

In G. Guo & C. Liu (eds.), *Scientific Explanation and Methodology of Science*, 2014, pp. 138-147.

## UNDERSTANDING AS INTEGRATION OF HETEROGENEOUS REPRESENTATIONS<sup>1</sup>

SERGIO F. MARTÍNEZ

*Instituto Investigaciones Filosóficas,  
Universidad Nacional Autónoma de México,  
sfmar@filosoficas.unam.mx*

The search for understanding is a major aim of science. Traditionally, understanding has been undervalued in the philosophy of science because of its psychological underpinnings; nowadays, however, it is widely recognized that epistemology cannot be divorced from psychology as sharp as traditional epistemology required. This eliminates the main obstacle to give scientific understanding due attention in philosophy of science. My aim in this paper is to describe an account of scientific understanding as an emergent feature of our mastering of different (causal) explanatory frameworks that takes place through the mastering of scientific practices. Different practices lead to different kinds of representations. Such representations are often heterogeneous. The integration of such representations constitute understanding.

1. The preeminence of knowledge as the main epistemic aim has been often formulated in the philosophy of science by saying that either understanding is a mere subjective phenomenon, or else it is subordinated to the establishment of factive knowledge.<sup>2</sup> Hempel, for example, thought that understanding could only be a subjective (psychological) phenomenon, and thus, that it was not apt to play a role in the philosophy of science (Hempel 1965). To the extent that understanding required making reference to persons and concepts understanding was a pragmatic concept with no place in epistemology (and in the philosophy of science). One of the most influential critics and followers of Hempel, Wesley Salmon, argued that many scientific explanations did not fit Hempel's model, to the extent that the identification of causal structures articulated in mechanisms was required to understand the force of explanations. Salmon considered that such identification of mechanisms was

---

<sup>1</sup>The research for this paper has been carried with support of grant CONACYT 133345.

<sup>2</sup>That knowledge is factive means that it is impossible to know that p unless "p" is true. This is the sense in which knowledge is tied-down to the truth. If one says that knowledge is not factive one claims that, for example we can have knowledge of ideal gases, knowledge that an ideal gas with a given molecular weight has a given property, for example, even though we accept that there are no ideal gases and that the proposition that an ideal gas with such molecular weight has a given property does not correspond to the facts, but it generates explanation.

“the key for our understanding of the world” (p. 160, 1984). Thus, understanding for Salmon is no more a mere subjective phenomena, but ultimately is an expression of the way Salmon thinks factive knowledge is structured. Understanding is objective to the extent that the paramount epistemic aim is the identification of causes.

This is a rather common view in the philosophy of science. Peter Lipton suggests something similar (Lipton 2004, p. 30). In this view, understanding can be taken to be something like the horizon of the causal structure of the world. That such horizon is an important dimension of understanding is no doubt correct, but it certainly does not exhaust what we try to capture in an epistemology that takes understanding as central. As Hempel was already aware, explanations of laws do not relate them to facts, but to other laws. Hempel suggested that explanations of laws consisted in situating the given law with respect to other laws, in a way that cannot be modeled as a deductive relation. Nor it reduces to processes of identification of causes.

Michael Friedman has proposed a view of understanding that picks up on Hempel’s idea and overcomes some of the shortcomings of the causal view of understanding. For him understanding is related to the reduction of the number of phenomena that we accept as independent facts (Friedman 1974). Friedman shows that the idea of independence of facts is a useful and productive notion in the context of some important physical theories, but the point for us is that the unification that generates understanding for Friedman depends on some a priori structure implicit in the mathematical formalism of the theories in question. Understanding is non-factive but it is restricted in any case to the sort of theories in which such sort of formalism is relevant. More importantly for us now, it does not seem to capture some notions of understanding that are common among scientists and that are related with the value attributed to the integration of explanations in a non-foundationalist manner (see below for an elaboration of this point).

**2.** In contemporary epistemology there are also several attempts to characterize understanding as a central epistemological concept.<sup>3</sup> A detailed proposal is that of Kvanvig 2003. For Kvanvig understanding is a mental state constituted by beliefs that have to be true. The difference with knowledge according to him is that understanding is “the internal “seeing or appreciating of explanatory relationships in a body of information” (p. 198, 2003). This idea is formally close to the non-causal unificationist characterization of understanding, like Friedman’s. They do not have to deny the heterogeneity of explanatory resources (causes in particular) but one way or another, they have to rely on the unificatory power of some (mathematical)

---

<sup>3</sup>I will proceed without apologies and hope that I will be vindicated once the reader understands the sense in which understanding is not tied down to the truth, but nonetheless is indirectly supported by the facts.

models in one case or some “internal seeing” in the other, to carry out the integration of resources and the desired characterization of understanding.

Whereas Friedman considers that understanding can be characterized by the sort of unification we can achieve through a reduction of independent facts, Kvanvig thinks that understanding can be spell out as some sort of coherence of a proposition to a whole body of information. If we say, I know that  $2 + 2$  or, we know that the Romans conquered most of Europe; we tend to think that we know an individual proposition. But if we say, I understand  $2 + 2$  or, I understand that the Romans conquered most of Europe, it suggests that what we are doing is somehow fitting the proposition into an integrated and coherent body of information.

The idea that understanding involves situating, or making coherent a given claim as part of a more comprehensive whole is a widespread idea and we agree that this is an important component of a fruitful epistemological concept of understanding. But as with independence of facts, something else have to be said in order to capture the non-aggregativity that seems to distinguish understanding from knowledge. Coherence is not enough. False theories are notoriously composed of coherent beliefs. Alchemy was a sophisticated set of coherent beliefs. Kvanvig is able to overcome obvious objections to understanding as coherence to the extent that he relies on the factive character of understanding. For example, Alchemy does not endow understanding to a proposition because the beliefs in question are largely false.

But overcoming the objection in this way excludes precisely the sort of possibility we are interested in bringing to the fore, it leaves out the the way in which coherent bodies of beliefs are the expression of constraints and standards articulated (often implicitly) in practices.

**3. Understanding and the Management of Perspectives.** Clearly understanding could not be identified with knowledge of the conjunction of true propositions constituting the body of beliefs with respect to which we understand a given proposition, the proposal that would be natural to make if we take understanding to be factive. Too often this body of beliefs include a false belief. That is why Kvanvig suggests that the organic feature of understanding is related to a conjunction of *central* propositions that are true, even if some peripheral propositions are false. But this suggestion leaves out of the question what are the norms that allow us to distinguish central from peripheral propositions, and more to the point, it fails to capture important uses of the concept of understanding in science. Consider the role played by diagrams in Euclidean Geometry. It seems that if we are willing to say that diagrams play a non-reducible (to propositions) epistemic role, then it is clear that diagrams have to be characterized as something more than conjunctions of propositions (see below). Analogously, state-charts (that are increasingly used in computer science, molecular biology, robotics, economics and other fields of knowledge) play an important role promoting the use of epistemic standards across disciplines, their role is not characterizable as promoting or identifying a core set of propositions, however.

For another example, take the sense in which Wimsatt argues for the importance of false models:

Scientific models are often as useful for the false assumptions they make as for what truths they might embody. One often designs a false model deliberately for use as a template to compare with data when one is interested not in how well it fits the data, but in the form of the residuals—where it does not fit, how and why. (Wimsatt 1987, 2007 chapter 6).

Wimsatt shows how models of blending inheritance can be used to show that an important objection to group selection (as a significant evolutionary force) starts with the assumption of a strong form of blending population that is equivalent to the assumption that there is no group structure. It is then no surprise that this assumption allow us to derive the consequence that group selection is not a significant evolutionary force. The general point that is relevant for our present discussion is that a false model can help us explain better the explanatory scope of a model, and in that sense it provide us with understanding.

Notice that a false model does not need to be (reducible to) a false belief. An image, a chart, or a diagram, are able to generate understanding beyond what true beliefs can, but they are not (necessarily) false beliefs. This is related to two interesting features of understanding that seem to be epistemically relevant.

The first one is that, as many philosopher, from Hempel to Kvanvig acknowledge, understanding seems always to involve a reference point (which is always some sort of organized whole, relative to which understanding takes place). We might praise the sophistication of a newspaper in handling the history of china, whereas a specialist might find too many errors and wrong characterizations of a given episode. We can say that a child in primary school has an admirable understanding of Darwin's theory, even though we would not consider such an understanding sufficient to pass an examination at the university level. We might find a proof of a theorem quite satisfactory in secondary school, but unacceptable in a course in undergraduate mathematics, and we might find that a proof based on a diagram is accepted in a course en theoretical physics, but the proof is not be accepted as a proof in a course of mathematical analysis.

The second feature that seems relevant for discussing the difference between knowledge and understanding is very well expressed in a recent article: “the hallmark of scientific understanding is the reduction of a natural phenomenon to simpler units (Regev A. Shapiro, *E. Nature*/vol. 419/26 sept. 202). Understanding is often associated with some sort of reduction, the sort of reduction that promotes intelligibility (not the sort of reduction aiming for foundations).<sup>4</sup> It is a sort of reduction

---

<sup>4</sup>One can distinguish two senses of reduction. A fundamentalist and a non-fundamentalist type. Fundamentalist reduction assumes that theories describe layers of reality that are neatly separated and are non-controversially (objectively) identifiable. Reduction consists of explaining away less fundamental layers of reality in terms of more fundamental ones. Non fundamentalist reduction

that often, as the quote suggests, is associated with finding some basic explanatory units covering a wide variety of phenomena. As a matter of fact, the aim of Regev and Shapiro is to suggest that the lack of an appropriate abstraction from which “the mountains of knowledge about the function, activity and interaction of molecular systems in cells” could be made intelligible is not allowing the sort of understanding that accompanies successful science. They suggest that such abstraction can be found in the ‘molecule-as-computation’ abstraction. Insofar molecular cell biology embraces such abstraction, they hope, the possibility of sharing, comparing and criticizing claims to knowledge would be simplified and consensus reached faster.<sup>5</sup> Understanding involves an standpoint, but such standpoint is not factive, it requires locating the appropriate abstraction.

Notice that both features are closely related. *The search for “simpler units” is not the search for foundations but the search for the right sort of understanding-promoting perspective.* The simple units are simple in the sense that are sufficiently abstract to serve as a common ground for a wide variety of explanations that are then seen as related to each other, and thus can ground further understanding. DNA-as-string and “ideal gas” are such sort of abstractions. Their value lies not in being true, but in providing relations among the different perspectives in such a way that a robust characterization of entitlements to inferences can be orchestrated.

Lest the reader think that such account of understanding plays a role only in some relatively strongly mathematical areas of science it might be worth recalling that for example Rudwick (in Rudwick (1985)) shows the crucial importance, that for the advancement of geology, had the invention of a series of techniques for the graphical representation of different aspects of the Earth. He convincingly shows that the consolidation of geology as a body of knowledge generating understanding owns a lot to the development of a visual language (charts, diagrams and other visual resources) that eventually made possible the representation of the history of the dynamics of geological strata; maps and other techniques of graphical representation cannot be seen as mere auxiliary tools used as illustrations of theory; rather they have to be seen as constitutive elements in a new kind of argument. Further more, important studies aiming to establish the spatial and visual origins of logical and mathematical reasoning provide very general kind of evidence for such a view. A related but independent sort of evidence comes from studies suggesting that reasoning in space (visuospatial reasoning) precedes language both philogenetically and ontogenetically (see Tversky 2005, for example). Such idea suggest that the

---

does not assume a neat layered reality, and considers that reduction is more a matter of finding appropriate abstractions allowing heterogeneous representations of collective phenomena, which in turn contribute to understanding of a complex reality (that is up for grabs). As Goldenfeld and Woese put it: “. . .the molecular reductionism that dominated twentieth-century biology will be superseded by an interdisciplinary approach that embraces collective phenomena” (Goldenfeld and Woese 2007).

<sup>5</sup>Regev and Shapiro have in mind abstract computer languages such as Petrinets, State-charts and other similar proposals. They claim that such languages, developed for the specification and study of systems of interacting computations can be used to represent biomolecular systems.

role of charts and diagrams and other representational tools are not dependent on (or reducible to) language.

4. Now let us go back the question posed above, how can false beliefs generate understanding? Catherine Elgin suggests the following answer: idealizations are designed to highlight subtle matters of fact. And they do so by exemplifying features they share with the facts (see Elgin 2007 and 2009). A paint sample card can exemplify a color or a rectangle; it depends on the function the card plays in a given context. In a classroom it might be taking as exemplifying a rectangle, but in the context of looking for the color we want to paint the kitchen the card exemplifies its color. The cognitive contribution of an exemplar consists in the exemplification of selected features, and thus, anything that exemplifies exactly those features can, in a suitable context, make the same contribution (see Elgin 2009). It is not difficult to see how this account can be extended to scientific idealizations. The ideal gas (according to a theory) exemplifies features that real samples share, and to that extent can be considered to make the same contribution to the phenomena in question, even though we cannot actually separate the contribution of the real system in (or from) the real environment. This seems to be part of what is non-factive understanding about, but it cannot be all. It makes the epistemic status of idealizations parasitic on theories that make sense of the facts. Elgin sees that, and thinks that is as it should be. If those theories are overthrown we lose our reason to accept the idealization they contain.

But this does not seem quite right. Often in the history of science idealizations far outlive the theories in which they were born. “Electron” for example, is a term used by many different theories, many of which are considered false nowadays. The idealization of an ideal gas is another example. Such idealizations were born in theories that have been long discarded (at least to some extent). Elgin could argue that the idealizations are useful insofar we accept theories that support them. One could argue that statistical mechanics support the idealization of an “ideal gas”. However, it is not obvious what this support amounts to. And in cases like the idealization of “electricity as a fluid” there is no specific theory in contemporary physics supporting the idealization. The idealization is supported by the analogy or water and electric circuits. In other cases, techniques or methods and not theories are the (main) support of successful idealizations in science. Take as an example the work of Carl Woese (and collaborators) leading to a revolution in our understanding of evolution (see Goldenfeld and Woese 2007). This is a clear case in which a set of technologies (of sequencing) have an impact on basic concepts like genome and species. Finally, as pointed out above, even a false idealization (unambiguously associated with a paradigmatically false theory), such as Darwin’s idealization of “blending inheritance”, is a powerful tool for understanding.

Idealizations are not as parasitic on theories as Elgin seems to think. To see what idealizations can be beyond its relation to specific theories let us consider again the proposal of Regev and Shapiro. Their idea is that ‘*molecule-as-computation*’ would

point to the properties that are relevant to their function in the explanations we seek (in the explanations that would advance understanding). They think that this is the right abstraction to the extent that this abstraction can account for the key properties of *phenomena characterized by heterogeneous representations* in different areas of research in molecular cell biology. Why it is thought that the idealization of molecule-as computation is the right abstraction is because it is claimed that has the capacity to integrate heterogeneous representations.

But this would not happen by sheer “convention”, or by making the idealization parasitic on a specific theory. In analogy to the way in which a cultural context makes clear that paint sample cards are intended to be used as exemplification of color, a scientific context, *which would need to be painstakingly constructed as supported by specific scientific practices (techniques, methods, models, and so on)*, would provide that conditions in which the exemplified properties would fit the abstraction (by projecting the structure of the idealized abstraction into the new field). For example, if the abstraction is to be of any use it should project the fundamental distinction between two levels of description for behavior: implementation (how the system is built) and specification (what the system does). Such distinction, as an expression of the given scientific context, would guide the ascription of biological function in a way that would constraint possible ambiguities. The ‘molecule-as-computation’ abstraction would thus function like the cultural context that in the case of the paint sample card makes plain to us its (paradigmatic) function.

The main claim of this paper can now be more clearly formulated. Scientific understanding can be best characterized as the integration of heterogeneous representations. The remainder of the paper is a first attempt to explain this claim. The first thing is to get clear on what are heterogeneous representations.

**5.** Heterogeneous representations are different representations of a thing or process that cannot be seen as aggregative descriptions, that is, they cannot be seen as representations of parts which can be seen as adding to a complete representation. If we represent a thermodynamic process using phenomenological thermodynamics we provide one sort of representation. If we characterize the same process using statistical thermodynamics we have a different representation of the same process. We might think that the phenomenological representation is subordinated to the statistical-thermo-dynamical one (or viceversa), in which case the representations are not heterogeneous. But if we think that both representations stand on its own, or at least, that both representations contribute to our understanding of the process in question (in a way that would need to be elaborated, of course), then these two representations would be heterogeneous. In this case, any one sort of representation by itself cannot capture the reality of the process in question. Of course, in order for this possibility to be taken seriously the assumption that there is, at least in principle, a single universal representational system has to go.

In theoretical contexts there is a tendency to think that different representations of the same process can ultimately be explained away and fit into a basic

representation, in other words, that the heterogeneity of representations is a transitory state of our knowledge. once we take seriously the way in which knowledge is organized in practices this is a highly questionable assumption.

Consider the example of the design of a computer chip.<sup>6</sup> There are many different type of factors that might go into its design considerations. Engineers have identified important dimensions of design, the point of view of control, logical gates, and of timing, among others. We could think that such way of approaching the problem answer to purely pragmatic considerations, and that radically different ways of representing the different dimensions which require attention in the design could be develop. This does not sound quite right. At least for engineers, such dimensions of design (of a chip) seem to be factive. The problem is addressed by using three separate representational systems: state-charts for the representation of control, circuit diagrams for the representation of gate information, and timing diagrams for the representation of timing. Each of these systems represents some structural features to which they are homomorphic.

But the same sort of mark represents different things in the different systems. For example, certain physical relationships among lines represent relationships among various aspects of the operating chip. But lines and relationships among lines have a different interpretation in the three different representational systems. In one system of representation a line represents wiring connections, in other possible transition between states, and in the other the value on a wire through time.

We might think that the need for using different systems of representation is a consequence of our cognitive limitations, and in particular a consequence of the sort of reasoning processes we are able to perform. This might be right, but this does not need to imply that the heterogeneity of representations is a mere subjective appearance. There is at least one sense in which the different representational systems cannot be part of one single system of representation. Suppose we want to translate one representation into the other. This would require finding a common abstraction for the different representations. But each different representation involves different levels of the granularity of information. The sort of information included in timing diagrams is of a sort that is not needed for the state diagrams laying out the algorithm being computed. We can say that even if cumbersome, the translation could be carried out. But what is crucial to establish the reality of the heterogeneity of representations is that in any case there is no unique translation between diagrams (see Fisler 1993). There are any number of circuits that will implement a state machine, and any number of timing diagrams that will be compatible with those circuits. At most we can aim to establish consistency of representations. But the heterogeneity of representations stands unquestioned.

The heterogeneity of representations is often related to availability of different abstractions and different sorts of representations-cum-reasonings. Even in mathematics, often the solution of problems requires the integration of different types

---

<sup>6</sup>This example is given in Barwise 1993.



of representation. So for example, Aaron Sloman shows in an article about mental diagrams that the way in which types of abstraction and types of reasoning combines is far from simple. Depending on the abstractions used we might emphasize or de-emphasize metrical or topological reasoning, differences that might have consequences for the way problems are framed and solutions found. He finds that often one requires to combine abstract spatial representations with abstract non-spatial representations (arithmetical and logical) in order to be able to transform possible solutions at one level of abstraction (purely spatial or topological) to lower-level solutions (with metrical information) by a sort of trial and error adding of details that would lead to a clear grasp of one particular solution. Simply using concrete spatial visualization capabilities do not seem to lead us to understanding of what a solution requires. This is a fairly common type of situation. The solution to a problem is really a path through different representations that allow us to move from one level of abstraction to another. Diagrams often serves as a guide through the forest of representations which are possible or available. I want to stress that diagrams are tools that are constitutive of specific practices that norm the the way in which abstractions and reasoning interact productively.

A related way of formulating the same point can be found in Netz 1999. Netz thinks that the diagrams of Euclidean geometry can be understood as practices that through their implicit norms promote a community of what we would end up identifying as Euclidean geometers. In particular such norms constrain the boundaries of what is taken to be a good inference, and to that extent diagrams are vehicles for logic, for the sort of logic that regulates the thematic inferences distinctive of Euclidean Geometry: “the diagram is not just a graph, in the sense of graph theory. It contains at least one other type of information, namely the straightness of straight lines; that points “stand on a line” is constantly assumed on the basis of the diagram (p. 34, Netz 1999). This assertion is closely related with our claim that the right abstractions, Euclidean diagrams in this case, are not subordinated to propositions, on the contrary, Netz tells us that “propositions are individuated by their diagrams”. To the extent that diagrams individuate propositions, diagrams play a role in identification of what is a good inference and thus in what is considered a proof. This is a way of suggesting that if we take seriously the importance in our cognition of visuospatial reasoning (as a sort of reasoning different from and not reducible to propositional reasoning) we are led to recognize the centrality of diagrammatic reasoning in achieving understanding.

More generally, understanding results from the integration of heterogeneous representations arising from the mastering of tools with representing capabilities (models, diagrams, charts). The tools and the representing capabilities are normed and form part of specific practices.

## References

- Anderson M, Meyer B, Oliver P. eds *Diagrammatic Representation and Reasoning*, Springer, 2002.

- Allwein G. and Barwise, J. e. (1993). *Working Papers on Diagrams and Logic*. Bloomington, Indiana University.
- Bairwise J. 1993, "Heterogeneous Reasoning", in Allwein and Bairwise 1993.
- Elgin C. Understanding and the Facts *Philos. Stud* 2007, 132: 33.
- Elgin C. Is Understanding Factive? in Haddock A, Millar A, Pritchard D, *Epistemic Value*, Oxford University Press.
- Kvanvig, J. *The Value of Knowledge and the Pursuit of Understanding*, Cambridge UP.
- Fisler K. 1993 A diagrammatic logic for reasoning about Asynchronous Hardware. In Allwein and Barwise 1993.
- Friedman M. 1974 Explanation and Scientific Understanding, *J. of Philosophy*, 71(1): 5–19.
- Goldenfeld and Woese 2007 *Nature* vol. 445 January 2007.
- Goodman, N. (1976). *Languages of Art*, Hackett.
- Grialou P. Longo G. Okada M. eds. 2005. *Images and Reasoning*, Keio U. P.
- Hempel 1965 *Aspects of Scientific Explanation and other Essays*. New York, Free Press.
- Holyoak K, Morrison R, 2005, eds, *The Cambridge Handbook of Thinking and Reasoning*, Cambridge U. P.
- Lipton 2004 *Inference to the best Explanation*, London, Routledge.
- Netz, R. (1999). *The Shaping of Deduction in Greek Mathematics*, Cambridge U. P.
- Regev A. and Shapiro E. 2002, "Cells as Computation" *Nature* vol 419, 26 September 2002.
- Rudwick, M. (1985). *The Great Devonian Controversy*, University of Chicago Press.
- Salmon 1984. *Scientific Explanation and the Causal structure of the World*, Princeton.
- Sloman A. Diagrams in the Mind? In Anderson M. *et al.* 2002.
- Tversky B. 2005a. Visuospatial Reasoning, in Holyoak and Morrison eds. 2005.
- Tversky B. 2005b, Some ways images Express and promote Thought, in Grialou *et al.* 2005.
- Wimsatt, W. (2007). Re-engineering philosophy for limited beings: Piecewise approximations to reality., Harvard U. P.
- Wimsatt 2002, Using False Models to Elaborate Constraints on ProcessesL Blending Inheritance in Organic and Cultural Evolution. *J. Philosophy of Science* 69, pp. S12–S24.
- Wimsatt 1987. False Models as means to Truer Theories, in M. Nitecki and A. Hoffman, *Neutral Models in Biology*. London: Oxford U. Presss, pp. 23–55.