even as they argue with prevailing consensus: on evolution, that Darwinian theories of mind and consciousness are incomplete:" T. Nagel, *Mind and Cosmos: Why the Materialist Neo-Darwinian Conception of Nature Is Almost Certainly False*, Oxford University Press, Oxford, **2012**; on origins of life: "*Animadversations of a Synthetic Chemist: A Critical Essay on Origin of Life Proposals*": J. M. Tour, *Inference* **2016**, 2; http://inference-review.com/article/animadversions-of-a-synthetic-chemist; also "*Two Experiments in Abiogenesis*": J. M. Tour, http://inference-review.com/article/two-experiments-in-abiogenesis; on anthropogenic climate change: "*CO_2: The Greatest Scientific Scandal of Our Time*," Z. Jaworowski, EIR (Executive Intelligence Review), March 16, 2007, 38–57; "*Climate Change: Incorrect information on pre-industrial CO_2*," Z. Jaworowski, http://www.warwickhughes.com/icecore/. On climate change, leads to the counterargument to the climate change doubters may be found in I. H. H. Zabel, D. Duggan-Haas, R. M. Ross, B. Brown-Steiner, A. F. Moore, *The Teacher-Friendly Guide to Climate Change*, Paleontological Research Foundation, Ithaca, **2017**.

[12] "This breaking force is not an accidental predicate but the very structure of the written text…the sign possesses the characteristic of being readable even if the moment of its production is irrevocably lost and even if I do not know what its alleged author-scriptor consciously intended to say at the moment he wrote it, that is, abandoned it to its essential drift." "Signature Event Context" in *Limited Inc*, J. Derrida, translated by Samuel Weber, Jeffrey Mehlman, Northwestern University Press, Evanston, **1981**, p. 5. "Cette force de rupture n'est pas un prédicat accidentel, mais la structure même de l'écrit…. Il appartient au signe d'être en droit lisible même si le moment de sa production est irrémédiablement perdu et même si je ne sais pas ce que son prétendu auteur-scripteur a voulu dire en conscience et en intention au moment où il l'a écrit, c'est-à-dire abandonné a sa dérive essentielle." "Signature Evénement Contexte" in *Marges de la Philosophie*, J. Derrida, Les Editions de Minuit, Paris, **1972**, p. 377.

[13] Readers knowledgeable in philosophy will see through the low level of our analysis. We more or less equate understanding and explanation, seeing differences twixt the two mainly in expression. Yet that difference has been the subject of philosophical debate that continues to this day. So the *Journal for General Philosophy of Science* has just published (Volume 50, Issue 3, 2019) a special issue on "Scientific Understanding," edited by J. Faye, H. de Regt. But perhaps we are not off the trend line. So the editors begin their essay: *Among philosophers working on scientific explanation, there seems to be a growing consensus that explanation is somehow connected to some form of understanding. This stands in sharp contrast to Hempel's view, according to which understanding is a psychological notion with no constructive bearings on an acceptable notion of scientific explanation.* "Introduction, Norms, Naturalism, and Scientific Understanding," J. Faye, H. de Regt, *J. Gen. Philos. Sci.* **2019**, *50*, 323.

[14] One of us has written elsewhere, with Saundra McGuire, about teaching and learning in the context of chemistry: "Learning and Teaching Strategies," R. Hoffmann, S. Y. McGuire, *Am. Sci.* **2010**, *98*, 378–382.

[15] See references in the Wikipedia pages on "Learning by Teaching," https://en.wikipedia.org/wiki/Learning_by_teaching, and the Lancasterian system, https://en.wikipedia.org/wiki/Monitorial_System, and references within. See also "Reinforcing Math Knowledge by Immersing Students in a Simulated Learning-By-Teaching Experience": *J. Artif. Intell. Educ.* **2014**, *24*, 216–250.

[16] D. L. Goodstein, *Feynman's Lost Lecture: The Motion of Planets Around the Sun*, Norton, New York, **1996**, p. 45.

[17] "Aquinas and His Understanding of Teaching and Learning," J. T. Ozoliņš in *Aquinas*, *Education and the East* (Eds.: T. B. Mooney, M. Nowacki), Springer, Dordrecht, **2013**, pp. 9–25. Paul B. Kantor also reminded us that the point about reasoning being awakened, not taught, was made earlier, in Plato's *Meno.*

[18] *Feu DIEU d'Abraham*, *DIEU d'Isaac*, *DIEU de Jacob, non des philosophes et des savants. Certitude. Certitude. Sentiment. Joie. Paix.* This is the **1662** "Mémorial" of Blaise Pascal. For its fascinating history see https://fr.wikipedia.org/wiki/M%C3%A9morial (Blaise Pascal).

[19] S. Singh, *Fermat's Enigma*, Anchor Books, New York, **1998**. p. 275.

[20] "Dispersion Forces: Neither Fluctuating nor Dispersing," D. J. Truhlar, *J. Chem. Educ.* **2019**, *96*, 1671–1675.

[21] E. F Taylor, J. A. Wheeler, *Spacetime Physics*, W. H. Freeman, San Francisco, **1966**, p. 60 The extended quotation is just as relevant: "WHEELER'S FIRST MORAL PRINCIPLE. *Never make a calculation until you know the answer.* Make an estimate before every calculation, try a simple physical argument (symmetry! invariance! conservation!) before every derivation, guess the answer to every puzzle. Courage: No one else needs to know what the guess is. Therefore make it quickly, by instinct. A right guess reinforces this instinct. A wrong guess brings the refreshment of surprise. In either case life as a spacetime expert, however long, is more fun!".

[22] Nero Wolfe was the fictional detective in Rex Stout's mid-20th century popular fiction. René Magritte, R. B. Woodward liked him.

[23] Charles A. Coulson, quoted by R. McWeeny, http://www.quantum-chemistry-history.com/Coulson1.htm.

[24] see also the articles on "Theoretical Terms in Science" by Holger Andreas, and "Theory and Observation in Science" by James Bogen, in the Stanford Encyclopedia of Philosophy, https://plato.stanford.edu/, accessed **2019**.

[25] "Nearly Circular Reasoning," R. Hoffmann, *Am. Sci.* **1988**, *76*, 182.

[26] "What Might Philosophy of Science Look Like If Chemists Built It?" R. Hoffmann, *Synthèse* **2007**, *155*, 321–336.

[27] James Bogen in Stanford Dictionary of Philosophy, article on "Theory and Observation in Science." https://plato.stanford.edu/

[28] E. F. Keller, *Making Sense of Life: Explaining Biological Development with Models, Metaphors, and Machines*, Harvard University Press, Cambridge, **2002**.

[29] See articles on "Realism and Theory Change in Science" by S. Psillos, and "Scientific Realism" by A. Chakravarti, in Stanford Encyclopedia of Philosophy, https://plato.stanford.edu/, accessed **2019**. The original article that prompted this discussion was "A Confutation of Convergent Realism", L. Laudan, *Philos. Sci.* **1981**, *48*, 19–49.

[30] See also "Eloge des Theories Fausses," J.-M. Lévy-Leblond, *Eur. J. Phys.* **1980**, *1*, 248–252.

[31] "Why Buy That Theory?," R. Hoffmann, *Am. Sci.* **2003**, *91*, 9–11.

[32] See Wikipedia article on "Tests of general relativity:" https://en.wikipedia.org/wiki/Tests of general relativity.

[33] "Sauter à pieds joints sur les calculs, grouper les opérations, les classer suivant leur difficulté et non suivant leur forme, telle est selon moi la mission des géomètres futurs." in C. Ehrhardt, *Évariste Galois, la fabrication d'une icône mathématique*, *En temps et lieux* (in French), Editions de l'Ecole Pratiques de Hautes Etudes en Sciences Sociales, **2011**; A. Astruc, *Évariste Galois*, Grandes Biographies (in French), Flammarion. **1994**. For the context to Connes, see an interview with Alain Connes: www.alainconnes.org/docs/Inteng.pdf. For an amusing discussion on the French phrase Galois uses, see https://forum.wordreference.com/threads/sauter-%C3%A0-pieds-joints.177369/.

[34] "On the nature of turbulence": D. Ruelle, F. Takens, *Commun. Math. Phys.* **1971**, *20*, 167–192; See Wikipedia article on "Attractor," https://en.wikipedia.org/wiki/Attractor, accessed **2019**.

[35] "Quand les Machines Font des Mathématiques": J.-P. Delahaye, *Pour la Science*, **2020**, *510*, 72–77; K. Chaudhuri, R. Salakhutdinov, Eds., Proceedings, International Conference on Machine Learning, 9–15 June **2019**, Vol. 97, http://proceedings.mlr.press/v97/; S. Colton, *Automated Theory Formation in Pure Mathematics*, **2002**, Springer.

[36] P. A. M. Dirac, *Proc. R. Soc. London Ser. A* **1929**, *123*, 714–733.

[37] The first textbook of quantum chemistry, Hans Hellmann's *Einführung in die Quantum Chemie*, was published in the year RH was born. A good introduction to the present status of the field is: L. Piela, *Ideas of Quantum Chemistry*, Elsevier, Amsterdam, **2007**.

[38] Regarding the coupled cluster method; the original reference is: "Short range correlations in nuclear wave functions", H. Coester, F. Kummel, *Nucl. Phys.* **1960**, *17*, 477–485. Born in Nuclear Physics, it rapidly became a standard method in Quantum Chemistry.

[39] The shape of the Coulomb hole was established in "Interpretation of the rapid convergence of correlated Wave functions", T. L. Gilbert, *Rev. Mod. Phys.* **1963**, *35*, 491–494. For the quantum chemistry treatments incorporating this effect one may refer to "Wave functions with terms linear in the interelectronic coordinates to take care of the correlation cusp I General theory", W. Kutzelnigg, W. Klopper, *J. Chem. Phys.* **1991**, *94*, 1985–2001.

[40] R. Thom, É. Noël, *Predire n'est pas expliquer*," Flammarion, Paris, **2009**.

[41] See "auspices ex tripudiis," in the Wikipedia entry on "Augury".

[42] Film (two sequels) directed by The Wachowskis https://en.wikipedia.org/wiki/The Matrix.

[43] Film directed by Alex Garland: https://en.wikipedia.org/wiki/Ex Machina (film).

[44] N. Goodman, *Languages of Art*, Hackett, Indianapolis, **1976**, p. 171.

[45] The fundamental idea of the model is that molecular geometries result from the repulsion of the valence electron pairs disposed around an atom in a molecule, R. J. Gillespie, *Molecular geometry*, Van Nostrand Reinhold, London, **1972**; "A Defense of the Valence Shell Electron Pair Repulsion (VSEPR) Model," R. J. Gillespie, *J. Chem. Educ.* **1974**, *51*, 367; "The electron-pair Repulsion Model for Molecular Geometry," R. J. Gillespie, *J. Chem. Educ.* **1970**, *47*, 18, An alternative explanation follows from molecular orbital theory, with a detailed analysis of the changes in energy of molecular orbitals (in turn dependent on the energies of the underlying atomic orbitals) as a function of geometry: "Molecular Geometry: Bonded versus Nonbonded Interactions," L. S. Bartell, *J. Chem. Educ.* **1968**, *45*, 754; B. M. Gimarc, *Molecular Structure and Bonding*, Academic Press, New York, **1979**; T. A. Albright, J. K. Burdett, M.-H. Whangbo, *Orbital Interactions in Chemistry*, 2nd Ed., Wiley, Hoboken, **2013**.

[46] "Crystal Structure of Homo-DNA and Nature's Choice of Pentose over Hexose in the Genetic System," M. Egli, P. S. Pallan, R. Pattanayek, C. J. Wilds, P. Lubini, G. Minasov, M. Dobler, C. J. Leumann, A. Eschenmoser, *J. Am. Chem. Soc.* **2006**, *128*, 10847–10856; "Searching for Nucleic Acid Alternatives: A. Eschenmoser, *Chimia* **2005**, *59*, 836–850, reprinted in: *Chemical Synthetic Biology* (Eds.: P. L. Luisi, C. Chiarabelli), Wiley, New York, **2011**, pp. 7–45; "The search for the chemistry of life's origin:" A. Eschenmoser, *Tetrahedron* **2007**, *63*, 12821–12844.

[47] "Planar Tetracoordinate Carbon": R. Hoffmann, R. W. Alder, C. F. Wilcox, Jr., *J. Am. Chem. Soc.* **1970**, *92*, 4992.

[48] "Ockham's Razor and Chemistry": R. Hoffmann, V. I. Minkin, B. K. Carpenter, *Bull. Soc. Chim. Fr.* **1996**, *133*, 117–130; Reprinted in *Hyle* **1997**, *3*, 3–28.

[49] "Le but ultime de la science n'est pas d'amasser indistinctement les données empiriques, mais d'organiser ces données en structures plus ou moins formalisées qui les subsument et les expliquent. Dans ce but, il faut avoir des idées «a priori» sur la manière dont se passent les choses, il faut avoir des modèles. Jusqu'à présent, la construction des modèles en Science a été avant tout une question de chance, de «lucky guess». Mais le moment viendra où la construction des modèles elle-même deviendra, sinon une science, du moins un art." R. Thom, É. Noël, *Predire n'est pas expliquer*," Flammarion, Paris, **2009**.

[50] Plato, *The Republic*, tr. R. E. Allen, Yale, New Haven, **2006**.

[51] "From Cubism to Futurism to Suprematism: The New Realism in Painting," K. Malevich, http://theoria.art-zoo.com/from-cubism-and-futurism-to-suprematism-the-new-realism-in-painting-k-malevich/.

[52] It is fair to notice that this journal hosted a special issue (piloted by Sason Shaik) to celebrate the 100th anniversary of the fascinating proposition by Lewis of electron pairing in local bonds and lone pairs. This crucial idea, formulated before quantum mechanics and later confirmed by it, was appropriately honored in a general chemistry review. Notice that it was a purely qualitative rationalization.

[53] Feynman diagrams are very useful to represent perturbative expansions; popular in Physics, they are not of common use among chemists.

[54] The requirement of application of highest level technical resources (usually demanding the greatest computational power, may constitute a barrier to the access of scientists from less-developed countries to the marketplace of so-called "state-of-the-art" computational chemistry.

[55] The story how one of us (RH) kept doing good chemistry for 25 years with the kinds of calculations to which all other calculations are superior is worth telling elsewhere. Also how colleagues' dissatisfaction with RH's level of theory was taken out on his students more than him.

[56] "The Right Answer for the Right Reason: My Personal Goal for Quantum Chemistry," E. R. Davidson, *Annu. Rev. Phys. Chem.* **2019**, *70*, 1–20.

[57] "…*quando orientur controversiae, non magis disputatione opus erit inter duos philosophus, quam inter duos computistas. Sufficiet enim calamos in manus sumere sedereque ad abacos, et sibi mutuo (accito si placet amico) dicere:* ***calculemus*** …" *De arte characteristica ad perficiendas scientias ratione nitentes* in Die philosophischen Schriften von Gottfried Wilhelm Leibniz (Ed.: C. I. Gerhardt) (7 vols. 1875–1890) VII 125.

[58] "Machine Learning in Materials Science: Recent Progress and Emerging Applications," T. Mueller, A. G. Kusne, R. Ramprasad, in *Reviews in Computational Chemistry, Vol. 29* (Eds.: A. L. Parrill, K. B. Lipkowitz), Wiley, New York, **2016**, Ch. 4, pp. 186–273.

[59] "AI is learning from Humans, Many Humans," C. Metz, *New York Times*, Aug. 16, 2019, https://www.nytimes.com/2019/08/16/technology/ai-humans.html.

[60] "Machine Learning Predictions of Molecular properties: Accurate Many-Body Potentials and Nolocality in Chemical Space, K. Hansen, F. Biegler, R. Ramakrishnan, W. Pronobis, O. A. von Lilienfeld, K.-R. Müller, A. Tkachenko, *J. Phys. Chem. Lett.* **2015**, *6*, 2326 – 2331.

[61] "Physically informed artificial neural networks for atomistic modeling of materials," G. P. Purja Pun, R. Batra, R. Ramprasad, *Nat. Commun.* **2019**, *10*, 2339.

[62] "Predicting reaction performance in C-N cross-coupling using machine learning," D. T. Ahneman, J. G. Estrada, S. Lin, S. D. Dreher, A. G. Doyle, *Science* **2018**, *360*, 186 – 190.

[63] "Universal Chemical Synthesis and Discovery with "The Chemputer"", P. S. Gromski, J. M. Granda, L. Cronin, *Trends Chem.* **2020**, *2*, 4 – 12.

[64] For a description of some problems in AI approaches, see "What are the limits of deep learning?", M. M. Waldrop, *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 1074 – 1077.

[65] https://www.darpa.mil/program/explainable-artificial-intelligence. XAI Program Update. David Gunning.

[66] "What the Near Future of Artificial Intelligence Could Be," L. Floridi, *Philos. Technol.* **2019**, *32*, 1 – 15.

[67] "A proposal for the Dartmouth summer research project on artificial intelligence, August 31, 1955" J. McCarthy, M. L. Minsky, N. Rochester, C. E. Shannon, *AI Magazine* **2006**, *27*, 12.

[68] "On the morality of artificial agents, " L. Floridi, J. W. Sanders, *Minds Machines* **2004**, *14*, 349 – 379.

[69] The statement, with slight variations, appears in many places in the literature (for instance in an Opinion piece in *Nature* **2000**, *403*, 345), but we have been unable to find the source.

[70] "One Giant Step for a Chess-Playing Machine: S. Strogatz, New York Times, Jan. 8, **2019**.

[71] https://www.darpa.mil/program/explainable-artificial-intelligence.

[72] https://beenkim.github.io/; https://ai.google/research/people/106045.

[73] "A New Approach to Understanding How Machines Think", J. Pavlus, https://www.quantamagazine.org/been-kim-is-building-a-translator-for-artificial-intelligence-20190110/.

[74] https://kyndi.com/products/.

[75] "Researchers Seek Smarter Paths to A. I." S. Lohr, New York Times, June 21, **2018**, p. B4.

[76] "Explainable AI: What Happens Inside the Black Box," S. Barsotti, https://www.cmu.edu/news/stories/archives/2018/october/explainable-ai.html.

[77] "Using attribution to decode binding mechanism in neural network models for chemistry," K. McCloskey, A. Talya, F. Montia, M. P. Brenner, L. J. Colwell, *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 11624 – 11629.

Manuscript received: March 4, 2019
Accepted manuscript online: November 1, 2019
Version of record online: May 27, 2020