MODELING SCIENTIFIC THEORIES. ON SOME NOTIONS OF STRUCTURALISM[1]

WERNER DIEDERICH

Often we hear people saying things like: "What you are proposing may be theoretically true, but not practically," or: "This is only a theory, it has no bearing on reality." If we press somebody saying this, he or she may admit that theories refer to reality *indirectly*, but what they are directly offering are mere *models* which only may have some resemblance with reality. Thus there is a certain pejorative use of the terms 'theory' and 'model'. I admit that theories usually have no *direct* impact on reality, but that they should have, and often indeed have, an *indirect* one, namely *via* models.

I'll argue that theories are so intimately connected with models that —in a certain sense— they even may be identified with models, or rather, classes of models. This is the basic idea of the so-called *non-statement view* of theories. To the extent to which this view is adequate the common sense is quite right in using 'theory' and 'model' interchangingly. But, I'll argue, models also must be *specified*, we have to conceive of them. Thus, although theories are no statements, but (loosely speaking) models, language comes into play nevertheless.[2] Also, of

---

[1] I am grateful for critical comments by Roberto TORRETTI on an earlier version of this paper. By 'structuralism' I refer to the approach developed by Sneed, Stegmüller, and others, cf. esp. SNEED [71], STEGMÜLLER [76], BALZER / MOULINES / SNEED [87], and BALZER / MOULINES (eds.) [96]; cf. also the bibliographies DIEDERICH / IBARRA / MORMANN [89e] and [94d].

[2] The view proposed might be wrightly called a *semantic view*. Unfortunately this term is often used in a narrower sense, excluding structuralism, e.g. by F. SUPPE, cf. his [89] and my review [94c] and article [96a]. In order to avoid misunderstandings I therefore prefer the expression *model theoretic approach* to cover, i.a., structuralism, VAN

course, a model should be modeling something, i.e., there are *claims* connected with models. Hence theories, although basically being models, have a propositional side as well; the non-statement aspect has been overemphasized.[3]

In this paper I am focusing on the smallest units of scientific theories, which structuralists call *theory-elements*.[4] Usually several theory-elements combine to form what structuralists call a *theory-net*. It is these theory-nets which resemble most closely with actual scientific theories. (There even is a still higher unit considered in structuralism: so-called *theory-holons*, consisting of several theory-nets, and likely to be of a rather interdisciplinary character.[5]) The task, then, of the following explication is to express in set theoretical or model theoretical terms, what a theory-element is.[6] (It may be more appropriate to use *category theory* instead of set theory.[7])

Models are structures satisfying certain conditions, usually called the *axioms* of a theory. Such structures typically comprise sets of higher logical types. This has become clear already in the early 5Oies when SUPPES and others have thoroughly axiomatized such elementary theories as *classical partical mechanics*. It turned out that the structures which a theory is about are of the form

(1) $< B_1, \ldots, B_l; R_1, \ldots, R_n >$

where the B_i are certain "base sets," i.e., sets of objects of various sorts,[8] and the R_j certain relations. However, the relations typically are not between elements of the B_i as such, but between elements of certain sets of higher order constructed out of these B_i: power sets and direct products

FRAASSEN's 'constructive empiricism', and Suppe's 'semantic view'. Cf. also the clarifying remarks in SCHEIBE [97], beginning of ch. II.

3 Cf. SCHEIBE, *op.cit.*, ch. II, p. 46 (cf. sec. III.1, p. 85), and my review, forthcoming in *JGenPhilosSci*.

4 In the following I often say 'theory' instead of 'theory-element', for short.

5 Cf. BALZER / MOULINES / SNEED [87], ch. VIII, sec. 1.

6 Cf. the appendix for a semi-formal explication of some key notions. HINST [96] gives a full-blown formal account.

7 Cf. MORMANN [96].

8 Usually some of the B_i are mathematical sets. Within physical theories these serve as *auxiliary base sets*. But here we need not distinguish between genuine (or "principal") and auxiliary base sets. For elucidating remarks on the interplay between physical and mathematical base sets cf. SCHEIBE, *op.cit.*, ch. II, sec. 3.

and combinations of them. In other words, the *structures* (1) are of a certain set-theoretical *type*, see **Df. 4**, below.

The typification of a theory's structures is but a first step. It doesn't say anything *specific* about such structures. This is only done by indicating a (structure) *species* Σ to which the structures should belong, see **Df. 6**. It is here where the theory's "axioms" come in to play (α_1, ..., α_s); they form the 'set-theoretic predicate' defining the species.[9] A species is something like a *concept* of higher logical order.[10] Thus we may simply say that the theory's *models* are the structures falling under this concept, i.e., belonging to the respective species (**Df. 7**). Since one and the same class of structures may be the class of models of different species, we better identify a theory not with a species but with the class of its models. We thereby abstract, so to speak, from different axiomatizations of the same theory. Hence, in a first approximation, a theory may be regarded as a *system of models* (**Df. 7**).[11]

However, there are two necessary complications. One is that a theory typically applies to structures not individually but as combined in a specific way. This trait is modelled by the notion of a *constraint,* see **Df. 5**. I'll return to this point later.[12] Thus, in a second approximation, a theory may be regarded as a pair $<\mathbf{M},C>$, i.e., as a *kernel* in the sense of **Df. 8**, where **M** is a system of models and *C* a constraint of the same type.

The other complication comes in from the distinction of *theoretical* and *non-theoretical* entities. Although this is one of the main constituents of structuralism, there still are ongoing controversies over this notion. I do not want to enter these discussions here. I just leave room for such a distinction by introducing *sub-types* (**Df. 9**) and allowing the so-called *domain of intended applications* of a kernel to be of a "smaller" type than the kernel itself (**Df. 10**). The standard structuralist account formally separates the introduction of the theoretical / non-theoretical distinction from conceiving of applications, e.g. by first introducing "theory-cores" $<\mathbf{M},C,\mathbf{M}_{pp}>$, where $\mathbf{M}_{pp}$ is a class of structures of the "smaller" type, then

[9] For a purely set-theoretical definition cf. HINST, *op. cit.*

[10] Cf. TORRETTI [90], sec. 2.8.4.

[11] Note, however, that this notion involves some axiomatization.

[12] There are several reasons for introducing constraints, first of all compatibility requirements in (the usual) case of various applications, but also the possibility of genuine second order laws, e.g. in thermodynamics (cf. BALZER / MOULINES / SNEED [87], III.5.4).

adding a domain **A** of intended applications of that type to form theory-elements.[13] Besides avoiding a certain redundancy in typification[14] my main reason for deviating from this line is that I am convinced that theoreticity is a deeply *pragmatic* notion and thus comes into play only when *applying* a kernel. Also I would like to leave open the possibility that in a full-blown theory (a "theory-net" in the structuralist jargon) one and the same kernel is applied to structures of different sub-types, i.e., a component R_j may be regarded as theoretical in some applications and as non-theoretical in others. (I suppress, in this paper, what structuralists capture by the idea of *links*, i.e., a further component, built into theory-elements, to establish certain relations to other theory-elements. Thus I am conceiving here of theory-elements as *monads*, so to speak, not reaching out to others.[15])

Although a theory-element **T**, consisting of a kernel and a domain **A**, in itself is no statement, it has a propositional side as well in that it is canonically connected with a *claim*, namely with

(2) $\mathbf{A} \in Con(\mathbf{T})$,

where $Con(\mathbf{T})$ is the "content" associated with **T** (cf. **Df. 10**). The statement (2) is sometimes called the theory element's "empirical claim." However, the cognitive status of (2) depends on how, if at all, **A** connects with reality / the phenomena, cf. below.

Note, that I have allowed for the possibility that there are no "theoretical" components at all. In this case a theory-element would just be a pair $\langle\langle \mathbf{M}, C\rangle, \mathbf{A}\rangle$ with **A** of the same type as **M**, connected with the claim '$\mathbf{A} \in Pow\mathbf{M} \cap C$'. If there are also no constraints,[16] the kernel, the theory-element, and its claim may be identified with just **M**, $\langle \mathbf{M}, \mathbf{A}\rangle$, and '$\mathbf{A} \subseteq \mathbf{M}$', respectively; i.e., we are back to a kind of model theoretical version of a pre-structuralist conception of theory.

After thus having neatly wrapped the theory's claim it is, may be, time to unwrap it a little bit again. Recall that **A** is a set of structures without the genuine theoretical components of the theory. Thus each

---

[13] I use '**A**' instead of the structuralists' 'I' in order to suggest an interpretability in VAN FRAASSEN's terms as well, cf. below.

[14] The dispensability of $\mathbf{M}_{pp}$ connects with that of $\mathbf{M}_p$, cf. below.

[15] Cf. MOULINES / POLANSKI [96].

[16] No contraints: that amounts to the same as constraints $C := PowStr\,\tau \setminus \{\emptyset\}$.

$a \in \mathbf{A}$ may represent a part, or an aspect, of reality which is described by means prior to the theory in question. In terms of VAN FRAASSEN (or as I reconstruct his ideas)[17], each $a \in \mathbf{A}$ should be isomorphic to some phenomenon φ, belonging to a realm of phenomena, $\mathbf{\Phi}$: $\varphi \in \mathbf{\Phi}$, $\varphi \approx a$. But to represent something by a set of structures, $\mathbf{A}$, is not yet *explaining* it, i.e., not yet '*saving the phenomena*'. The theory does not claim something about $\mathbf{A}$ as such, but only as a set of sub-structures[18] of models of the theory. In structuralist terms the theory claims the existence of a set of models which have the $a \in \mathbf{A}$ as substructures and which together fulfill the constraints. If, for the moment, we forget about van Fraassen's "isomorphy condition," the theory's claim, in his view, seems to be much simpler: There is a model which comprises all $a \in \mathbf{A}$ as substructures:

(3) $(\exists \mu \in \mathbf{M})(\forall a \in \mathbf{A})(a \subseteq \mu)$,

instead of the structuralist's

(4) $(\forall a \in \mathbf{A})(\exists \mu \in \mathbf{M})(a \subseteq \mu)$,
"plus constraints."

(Here "$\subseteq$" denotes "is substructure of.")

As is well known from elementary logic, (3) is a stronger claim than (4), *if* we forget about constraints. What structuralists have modelled by this notion are the *connections* between the various applications of a theory. In van Fraassen's (3) these connections are depicted by the "one model condition." Isn't that much simpler than the clumsy structuralist concept of constraints? And also more natural? After all, if we regard the models of a theory as *possible worlds,* the theory's claim should be that the *actual world* (under certain aspects) is one of these possible worlds, and (3) seems to claim just this. A short reflection, however, may show that it is van Fraassen's rendering of a theory's claim which is clumsier:

If a theory is successfully applied to a domain $\mathbf{A}(t)$ at time t and intended to be applied to a larger domain $\mathbf{A}(t') \supset \mathbf{A}(t)$ at time $t'>t$, it does not suffice to check the additional $a \in \mathbf{A}(t') \setminus \mathbf{A}(t)$ independently from the

---

[17] Cf. esp. his [80].

[18] I here use van Fraassen's term, although I suspect that what he refers to in this context is what model theorists sometimes call *reducts,* while "sub-structures" in model theory are structures of the same type as the structures whereof they are sub-structures, such as subgroups of groups.

already established applications in $\mathbf{A}(t)$. The new applications must, in van Fraassen's terms, be isomorphically embedded into *one* model that also covers $\mathbf{A}(t)$. This may be true for a certain model which would not work for a differently enlarged $\mathbf{A}^*$ (neither including, nor being included in, $\mathbf{A}(t')$ and *vice versa*. Thus all conceivable domains $\mathbf{A}$ to which actually (platonically) the theory can successfully be applied, may not be ordered chain-like, and their union may not itself be a domain for successful application. Hence, the 'empirical development' of a theory by expansion of its domain of application may not, in general, be conceived of as the exhaustion of one maximum domain of actual applicability, say $\mathbf{A}_\omega$. That a theory is applicable to $\mathbf{A}$ may thus not be expressed by the inclusion statement $\mathbf{A}\subseteq\mathbf{A}_\omega$, but only by a clumsier membership statement $\mathbf{A}\in\mathcal{A}$, where $\mathcal{A}$ is the class of all domains to which the theory actually is applicable.[19] —Also structuralism better reflects the factual plurality of science.

Those of you who are familiar with the structuralist approach certainly have wondered why I didn't mention so-called *potential models* so far. The reason simply is: I don't think they are necessary. A further reason is: I find the usual account of potential models[20] not very convincing. Also it is open to criticism like that of TORRETTI in his *Creative Understanding* (1990).[21] Torretti proposes a different account, and I am going to propose a still different one just for the case that, nevertheless, there should come up any need for the concept of potential models. The problem behind this confusing situation probably is that there are *several explicanda* for this concept. The idea with which the structuralist explication started is that

(i) the potential models specify the *conceptual frame* of a theory.

This goes along with the idea that

(ii) potential models are *non-committing/vacuous* in the sense that every structure intended for application can be extended to a potential model.

Torretti's proposal refers to a theory's development in a so-called theory-net:

---

[19] This argument is taken over from my [96a], p.19.

[20] Cf. BALZER / MOULINES / SNEED [87], ch. I.

[21] TORRETTI [90], cf. below.

(iii)Potential models of a theory-net are just the models of its basic element which serves as a *frame*.

Let's choose, as an example, the one Torretti chooses, which also is a kind of paradigmatic example of the strucuturalist approach: *classicle particle mechanics* (**CPM**[22]). In a simplified version the "axioms" for appropriate systems $<P,T,s,m,f>$ are:

(1) P is finite and non-empty,

(2) $T \in \mathrm{Int}(\mathbf{R})$,

(3) $s : P \times T \rightarrow \mathbf{R}^3$ (twice differentiable),

(4) $m : P \rightarrow \mathbf{R}^+$,

(5) $f : P \times T \rightarrow \mathbf{R}^3$,

(6) $f = m \cdot s$.

Evidently (6) is the only *genuine* law, while (1)–(5) just categorize the five components of an alleged system of classical mechanics. Only quintuples satisfying (1)–(5) we would want to put to a test whether they really are mechanical systems. In this sense (1)–(5) are only *conceptual,* while (6) is propositional (lawlike). The traditional structuralist approach therefore takes (1)–(5) as defining *potential models,* while *models* are those potential models which (also) satisfy (6). Thus it seems that the demarcation can easily be drawn. In general, however, it is not at all self-understanding which of the conditions defining models are substantial and which are not. BALZER / MOULINES / SNEED try to draw the line between conditions which, like (1)–(5), refer to just one component each, so-called *characterizations,* and those which connect several components, like (6). However, as Torretti has shown convincingly, there are severe problems connected with this idea.[23] Quite generally, I would say, this attempt dwells too heavily on the particular way the models are defined.

My main argument against this proposal relates to the explicandum (ii). Characterizations like (1)–(5) may still be too restrictive to fulfill (ii). We might for instance want to include among tentative applications more bizarre motions than (3) allows. In general, the domain **A** chosen for ap-

---

22 *Op. cit.*, 3.4; cf. BALZER / MOULINES / SNEED [87], III.3.

23 *Op. cit.*, 3.3.

plication should possibly be any set of systems of an appropriate subtype ρ (i.e., any subset of $Str(\rho)$).

Hence, the "projection" of the class of potential models, $\mathbf{M}_p$, should cover the whole class of structures of type ρ:

(I) $r(\mathbf{M}_p) \supseteq Str(\rho)$.

Now, if we do not want to exclude the limiting case, $r{=}id$, i.e., the case where there are no theoretical components at all, we have to choose the widest possible candidate for $\mathbf{M}_p$, i.e., $Str(\tau)$ itself. Thus my radical proposal is

(II) $\mathbf{M}_p := Str(\tau)$.

It is evident that this proposal does not have to enter the concepts of a kernel: it doesn't at all refer to a particular theory or species.[24] And it is compatable with Torretti's proposal insofar typical basic elements (frames) may be vacuous, like **CPM**, in that they do not impose any effective restrictions: if $r(\mathbf{M}) = Str(\rho)$, then, of course, $\mathbf{M}_p := \mathbf{M}$ meets requirement (I).

My proposal (II) also matches my deviation in defining *constraints*. Usually constraints are regarded as subsets of $Pow(\mathbf{M}_p)$, i.e., are defined with respect to a structure species.[25] In accordance with (II) I prefer the purely set-theoretical **Df. 5**. Constraints, by the way, may be an obstacle to Torretti's solution, if they are not "wide enough": they should not narrow down the frame and therefore should be taken in my more general sense. Also, if potential models are to specify the conceptual frame of a theory, one wonders why they are not defined with respect to constraints as well.[26] Hence, the usual account doesn't convincingly fulfill idea (i).

---

[24] Cf. HINST [96], Def. 3.23.

[25] Cf., e.g., BALZER / MOULINES / SNEED [87], II.2.3. —HINST [96], p. 252, does, in effect, define constraints as I do.

[26] In the spirit of defining potential models by singling out the "characterizing" clauses within the definition of models, one would want to put the (restrictions of the second level) "characterizing" conditions within the definition of constraints into the definition of potential models as well, i.e., constraints, like models, are methodologically prior to potential models.

Semi-Formal Appendix

The following "definitions" are highly abstract and sketchy. They are not intended to sufficiently characterize the respective concepts. Esp. the last one, **Df. 10**, is best to be regarded as being *partial*, i.e., giving only *necessary* conditions for something to be a *theory-element* or its *claim*. Otherwise all kinds of weird structures would count as "models," "constraints," "theory-elements," and so on. But relevant pieces of real-life theories should be *reconstructable* along the indicated lines.

Df. 1: If l is a positive integer, *l-echelons* are recursively defined schemes:

(i) each numeral i, $1 \leq i \leq l$, is a l-echelon,

(ii) if σ is a l-echelon, so is $>Pou\sigma<$,

(iii) if σ_1 and σ_2 are l-echelons, so is $>(\sigma_1 \times \sigma_2)<$.

(Take '>' and '<' as Quinean "corner quotes.")

Df. 2: If σ is a l-echelon and $B_1, \ldots, B_l$ are sets, the *echelon set* $\sigma(B_1, \ldots, B_l)$ is defined correspondingly.

Df. 3: If l and n are positive integers and $\sigma_1, \ldots, \sigma_n$ l-echelons, $\tau := <l, \sigma_1, \ldots, \sigma_n>$ is a *type* (of *order* $\mathrm{ord}(\tau) := n$).

Df. 4: If $\tau = <l, \sigma_1, \ldots, \sigma_n>$ is a type and $B_1, \ldots, B_l, R_1, \ldots, R_n$ are sets, $<B_1, \ldots, B_l, R, \ldots, R_n>$ is a *structure* of type τ

iff $R_i \in \sigma_i(B_1, \ldots, B_l)$ for all i, $1 \leq i \leq n$.

$Str(\tau)$ is the class of structures of type τ.

Df. 5: If τ is a type and $C \subseteq Pow\, Str(\tau)$, C is a *constraint* of type τ

iff (i) C is non-empty,

(ii) $\varnothing \notin C$,

(iii) $\{x\} \in C$ for all $x \in Str(\tau)$.

Probably all "real-life" theories' constraints are also "transitive," i.e., obey also the condition

(iv) If $X \subseteq Y \in C$, then $X \in C$

(*cf.* TORRETTI [90], p. 301, n. 27). —There might be a problem in case $Str(\tau)$ is a proper class (cf. Hinst [96], pp. 252 f.).

Df. 6: If τ is a type and $\alpha_1,\ldots,\alpha_s$, for some positive integer s, are set-theoretical formulas "applicable" to structures of type τ, $<\tau,\alpha_1,\ldots,\alpha_s>$ is a (structure) *species* of type τ. (A less sloppy account should, of course, include invariant conditions; cf. e.g. HINST [96], p. 244ff, or SCHEIBE [97], sec. II.3, pp. 65f.)

Df. 7: If $\Sigma = <\tau,\alpha_1,\ldots,\alpha_s>$ is a species and $y \in Str(\tau)$, y is a *model* of Σ iff all α_j, $1 \leq j \leq s$, "apply" to y. $Mod(\Sigma)$ is the class of models of Σ. **M** is a *system of models* iff there is a species Σ such that $\mathbf{M} = Mod(\Sigma)$.

Df. 8: If **M** is a system of models and C a constraint of the same type, $\mathcal{K}=<\mathbf{M},C>$ is a *kernel* (of that type).

Df. 9: If $\tau = <l,\sigma_1,\ldots,\sigma_n>$ is a type, p a positive integer $\leq n$, and ρ is obtained from τ by deleting all but p of the σ_i, ρ is a *sub-type* of τ (of order p). A pair $<\tau,\rho>$ thus induces a function $r : Str(\tau) \rightarrow \mathrm{Str}(\rho)$, which deletes the respective relations in structures; 'r' is also to denote higher level functions corresponding to r.

Df. 10: If $\mathcal{K}= <\mathbf{M},C>$ is a kernel, its *content* is $Con\ K := Pow\ \mathbf{M} \cap C$. If K is of type τ, ρ a sub-type of τ, and

$\mathbf{A} \subseteq Str(\rho)$ a set (representing a) "domain of intended applications," then $\mathbf{T} := < \mathcal{K},\mathbf{A} >$ is a *theory-element* with *content* $Con\mathbf{T} := r\ Con\mathcal{K}$ and *claim* '$\mathbf{A} \in Con\mathbf{T}$', where r is the "projecting" function induced by $<\tau,\rho>$. (BALZER / MOULINES / SNEED [87] would call such a **T** an "idealized theory-element" (without links), because it doesn't comprise a component for approximations; cf. their definition D II-17, p. 89, entering D VII-8, p. 352.)

University of Hamburg, Germany

References

BALZER, Wolfgang / MOULINES, C. Ulises (eds.) [96]: *Structuralist Theory of Science. Focal Issues, New Results*. Berlin etc.: de Gruyter 1996.

BALZER, Wolfgang / MOULINES, C. Ulises / SNEED, Joseph D. [87]: *An Architectonic for Science. The Structuralist Program*. Dordrecht etc.: Reidel 1987.

DIEDERICH, Werner / IBARRA, Andoni / MORMANN, Thomas [89e]: 'Bibliography of Structuralism', *Erkenntnis* ***30*** (1989), 387–407.

DIEDERICH, Werner / IBARRA, Andoni / MORMANN, Thomas [94d]: 'Bibliography of Structuralism II (1989–1994 and Additions)', *Erkenntnis* **41** (1994), 403–418.

DIEDERICH, Werner [94c]: Review of SUPPE [89], *Erkenntnis* **41** (1994), 421–426.

DIEDERICH, Werner [96a]: 'Structuralism as Developed Within the Model-Theoretical Approach in the Philosophy of Science', in BALZER / MOULINES (eds.) [96], 15–21.

HINST, Peter [96]: 'A Rigorous Set Theoretical Foundation of the Structuralist Approach', in: BALZER / MOULINES (eds.) [96], 233–263.

MORMANN, Thomas [96]: 'Categorical Structuralism', in: BLAZER / MOULINES (eds.) [96], 265–286.

MOULINES, C. Ulises / POLANSKI, Marek [96]: 'Bridges, Constraints, and Links', in: BALZER / MOULINES (eds.) [96], 219–232.

SCHEIBE, Erhard [97]: *Die Reduktion physikalischer Theorien. Ein Beitrag zur Einheit der Physik. Teil I: Grundlagen und elementare Theorie*. Berlin etc.: Springer 1997. (English translation forthcoming)

SNEED, Joseph D. [71]: *The Logical Structure of Mathematical Physics*. Dordrecht etc.: Reidel 1971, 2nd ed., 1979.

STEGMÜLLER, Wolfgang [76]: The Structure and Dynamics of Theories. New York etc.: Springer 1976.

SUPPE, Frederick [89]: The Semantic Conception of Theories and Scientific Realism. Urbana etc.: University of Illinois Press, 1989.

TORRETTI, Roberto [90]: *Creative Understanding. Philosophical Reflections on Physics*. Chicago etc.: University of Chicago Press 1990.

VAN FRAASSEN, Bas C. [80]: *The Scientific Image*. Oxford: Oxford University Press, 1980.