HANS ROTT

REDUCTION: SOME CRITERIA AND CRITICISMS OF THE STRUCTURALIST CONCEPT*

PROLOGUE

Once upon a time there was a concept of reduction that was of a miraculous clarity. It was advocated by Nagel (1949, 1961) and Hempel (1965, 1966) and it said that a theory T' is reducible to another theory T if and only if there are conditions of application and "bridge principles" collected in a set of sentences A such that

(d) $T, A \vdash T'$.

Since the bridge principles sometimes were called *d*efinitions (of T'-terms in the language of T) and since "$\vdash$" of course signifies the relation of *d*erivability between sets of sentences, we shall call this idea the *d-concept*. Deplorably, those happy days are gone. The party most guilty of this sad state of affairs is Feyerabend, who made out two more dogmas of empiricism presupposed by the d-concept of reduction, viz., the *consistency condition* and the *condition of meaning invariance* (cf. Feyerabend 1962, pp. 43f and 1963, p. 10). The bloom of structuralism in the philosophy of science during the last 15 years may be regarded as a reaction to the challenge of the second of these conditions questioning the comparability or commensurability of scientific theories. The problem of shifts in the meaning of scientific expressions was circumvented by renouncing the explicit description of the languages of scientific theories.

The structuralist turn caused by Sneed was very effective in so far as the notion of reduction has been discussed – at least in West Germany – almost exclusively in a structuralist setting since the mid-seventies. The gap to the traditional statement view in the philosophy of science, however, is easily bridged. Adams (1959), the pioneer of the structuralist concept of reduction, explicitly wants to satisfy the adequacy conditions of the d-concept, and the translatability of statement and non-statement views goes much further than has sometimes been supposed.[1] The following, therefore, may also interest a convinced non-structuralist. Let me emphasize in advance that the present paper

Erkenntnis **27** (1987) 231–256

is restricted to the problems of strict reduction and leaves aside those of approximations.

The structuralist view of the dynamics of scientific theories has been subject to criticism in quite regular intervals. It seems to me, that Mayr (1976), Tuomela (1978), Niiniluoto (1980), Pearce (1982), Hoering (1984), Kamlah (1985), and Mormann (1984) are of special significance. The presentations of the structuralists as well as of their critics both have the drawback that practically every essay introduces new notations and minor changes of the definitions. Thus the reader – if not the author – runs the risk of losing sight of the continuity of discussion. In the following I shall try to show how the scattered criteria and criticisms of the structuralist concept of reduction are to be evaluated in their context by using a uniform notation and simplified definitions. The simplifying assumptions are essentially intended to avoid the intricate devices introduced by Sneed for his interpretation of theoretical terms. Even though the supplementation of these devices to the following definitions poses no essential problems it would make the immediate understanding more difficult. I will – like e.g. Kamlah – go back to the old Adamsian proposal and – neglecting theoretical terms and constraints – take a theory to be an ordered couple (M, I), where M is the class of models of T and I the class of its intended applications. M and I are always meant to be nonempty classes of structures of the same similarity type τ, and the class M_p of potential models of T is simply to be identified with the class of all structures of type τ (modulo a suitable logic L).[2]

Now let $T = (M, I)$ and $T' = (M', I')$ be two theories. The essence of the official structuralist definition of "T' is reducible to T" is the existence of a so-called *reduction-relation*, which most expediently is given as a function

(s) $\mathscr{F}: M_p^0 \to M'_p$, where $M_p^0 \subseteq M_p$.

In general, it is demanded that $\mathscr{F}$ is *onto* and that M_p^0 is a *proper* subset of M_p;[3] it will be indicated in the sequel where these demands are employed. Representing reduction relations as functions allows us to define images under $\mathscr{F}$

$$\mathscr{F}[X] := \{x' \in M'_p : \exists x \in M_p^0 \cap X\ (x' = \mathscr{F}(x))\}$$

for *every* $X \subseteq M_p$ (i.e., $X \subseteq M_p^0$ is not required) and inverse images under $\mathscr{F}$

$$\mathscr{F}^{-1}[X'] := \{x \in M_p^0: \mathscr{F}(x) \in X'\}$$

for every $X' \subseteq M'_p$, and $\mathscr{F}^{-1}(x') := \mathscr{F}^{-1}[\{x'\}]$ for every $x' \in M'_p$. Classes like these can be used very comfortably to formulate suggestive conditions.

Before turning to more detailed considerations, I would like to anticipate a fundamental objection to the structuralist idea. When Adams and Sneed speak of the existence of a function $\mathscr{F}$ they do not mean the mere mathematical sense. Such a function has to meet the intuitive requirement that $\mathscr{F}$ assigns to every potential model of the reducing theory T a potential model of the reduced theory T' which is – considered from an external point of view – "identical with" or "composed of the individuals of" the former model. Ascertaining such an identity or composition is an empirical matter and afflicted with problems that seem to be beyond the reach of a purely formal analysis. Accordingly, Adams' and Sneed's formal criteria – just as those discussed below – are meant only as necessary conditions for the presence of an intuitive reduction of theories.[4] Speaking of the existence of a function $\mathscr{F}$ in the sequel, I tacitly assume that this function fulfils the intuitive condition just mentioned.[5] Thus I want to avoid debating with all those critics who regard the structuralist idea as much too weak and therefore as unacceptable from the outset, because the existence of a function as indicated in (s) (with properties of the kind discussed below) can concern only questions of cardinality – and that is certainly insufficient by far for a genuine reduction.[6]

Now the way is paved for a more exact examination of the criteria to be satisfied by reductions and their use in the structuralist approach. For this purpose the papers of Tuomela, Niiniluoto, and Hoering that argue on a more general level can be kept in the background, which leaves us with the essays of Mayr, Pearce, Kamlah, and Mormann. These essays are interesting and valuable, but for the most part quite complicated, and their presentation obviously mirrors the idiosyncratic preferences of the authors. I would like to present their various efforts as an integral whole and leave the dramaturgical structure of the criticism unchanged, despite the small loss in systematic stringency. The main point is that it is a true, coherent story, told in a simple language and thus easy to understand in spite of the considerable aberrations and confusions to be overcome. Then the moral of the matter will take care of itself.

FIRST ACT: EXPOSITION

In his trail-blazing article, Adams (1959, pp. 256, 260f) names exactly the adequacy conditions of the d-concept of reduction. T' is reducible to T only if the following conditions hold:

(C1) The fundamental concepts of T' are definable[7] in terms of the fundamental concepts of T;

(C2) the fundamental laws of T' are derivable[8] from the fundamental laws of T together with the definitions mentioned in (C1).

As Adams emphasizes with reference to Nagel, it is important to note that (C1) has the character of an empirical hypothesis, whereas (C2) is a contention which may be decided a priori. This dividing of reduction into an "applied" and a "formal" aspect can be recognized in the structuralist "analogues" introduced by Adams. They can be expressed very conveniently by using a suitable function $\mathcal{F}$ according to (s):

($C1^s$) $\mathcal{F}[I] \supseteq I'$;

($C2^s$) $\mathcal{F}[M] \subseteq M'$.

($C1^s$) says that for every intended application of the reduced theory there exists a (i.e., at least one) "corresponding" intended application of the reducing theory. I cannot agree with Adams' (1959, pp. 260f) and Sneed's (1971, p. 217) suggestion that ($C1^s$) is an exact counterpart of (C1). While (C1) presupposes the possibility of a precise indication as to *how* T-applications are constructed out of T'-applications, ($C1^s$) simply implies *that* corresponding T-applications can be found. Again, in (C1) there is no restriction to *intended* applications. Therefore, ($C1^s$) is weaker, and one might expect it to be called in question in the following. ($C2^s$), stating that T-models may only be transformed into T'-models, however, seems to be a very faithful image of the basic criterion (C2): the syntactical relation of derivability is mirrored by the inclusion relation on the model level, where $\mathcal{F}$ takes the rôle of the "mediating" definitions between T' and T, and its domain M_p^0 characterizes the range of application of T'. Surprisingly it is not ($C1^s$) but ($C2^s$) which draws the fire of heavy criticism.[9]

There are other criteria which apparently played a rôle in the motivation of the structuralist concept of reduction marked by (s), ($C1^s$), and ($C2^s$). Adams (1959, p. 261) considers the following (necessary) condition to be the most important for the reduction of T' to T:

(C3) If T is true ("correct"), then T' is true ("correct") as well.

To judge (C3) we must know what it means for a theory to be true. Here is Adams' (1959, p. 260) plausible definition:

(D1) A theory $T=(M, I)$ is true (correct) iff $I \subseteq M$ (or equivalently, iff $I \backslash M = \emptyset$).

Hence, in structuralist wording, (C3) is expressed as follows:

($C3^s$) $I \subseteq M \Rightarrow I' \subseteq M'$ (or equivalently, $I \backslash M = \emptyset \Rightarrow I' \backslash M' = \emptyset$).

Clearly, ($C3^s$) is guaranteed by ($C1^s$) and ($C2^2$). For from $I \subseteq M$ it follows that $\mathscr{F}[I] \subseteq \mathscr{F}[M]$ and hence, by ($C1^s$) and ($C2^s$), $I' \subseteq \mathscr{F}[I] \subseteq \mathscr{F}[M] \subseteq M'$. Thus (C3) supports Adams' suggestion.

Another criterion is to be found in Sneed (1971, p. 218) and Stegmüller (1973, p. 143):

(C4) Everything that can be explained ("systematized") by T' can be explained (systematized) by T as well.

For the application of this requirement an adequate concept of explanation (more generally, of systematization) would have to be provided.[10] It is not pointed out, however, how this might be fitted into the structuralist approach. Sneed and Stegmüller must be held responsible for having advanced a noteworthy criterion which they do not know what to make of in this context. At this point, there is no ($C4^s$). We shall have to wait until the fifth act.

Finally, Sneed (1976, p. 139) gives the so-called "preservation property" as an additional desideratum. When T' is reducible to T, the following condition is said to hold:

(C5) For every specialization T'_1 of T', there is a specialization T_1 of T such that T'_1 is reducible to T_1 (analogously as T' is reducible to T).

First of all, we recall what a specialization is:

(D2) $T_1=(M_1, I_1)$ is a *specialization* of $T=(M, I)$ ($T_1 \leqslant T$) iff $M_1 \subseteq M$ and $I_1 \subseteq I$.

(D2) is to express that a specialization has more laws and less intended applications than the more general theory.

An obscurity in (C5) remains, which is due to the fact that the idea of this criterion is already made to order to the structuralist instruments: "analogously" simply means "by the same reduction function $\mathscr{F}$". Now (C5) may be translated in this way ($\mathscr{F}$ be the usual reduction function for T' and T):

$$(\mathrm{C5^s}) \qquad \forall T'_1 \leq T' \exists T_1 \leq T\ (\mathscr{F}[I_1] \supseteq I'_1\ \&\ \mathscr{F}[M_1] \subseteq M'_1).$$

Sneed states, and Balzer and Sneed (1978, pp. 188–192) prove, the theorem that their reduction concept – a version of (s), (C1^s), and (C2^s) complicated by the use of theoretical terms and constraints – satisfies (C5^s). That this must be considered as an error is one of the main arguments of Mayr's criticism.

SECOND ACT: MAYR

Thanks to our simplifying assumptions we are able to give a counterexample to (C5^s) far more easily than Mayr (1976, p. 285). Let T and T' be theories and $\mathscr{F}$ a reduction function such that $\mathscr{F}[M] \subsetneq M'$ and $\mathscr{F}[I] = I'$, and let $T'_1 = (M'_1, I'_1) = (M' \backslash \mathscr{F}[M], I')$. Obviously $T'_1 \leq T'$, but there is no *non-empty* $M_1 \subseteq M$ with $\mathscr{F}[M_1] \subseteq M'_1$, since $\mathscr{F}[M_1] \subseteq \mathscr{F}[M]$ and $\mathscr{F}[M] \cap M'_1 = \emptyset$. Therefore, there is no suitable T_1 for (C5^s). Indeed, the M_1 in the proof of Balzer and Sneed, which is defined as $M \cap \mathscr{F}^{-1}[M'_1]$, is empty in this example. Unfortunately, the authors do not seem to have noticed that Mayr essentially had shown that in certain cases every M_1 appropriate for (C5^s) *must* be empty. Still, it cannot be desired to save (C5^s) trivially by employing inconsistent "reducing" specializations T_1. For that reason I have presupposed M to be non-empty in every theory T; moreover, (C5^s) should be amended by the additional requirement $\mathscr{F}[M_1] \neq \emptyset$ (the signature "(C5^s)" will refer to this improved form in the sequel). Accordingly, (C5) doesn't support the explication of reduction by (s), (C1^s), and (C2^s), which seemed to be so satisfactory. On the contrary, it scores a heavy blow against it.

A second important intuitive argument that Mayr puts forward is this:[11]

(C6) If T is a specialization of T' or if T' is a specialization of T, T' and T do not stand in a reduction relation.

Even though this form of the criterion will be used, for the sake of uniformity we should express the criterion as a necessary condition for theories in a reduction relation:

(C6^s) $T \nleq T'$ & $T' \nleq T$.

(C6) also presents some difficulties, but not exactly for the concept of reduction characterized by (C1^s) and (C2^s). Nevertheless its discussion is instructive. Suppose that T is a specialization of T'. If we would consider nothing but the reduction of the theory cores (which are identical with the model classes in our simplified theory concept) given by (s) and (C2^s) alone, then T' would be reducible to T: simply choose $\mathscr{F} = \mathrm{id}|_{M_p}$ (the identity mapping on $M_p^0 = M_p$), and evidently $\mathscr{F}[M] = M \subseteq M'$. Conversely, let T' be a specialization of T; then, according to the *initial* reduction concept of Sneed and Stegmüller (cf. note 9), T' is reducible to T by $\mathrm{id}|_{M_p}$ as well. In this respect at least Adams' old version remains immune to Mayr's criticism; it is interesting to check why. While the first case ($T \leq T'$) is disposed of trivially by the help of (C1^s), in the second case ($T' \leq T$) you might suppose that a suitable restriction of $\mathrm{id}|_{M_p}$ could establish a reduction. But neither $\mathrm{id}|_{I'}$ nor $\mathrm{id}|_{M' \cup \mathscr{C}M}$ (where $\mathscr{C}M$ is the class complement of M), which are tailored to (C1^s) and (C2^s) respectively, optimizing the chances for the other criterion, can guarantee the simultaneous satisfaction of both criteria (without even mentioning that either M'_p couldn't be defined as the class of all structures of the same type as the T'-models or one would have to give up the requirement that $\mathscr{F}$ is onto). A little sketch[12] shows that this is due to the fact that $I' \setminus (M' \cup \mathscr{C}M)$ need not be empty.[13] It is advisable to take into account not only intended applications and models separately, but also combinations of these classes. For that purpose the following partition of the class of intended applications is useful:

(D3) Let $T = (M, I)$ be a theory. The elements of $I \cap M$ are called *successful applications of* T, and elements of $I \setminus M$ are called *anomalies of* T.

Using these terms, the situation can be formulated in this way: (C6) constitutes a problem for (C1^s) and (C2^s) if T' is a specialization of T and no successful application of T is at the same time an anomaly of T'. As this is not an implausible condition for specializations, one has to be aware that (C6) can become dangerous to the Adamsian proposal as well. This seems to have escaped Mayr's notice (but cf. (C7) below).

Before turning to the third and, as far as I can see, last main argument of Mayr's criticism, two remarks are in order. First, (C6) shows how important it is to understand clearly the ambiguous rôle of the domain M_p^0 of $\mathscr{F}$. On the one hand, it undertakes the restricting of the wider range of application of the reducing theory T, corresponding to the (non-lawlike) description of the initial conditions in the statement view. On the other, M_p^0 also allows smuggling additional *laws* over and above those of T into T' so that the "reduced" theory can be much stronger than the "reducing" theory – undoubtedly a counter-intuitive consequence. One must look for rules concerning the specification of M_p^0. Second, (C6) demonstrates that it is indispensable to think about the relation between specialization – which is sometimes regarded as typical for the change within normal science – and reduction – which is regarded as characteristic of scientific revolutions.

Finally, Mayr's last desideratum for reductions is

(C7) The reducing theory can resolve ("explain") anomalies of the reduced theory.

According to Mayr, this means that there should be successful applications of T corresponding to anomalies of T'. Applying (D3) we transcribe (C7) into

(C7ˢ) $\mathscr{F}[I \cap M] \cap (I' \backslash M') \neq \emptyset$.

But this is made altogether impossible by the concept of reduction in question. For, according to (C2ˢ), $\mathscr{F}[M] \cap \mathscr{C}M' = \emptyset$ holds, and since $\mathscr{F}[I \cap M] \subseteq \mathscr{F}[M]$ and $I' \backslash M' \subseteq \mathscr{C}M'$, a fortiori $\mathscr{F}[I \cap M] \cap I' \backslash M' = \emptyset$.

Mayr does not confine himself to destructive criticism, but offers, above all to warrant (C5) and (C7), a constructive counter-proposal. His alternative to (C1ˢ) and (C2ˢ) consists of the following two requirements (Mayr 1976, p. 289):

(C1ˢ′) $\mathscr{F}[I] = I'$,

(C2ˢ′) $\mathscr{F}[M] \supseteq M'$.

I'm not sure whether Mayr has noticed that his attempt to replace the Sneedian many-many relation $\bar{\rho}$ by his many-one relation $\tilde{\rho}$ has the result that ($C1^s$) is intuitively strengthened to ($C1^{s\prime}$). It should be noted that only this change blocks a violation of Mayr's own criterion (C6), since ($C1^s$) and ($C2^{s\prime}$) allow every specialization T' of T to be reduced to T by means of $\mathrm{id}|_{M_p}$.[14] In fact, Mayr's ($C2^{s\prime}$) turns out to be equivalent to the basic idea ($C2^{s*}$) of the original Sneed–Stegmüller approach (cf. note 9) with which Mayr himself found fault. Next, pleading for ($C1^{s\prime}$) and ($C2^{s\prime}$), one should correct ($C5^s$) in such a way that the relation between T_1 and T_1' is governed by ($C1^{s\prime}$) and ($C2^{s\prime}$) as well. Fortunately this new transcription of (C5) is just as guaranteed by ($C1^{s\prime}$) and ($C2^{s\prime}$) (take $M_1 := M$ and $I_1 := \mathscr{F}^{-1}[I_1'] \cap I$) as the old ($C5^s$) is (take $M_1 := \mathscr{F}^{-1}[M_1'] \cap M$ and $I_1 := I$). At last, the central condition (C2) is no longer ensured by ($C2^{s\prime}$), but – on the contrary – is necessarily violated if (C7) is to have an effect. Mayr failed to see this, for he maintained the contrary (1976, pp. 276, 286f).

THIRD ACT: PEARCE

At this juncture it cannot be decided which proposal should be preferred: that of the first or that of the second act.[15] Pearce's (1982) discussion sheds more light on the situation. We need not give an account of his decidedly pro-linguistic tendencies, nor of his strong logical apparatus, but can be content with considerations on the model level. Pearce (1982, p. 308) chooses a formulation of Stegmüller's (1973, p. 146; English 1976, p. 128) as his starting point. If T' is reducible to T it should hold that

(C8) For every sentence ϕ of T, if ψ is the corresponding sentence of T', then ϕ is true only if ψ is true.

Stegmüller believes (C8) to be an application oriented paraphrase of (C2). The phrase "sentence of T" – suspiciously sounding like the statement view – presents no problems; all what is needed in the following is that for any sentence ϕ of T, an extension $\|\phi\| \subseteq M_p$ (i.e., the class of potential models of T in which ϕ holds) can be assigned. To be able to judge (C8) completely, we still need some precise structuralist concepts. Pearce (1982, p. 323) represents with very good reasons in this context, I think, the truth of a sentence as theory-dependent, and he does this in two ways:

(D4) A sentence ϕ of a theory $T = (M, I)$ is called
(a) *true in* M iff $M \subseteq \|\phi\|$,
(b) *true in* T iff $M \cap I \subseteq \|\phi\|$.

The interpretation of "the corresponding sentence" in (C8), apparently presupposing the idea of translatability (and, therefore, commensurability?) of the theories in question, is a little more delicate. But if we again confine ourselves to the model level, (C8) seems to fit the present frame very well:

($C8^s$) For every sentence ϕ of T,
(a) $M \subseteq \|\phi\| \Rightarrow M' \subseteq \mathcal{F}[\|\phi\|]$,
(b) $M \cap I \subseteq \|\phi\| \Rightarrow M' \cap I' \subseteq \mathcal{F}[\|\phi\|]$.

Here the only demand on "the corresponding sentence" ψ mentioned in (C8) is that $\|\psi\| = \mathcal{F}[\|\phi\|]$. Testing this, we could define a "translation of T into T' (relative to $\mathcal{F}$)" abstractly as a function from T-sentences into T'-sentences assigning to every ϕ a ψ with the property $\|\psi\| = \mathcal{F}[\|\phi\|]$. Why we do not do this, will become clear below.

A confirmed proponent of structuralism will wish to get rid of the linguistic remains in ($C8^s$). For this purpose, he needn't even suppose that every "proposition" of a theory (i.e., every subclass of M_p) is expressible linguistically. The precondition that the class of models, and respectively the class of successful applications, of T is definable in T is sufficient to show that (C8) can be reformulated without any reference to languages. For ($C8^s$) is then equivalent to

($C8^{s\prime}$) (a) $\mathcal{F}[M] \supseteq M'$,
(b) $\mathcal{F}[M \cap I] \supseteq M' \cap I'$.

To get ($C8^{s\prime}$) (a) from ($C8^s$)(a), simply take a sentence ϕ defining the class of T-models (i.e., $\|\phi\| = M$) in ($C8^s$)(a). For the other direction, let ϕ be such that $M \subseteq \|\phi\|$; hence, $\mathcal{F}[M] \subseteq \mathcal{F}[\|\phi\|]$ and, by ($C8^{s\prime}$)(a), also $M' \subseteq \mathcal{F}[\|\phi\|]$. Case (b) is of course analogous to (a).

In this way, (C8) would yield two conditions: the Adams-Sneedian ($C2^s$) is inverted to ($C2^{s\prime}$), which already was preferred by Mayr, and in addition a corresponding condition regarding the successful application emerges: for every successful application of the reduced theory at least one successful application of the reducing theory corresponds (via the reduction function $\mathcal{F}$).

Unfortunately this result cannot be maintained in its whole simplicity. As the reducing theory always is supposed to have a greater expressive power and to allow linguistic distinctions which are not reproducible in the reduced theory, the foregoing idea cannot be an intuitively adequate explication of translation. "The corresponding sentence" ψ, presupposed by (C8), does not exist at all. Consequently, the direction of the translation as defined by Pearce (1982, p. 314) is exactly the other way round. Thus we define:

(D5) Let the theory $T' = (M', I')$ be reduced to the theory $T = (M, I)$ by a suitable function $\mathcal{F}$.

(a) A sentence ϕ of T is called a *translation of the sentence ψ of T' (relative to $\mathcal{F}$)* iff $\|\phi\| = \mathcal{F}^{-1}[\|\psi\|]$.

(b) A function $\mathcal{T}$ of T'-sentences into T-sentences is called a *translation of T' into T (relative to $\mathcal{F}$)* iff for every T'-sentence ψ $\mathcal{T}(\psi)$ is a translation of ψ (relative to $\mathcal{F}$).[16]

The following is a plausible consequence of (D5): if a translation $\mathcal{T}$ of T' into T exists, then all potential models x and y of T that are mapped onto the same potential model of T' cannot be distinguished by translations of T'-sentences, i.e., $\mathcal{F}(x) = \mathcal{F}(y)$ implies $x \in \|\mathcal{T}(\psi)\| \Leftrightarrow y \in \|\mathcal{T}(\psi)\|$ for every sentence ψ of T'.

This change brings it about that ($C8^s$) is replaced by

($C8^{s''}$) For every sentence ψ of T',

(a) $M \subseteq \mathcal{F}^{-1}[\|\psi\|] \Rightarrow M' \subseteq \|\psi\|$,

(b) $M \cap I \subseteq \mathcal{F}^{-1}[\|\psi\|] \Rightarrow M' \cap I' \subseteq \|\psi\|$.

The linguistically inspired conditions ($C8^{s''}$) are no longer equivalent to the simple inclusions ($C8^{s'}$). ($C8^{s'}$) gives only sufficient, but not necessary conditions for ($C8^{s''}$). Let us consider case (a) once more. On the one hand, ($C8^{s'}$) is sufficient: from $M \subseteq \mathcal{F}^{-1}[\|\psi\|]$ it follows that $\mathcal{F}[M] \subseteq \mathcal{F}[\mathcal{F}^{-1}[\|\psi\|]] \subseteq \|\psi\|$ and, by ($C8^{s'}$), $M' \subseteq \|\psi\|$ follows. On the other, ($C8^{s'}$) would only be necessary on making two implausible assumptions. Assume firstly that $\mathcal{F}[M]$ can be defined in T', i.e., that there is a T'-sentence ψ with $\|\psi\| = \mathcal{F}[M]$, then the consequent of ($C8^{s''}$) is identical with ($C8^{s'}$); but the antecedent of ($C8^{s''}$), viz., $M \subseteq \mathcal{F}^{-1}[\mathcal{F}[M]]$, would be valid only if we assumed secondly that $M \subseteq M_p^0$ (note that ($C8^{s''}$) is satisfied trivially if $M \not\subseteq M_p^0$!). But since the first assumption is rather strong and the second even is undesirable, one mustn't consider ($C8^{s'}$) and ($C8^{s''}$) to be equivalent.

Nevertheless, Mayr's condition ($C2^{s'}$) is supported by Pearce in a completely independent way. Whichever criterion is favoured – ($C2^{s}$) or ($C2^{s'}$) – it seems that Stegmüller's intuitions concerning the adequacy condition (C8) for reductions have to be corrected twice: first, (C8) is not formulated accurately enough – the translation must go in the opposite direction. And second, as pointed out by Pearce (1982, p. 328), (C8) is far from being equivalent to (C2).

It finally remains to state that Pearce (1982, p. 236) regards Mayr's ($C1^{s'}$) as too strong and upholds ($C1^{s}$) again. He adds ($C8^{s'}$)(b) as a third condition which derives from (C8) just as ($C2^{s'}$) does. ($C8^{s'}$)(b) does not follow from the other conditions, which say that a successful application x' of T' must have a corresponding intended application $x_1 \in I$ and a corresponding model $x_2 \in M$, but $x_1 = x_2$ need not hold. It should be noted, however, that ($C8^{s'}$)(b) also isn't able to prevent that every specialization T' of T can be trivially reduced to its general theory T by $\mathrm{id}|_{M_p}$ if Pearce's proposal is adopted. As the more informative, hence better theory is reducible to the weaker one then, (C6) is violated in its critical direction.

FOURTH ACT: KAMLAH

Kamlah's essay also can be made to suit the present frame, without any individual characteristics and without his treatment of approximations. Starting out from a quotation of Hempel (1965, p. 344), which says that the reducing theory implies the laws of the reduced theory only within a limited range,[17] Kamlah's argumentation is to the effect that not the laws (as in (C2)) but the empirical claims of the theories are put into a consequence relation. Thus, if T' is reducible to T, it should hold that

(C9) The empirical claim of T' follows from the empirical claim of T.

In our condensed theory concept, the empirical claim of a given theory T is simply equated with the truth condition for T mentioned in (D1), viz., $I \subseteq M$. On the model level this meta-theoretical requirement is reflected by the "proposition" of the potential models "allowed" by T, i.e., $\mathscr{C}I \cup M$. Accordingly, (C9) can be made precise in the following structuralist way (cf. Kamlah's (C1$^-$) and (I) in 1985, pp. 135f):

(C9^s) $\quad \mathcal{F}[\mathcal{C}I \cup M] \subseteq \mathcal{C}I' \cup M'$.

First of all, the connection with the related criterion (C3) is interesting: if we presuppose that $\mathcal{F}$ is onto (recall that this is a usual stipulation), (C9^s) implies (C3^s): from $I \backslash M = \emptyset$ we get, by (C9^s), $\mathcal{F}[M_p] \subseteq \mathcal{C}I' \cup M'$, and therefore, since $\mathcal{F}$ is onto, $I' \backslash M' = \emptyset$.

We can express (C9^s) also as a criterion concerning anomalies. Indeed, it is equivalent to $I \backslash M \supseteq \mathcal{F}^{-1}[I' \backslash M']$,[18] i.e., it says: any anomaly of the reduced theory has only anomalous correlates in the reducing theory. Now, it is clear that (C9) is in contradiction with (C7). From (C9^s), we have, on account of $I \cap M \subseteq M \subseteq \mathcal{C}I \cup M$, $\mathcal{F}[I \cap M] \subseteq \mathcal{C}I' \cup M'$, which is tantamount to $\mathcal{F}[I \cap M] \cap (I' \backslash M') = \emptyset$.

Furthermore, Kamlah demands a counterpart to (C1) in his final definition of reduction (1985, p. 138):

(C1$^{s\prime\prime}$) $\quad I \supseteq \mathcal{F}^{-1}[I']$.[19]

Assuming again that $\mathcal{F}$ is onto, (C1$^{s\prime\prime}$) entails $\mathcal{F}[I] \supseteq \mathcal{F}[\mathcal{F}^{-1}[I']] = I'$, i.e., (C1^s). However, (C1^s) does by no means warrant (C1$^{s\prime\prime}$); this would be the case only if $I \supseteq \mathcal{F}^{-1}[\mathcal{F}[I]]$ were true – but for this $\mathcal{F}^{-1}$ would also have to be a function (at least on $\mathcal{F}[I]$) what is generally denied.[20] For that reason (C1$^{s\prime\prime}$) is in effect stronger than (C1^s), and consequently more difficult to justify. Kamlah's own remarks pertinent to (C1$^{s\prime\prime}$) are certainly not satisfactory.

It is interesting that Kamlah's conditions (C1$^{s\prime\prime}$) and (C9^s) can be substantiated by Adams' (C1^s) and (C2^s), if one accepts an additional assumption looking far less objectionable than (C1$^{s\prime\prime}$) (in fact, it follows from (C1$^{s\prime\prime}$)). I think it perfectly possible to determine an intuitive reduction function such that, for every intended application of the reduced theory, the inverse images all turn out to be either intended applications, or to be no intended applications of the reducing theory.[21] The additional assumption reads thus:

(A) $\quad \forall x' \in I'(\mathcal{F}^{-1}(x') \subseteq I \vee \mathcal{F}^{-1}(x') \subseteq \mathcal{C}I)$.

With the help of (A), (C1^s) implies (C1$^{s\prime\prime}$): if, for every $x' \in I'$, there is at least one $x \in I$ with $\mathcal{F}(x) = x'$, then, by (A), even $\mathcal{F}^{-1}(x') \subseteq I$, and we have (C1$^{s\prime\prime}$).

More important is that Kamlah's (C9^s) is deducible from Adams' conditions together with (A). Let $x \in \mathscr{C}I \cup M$; we have to show that $\mathscr{F}(x) \in \mathscr{C}I' \cup M'$ (if $x \in M_p^0$). First case: $x \in M$; but then $\mathscr{F}(x) \in M' \subseteq \mathscr{C}I' \cup M'$, by ($C2^s$), and we are done. Second case: $x \in \mathscr{C}I$; suppose that $\mathscr{F}(x)$ were not an element of $\mathscr{C}I' \cup M'$, i.e., $\mathscr{F}(x) \in I' \setminus M' \subseteq I'$; but then, by ($C1^{s''}$) which is already verified, we have $\mathscr{F}^{-1}(\mathscr{F}(x)) \subseteq I$, contradicting $x \in \mathscr{C}I$, and ($C9^s$) is proven.

Hence, if one is ready to accept (A), ($C9^s$) cannot express interesting properties beyond those of the Adamsian proposal. As, on the other hand, ($C2^s$) is not derivable from ($C1^{s''}$) and ($C9^s$), Kamlah's idea ends in a genuine liberalization of the original concept of reduction – a liberalization strong enough, however, to be still contradictory to the criterion (C7).

Let us return to Kamlah's innovation (C9). The underlying idea is anticipated very exactly in the last part of Niiniluoto's (1980, p. 36) rewording of the Sneed–Stegmüller definition of reduction; he formulates as definiens:

> there is a many-one relation R from M'_{pp} to M_{pp} [our $\mathscr{F}^{-1}$] such that the intended applications of T' are correlated with intended applications of T and what T' says *about these applications* is entailed by what T says *about the corresponding applications*. [my italics]

The difference between ($C2^s$) and ($C9^s$) seems to have escaped Niiniluoto's notice. But if you search for the source of this inaccuracy, you will come across similar sentences in Sneed (1971, p. 218, ll. 6–3 from the bottom; 1976, p. 136, ll. 18–22), Balzer and Sneed (1977, p. 202, ll. 8–5 from the bottom), and Stegmüller (1979, p. 36, ll. 23–25), which are much more puzzling than that of Niiniluoto. Let me quote Balzer and Sneed as an example

(C10) everything the reduced theory says about a given application is entailed by what the reducing theory says about any corresponding application.

Except for the fact that in Sneed's *full* theory model a theory says nothing about a *single* application (this is due to the constraints), the formulation of (C10) presupposes at least[22] the following (because the reducing theory mustn't state different things about different corresponding applications):

(A′) $\forall x' \in I'(\mathscr{F}^{-1}(x') \subseteq M \vee \mathscr{F}^{-1}(x') \subseteq \mathscr{C}M)$.

I am in doubt whether this is a desirable additional assumption.[23] But let us take (A′) for granted. Then, (C10) can be made precise, I think, in at least three different ways:

$$(\mathrm{C}10^{s})\quad \forall x' \in I'((\mathcal{F}^{-1}(x') \subseteq M \Rightarrow x' \in M') \,\& \\ \& \,(\mathcal{F}^{-1}(x') \subseteq \mathcal{C}M \Rightarrow x' \in M')).$$

$$(\mathrm{C}10^{s'})\quad \forall x' \in I'(\mathcal{F}^{-1}(x') \subseteq M \Rightarrow x' \in M').$$

$$(\mathrm{C}10^{s''})\quad \forall x' \in I'((\mathcal{F}^{-1}(x') \subseteq M \Rightarrow x' \in M') \,\& \\ \& \,(\mathcal{F}^{-1}(x') \subseteq \mathcal{C}M \Rightarrow x' \in \mathcal{C}M')).$$

($\mathrm{C}10^{s}$) agrees best with the structuralist account that the empirical claim of $T' = (M', I')$ is $I' \subseteq M'$, i.e., that for every $x' \in I'$, T' can only "say" that $x' \in M'$. ($\mathrm{C}10^{s'}$) strengthens this idea by requiring that every intended application of T' has to be a model "in the light of the reducing theory T" as well, i.e., that for every $x' \in I'$, T, too, can only "say" that $\mathcal{F}^{-1}(x') \subseteq M$. ($\mathrm{C}10^{s''}$), however, allows that some $x' \in I'$ are no models "in the light of T", and moreover interprets (C10) in such a way that T' "says the same" about those x', viz. $x' \notin M'$. Thus ($\mathrm{C}10^{s''}$) violates the prescipt that $I' \subseteq M'$ be the empirical claim of T'. Perhaps this can be justified by the fact that open-minded T'-theoreticians are quite willing to acknowledge that a certain portion of the intended applications I' consists of anomalies, and to cease from contending that all elements of I' are elements of M' as well.

Be that as it may, none of the above-mentioned conditions makes clear at first glance that the protagonists of structuralism almost always have in mind the Adamsian condition ($\mathrm{C}2^{s}$) (but cf. notes 9 and 15; ($\mathrm{C}1^{s}$) is treated as an extra requirement). A second glance I am going to take at ($\mathrm{C}10^{s}$)–($\mathrm{C}10^{s''}$) will indeed reveal that (C10) is not suitable for establishing ($\mathrm{C}2^{s}$).

Our most faithful transcription plainly cannot be a criterion for an interesting inter-theoretical relation between T and T'. Due to (A′), one of the antecedents is always satisfied, and ($\mathrm{C}10^{s}$) is simply equivalent to the "empirical claim" of T', viz., $I' \subseteq M'$.

The second version of (C10) is weaker and doesn't lead to (C2) either, but surprisingly to ($\mathrm{C}9^{s}$):

$$\begin{aligned}(\mathrm{C}10^{s'}) &\Leftrightarrow \forall x' \in I'\,(\mathcal{F}^{-1}(x') \nsubseteq M \vee x' \in M') \\ &\Leftrightarrow \forall x' \in I'\,(\mathcal{F}^{-1}(x') \subseteq \mathcal{C}M \vee x' \in M') \quad \text{(by (A′))} \\ &\Leftrightarrow \forall x' \in I'\,(x' \notin \mathcal{F}[M] \vee x' \in M') \\ &\Leftrightarrow I' \subseteq M' \cup \mathcal{C}\mathcal{F}[M] \\ &\Leftrightarrow I' \cap \mathcal{F}[M] \subseteq M' \\ &\Leftrightarrow (\mathrm{C}9^{s}) \quad \text{(cf. note 19).}\end{aligned}$$

Likewise the third reading – though stronger than ($C9^s$) – is not sufficient for ($C2^s$). We once more presuppose that $\mathscr{F}$ is onto in the following:

$$
\begin{aligned}
(C10^{s'''}) \quad &\Leftrightarrow I' \cap \mathscr{F}[M] \subseteq M' \;\&\; I' \cap \mathscr{F}[\mathscr{C}M] \subseteq \mathscr{C}M' && \text{(just as in } (C10^{s'})\text{)}\\
&\Leftrightarrow I' \cap \mathscr{F}[M] \subseteq M' \;\&\; I' \cap \mathscr{C}\mathscr{F}[M] \subseteq \mathscr{C}M' && \text{(by (A') and } \mathscr{F}\text{'s being onto)}\\
&\Leftrightarrow I' \cap \mathscr{F}[M] \subseteq M' \;\&\; I' \cap M' \subseteq \mathscr{F}[M]\\
&\Leftrightarrow I' \cap \mathscr{F}[M] = I' \cap M'.
\end{aligned}
$$

This condition determines $I' \cap \mathscr{F}[M]$ more exactly than ($C10^{s'}$). Roughly, ($C10^{s''}$) says that T and T' mark out the "same" models relative to the applications intended for T'. Outside I', however, there is no demand on the relation between $\mathscr{F}[M]$ and M'.

Now we have seen that the more prominent structuralists were not fully aware of the differences between absolute and application-relative criteria for the class of models. But unfortunately, Kamlah's (1985, p. 140) own comparison of his approach with that of Adams and Sneed is no great help either. First of all, he apparently considers ($C1^{s''}$) as equivalent to ($C1^s$) – an error, as we have seen. More important, the sole argument he advances against ($C2^s$) is no counter-argument at all: the existence of T'-models that "contradict" the "mathematical parts" of T, i.e., that have at most T-nonmodels as inverse images under $\mathscr{F}$, is of course perfectly compatible with ($C2^s$). And trying to make use of Kamlah's idea by changing the rôles of T and T', one only faces a problem if one presupposes, like Kamlah (1985, p. 133), that $M_p^0 = M_p$ – but usually this is explicitly denied.

FIFTH ACT: MORMANN

The last part of the criticism of the structuralism concept of reduction is contributed by Mormann. His starting point is a consideration of intuitive adequacy conditions for reductions, and he is the first to discover a (exactly one) contradiction. One of his four criteria is new to us, and finally the concept of explanation enters the stage. Prompted by a passage of Sneed (1976, pp. 138f) and Balzer and Sneed (1977, p. 204), Mormann (1984, p. 14) formulates a novel criterion he calls "the condition of the potential tightening of explanations":

(C11) Not all explanations of T' remain valid in the light of T.

(C11) is to interpret the intuitive requirement that T be a better theory giving more accurate explanations than T'. The concept of explanation proposed by Mormann (1984, p. 15) is the following (the term "explanation" admittedly sounds more appropriate when the full Sneedian machinery is used):

(D6) Let $T = (M, I)$ be a theory.
(a) A potential model $x \in M_p$ is said to be *explained by* T iff $x \in M$;
(b) an intended application $x \in I$ is said to be *explained by* T iff $x \in M$.

Depending on whether (a) or (b) is preferred, *explanations of* T can simply be equated with (sets of) models or (sets of) successful applications of T. Mormann's usage varies a little, and we shall therefore follow both variants.[24] The explication suggested by Mormann enables us to rewrite (C11) in a structuralist manner:

(D7) An explanation $E' \subseteq M'(\cap I')$ of T' is called *valid in the light of* T iff there is an $E \subseteq M$ with $\mathscr{F}[E] = E'$.

Obviously, the definiens of (D7) is equivalent to $\mathscr{F}[M] \supseteq E'$ (take $E := \mathscr{F}^{-1}[E'] \cap M$). And the quantification in (C11) can be dispensed with if the "maximal explanation" of T', viz., $M'(\cap I')$, is used. Thus the two reformulations of (C11) are

(C11^s) $\mathscr{F}[M] \not\supseteq M'$,

(C11$^{s'}$) $\mathscr{F}[M] \not\supseteq M' \cap I'$.

(C11^s) indeed is the condition of the passages of Sneed and Balzer/Sneed mentioned; the stronger condition (C11$^{s'}$), however, is new. I do not regard (D7) as the most natural definition and would rather propose a more obvious alternative. As we refrain from using constraints, it is sufficient to consider explanations point by point instead of collecting them into sets.

(D7′) An explanation $x' \in M'$ (or $x' \in M' \cap I'$) is called *valid in the light of* T iff there is an explanation $x \in M$ (resp., $x \in M \cap I$) of T with $\mathscr{F}(x) = x'$.

There is no change if the interpretation of (C11) is based on (D6)(a), but (C11$^{s'}$) is weakened to a third version of (C11) if (D6)(b) – which presumably fits the intuition better – is applied:

(C11$^{s''}$) $\mathscr{F}[M \cap I] \not\supseteq M' \cap I'$.

Even the nonformal wording of (C11) makes it evident that this criterion is inconsistent with (C4). According to our "double entry" in (D6), (C4) is now easily transformed into these inclusions:

(C4^{s}) $\mathscr{F}[M] \supseteq M'$,

(C4$^{s'}$) $\mathscr{F}[M \cap I] \supseteq M' \cap I'$.

As was to be expected, (C4^{s}) and (C4$^{s'}$) are negations of (C11^{s}) and (C11$^{s''}$). A more interesting observation is that the two versions of (C4) already have found quite independent support: (C4^{s}) turns out to be the Mayr–Pearcean condition (C2$^{s'}$), and (C4$^{s'}$) exactly is the Pearcean supplementary condition (C8$^{s'}$)(b). In this respect, (C4) surprisingly provides the critics of Adams and Sneed with arguments, while (C11) works in their defence. But note that the strongest interpretation of (C11), i.e., Mormann's (C11$^{s'}$), is in contradiction to the strongest interpretation of the Balzer–Sneed condition (C10), i.e., my (C10$^{s''}$).

Mormann thinks that he can resolve the inconsistency of (C11^{s}) (and (C11$^{s'}$)) with (C2$^{s'}$). For this purpose he extensively draws on constraints, so his attempt is beyond the purview of this paper. The present diagnosis is far more critical than his one anyhow.

THE MORAL

Perhaps it is not enough to make a well-knit dramatic play out of five short stories. The reader may have failed to. keep track of the dénouement. We had criteria for intended applications ((C1)), for models ((C2), versions of (C4), (C8), and (C11)), for intended applications *and* models ((C5) and (C6)), for anomalies ((C3), (C9), and a version of (C10)), for successful applications (versions of (C4), (C8), (C10), and (C11)) and for anomalies *and* successful applications ((C7)). Let us make the structure of the plot somewhat more transparent.

THEOREM 1 (implications).

(a) (C1$^{s'}$) $\Rightarrow$ (C1^{s}), (C1$^{s''}$) $\Leftrightarrow$ (C1^{s}) & (A);

(b) $(C1^{s''})$ & $(C2^{s}) \Rightarrow (C9^{s}) \Rightarrow (C3^{s})$;

(c) $(C1^{s})$ & $(C2^{s'}) \Leftrightarrow (C5^{s})$;

(d) $(C2^{s'}) \Leftrightarrow (C4^{s}) \Leftrightarrow (C8^{s'})(a) \Rightarrow (C8^{s''})(a)$;

(e) $(C4^{s'}) \Leftrightarrow (C8^{s'})(b) \Rightarrow (C8^{s''})(b)$;

(f) $(C10^{s''}) \Rightarrow (C10^{s'}) \Leftrightarrow (C9^{s})$;

(g) $(C11^{s'}) \Rightarrow (C11^{s})$, $(C11^{s'}) \Rightarrow (C11^{s''})$;

(h) $(C8^{s'})(b)$ & $(C9^{s}) \Rightarrow (C1^{s})$.

This list is, as far as I can see, complete. The criteria not mentioned or not interwoven here come into effect in the next theorem, which can be extended with the help of theorem 1. Note especially the somehow complementary rôles of (C7) and (C11) in the criticism of the Adams–Sneedian and the Mayr–Pearcean proposals respectively.

THEOREM 2 (incompatibilities). The following pairs are logically incompatible:

(a) $(C2^{s}) - (C7^{s})$;

(b) $(C2^{s'}) - (C11^{s})$ (or equivalently, $(C4^{s}) - (C11^{s})$);

(c) $(C6^{s}) - (C1^{s})$ & $(C2^{s'})$ (& $(C8^{s'})(b)$);

(d) $(C8^{s'})(b) - (C11^{s''})$ (or equivalently, $(C4^{s'}) - (C11^{s''})$);

(e) $(C9^{s}) - (C7^{s})$;

(f) $(C10^{s''}) - (C11^{s'})$;

the following pair is practically incompatible:

(g) $(C3^{s}) - (C7^{s})$ (since it necessitates $I \not\subseteq M$).

Those who, like myself, think that the criticisms of the structuralist concept of reduction advanced by Mayr, Pearce, Kamlah, and Mormann all have some plausibility may try to replace the Adams–Sneedian criteria by a new approach. On the one hand, $(C1^{(s)})$ can be given up without any loss, since – prima facie – not every intended application of T' must have an $\mathcal{F}$-correlate: at first one is not at all anxious for the anomalies of T' to be mirrored in T. On the other hand, one need not insist on $(C2^{(s)})$ if the (artificial?) separation of the applied and the formal aspect of a theory is rescinded:[25] instead of the laws, the empirical claims – as suggested by Kamlah – are required to stand in an entailment relation. Accordingly, let us combine ($C9^s$) with the quite efficient basic criterion ($C2^{s'}$) favoured by Mayr and Pearce:

$$\text{(C12)} \quad M' \subseteq \mathscr{F}[M] \subseteq \mathscr{F}[\mathscr{C}I \cup M] \subseteq \mathscr{C}I' \cup M'.$$

The second inclusion is of course trivial. For illustration I once again add what mustn't happen concerning the models and anomalies if (C12) is operative: first, no T-model is allowed to represent a T'-anomaly by means of $\mathscr{F}$ (if this were the case, then remove the T-model in question from M_p^0), and second, the inverse image of a T'-model mustn't consist of T-anomalies only.

As ($C11^{s''}$) seems to be the best reading of (C11), the inconsistency of (C12) with Mormann's ($C11^s$) (and ($C11^{s'}$)) is not very annoying. But we also know that (C12) is incompatible with Mayr's idea of the "explanation of T'-anomalies by T". *After* verifying the existence of an intuitively convincing function $\mathscr{F}$ obeying (C12), however, it is possible to investigate whether there are some T'-anomalies $x' \in I' \backslash M'$ that have T-counterparts in $(I \cap M) \backslash M_p^0$. If this test proves positive then, in a second reduction step, enlarge M_p^0 suitably so that a function $\mathscr{F}^*$ is formed which satisfies ($C7^s$) (but, of course, no longer (C12)).[26]

Thus we have managed to furnish a round dozen of criteria, but I must confess that even this last two-stage plan to save the structuralist concept of reduction doesn't seem watertight to me. Anyhow, *every* variation of Adams' conditions loses the so simple and clear idea of the old empiricist definability-cum-derivability concept. For that reason I am inclined to draw entirely different conclusions in the case of reduction from the above discussion.

After all we have a result: the (necessary) conditions of adequacy proposed by the various parties are incompatible. On visualizing that in truely progressive reductions we have to expect $\subsetneqq$ (or $\supsetneqq$) instead of $\subseteq$ (or $\supseteq$), it becomes evident that not even ($C2^s$) and ($C2^{s'}$) are compatible. With the exception of Mormann's essay, I have found no indication of this problem in the literature.[27] One can only speculate as to why there is so little awareness here. For one thing, the continual changes in presentation (notations and conventions) are impeding the immediate grasp of the matter. For another, the technical expenditure caused by constraints and the distinction of partial and full potential models is adding to the embarrassing state of affairs. Though well-founded as such, constraints and theoretical functions would seem superfluous especially for the discussion of reduction.[28] Perhaps my result was obtained with so small an effort only because I confined myself to the simple theory model of Adams and formulated the criteria as relations between (inverse) image classes that are easy to handle. But this is no explanation for the fact that specialists have given distorted[29] or even false[30] pictures of the interrelations. That leads us to assume that the concept of reduction in its structuralist appearance is based on too immature intuitions; that all criteria are legitimate in one way or the other, but are nothing but an expression of a family resemblance of the relations between replaced and replacing theories; in short, that we have to look for a new pattern for the rational reconstruction of theory dislodgement.

For fear this paper might look too destructive, I would like to sketch briefly what pattern I have in mind. I think it is quite possible to revive the d-concept of reduction, at once retaining its intuitive appeal and repairing its shortcomings. The discussion reviewed above indeed is somewhat fictitious, since T' is generally in only approximate agreement with T, even within its own range of application, i.e., strictly speaking, T' is inconsistent with T. This is the first and foremost of the two challenges of Feyerabend.[31] In order to solve this problem philosophers of science used to replace T' or T in (d) by *approximate*, and respectively *extended*, "versions" T'^* or T^*. Such a procedure is plausible but not very instructive. I consider it better to make a transition from (d) to

(d*) $\quad T^*_{\mathrm{A}} \vdash T'$,

where T^*_{A} is a ***revision*** of T needed to accept A, in the sense of Gärdenfors and his collaborators.[32] This term refers to a process of theory change which has recently been developed for the semantics of (counterfactual) conditionals and can be described rather constructively. It allows A to designate contrary-to-fact conditions (i.e., conditions contrary to what T says is a fact),[33] and hence T and T' to be inconsistent. Of course, (d*) must first prove useful in case studies, and then it must show how we can get out of the dilemma when-to-apply-which-criteria-for-what-reasons.

NOTES

* I wish to thank David Pearce for enlightening criticism and Winfred Klink for correcting my English.

[1] Cf. for instance Tuomela (1978, pp. 222f), Niiniluoto (1980, p. 25), Pearce (1981, pp. 24f, 29f; 1982, p. 330). Even in Stegmüller (1979, pp. 48, 87f) similar remarks can be found.

[2] These are not the original assumptions, but amendments first suggested by Rantala. Cf. Niiniluoto (1980, pp. 9–11), Pearce (1982, pp. 312f) and the programmatic essay of Pearce and Rantala (1983), where in particular the terms "similarity type" and "general logic" are explained.

[3] For motivation, see e.g. Stegmüller (1973, p. 145). In Adams (1959) both conditions still are absent; in Sneed (1971, p. 221) $\mathscr{F}$ is not required to be onto, but in Sneed (1976, pp. 122, 136f) it is. Kamlah (1985, p. 133) stipulates that $M_p^0 = M_p$, but not that $\mathscr{F}$ be onto. In the applications of Pearce and Rantala both requirements appear, but always without the index "p", cf. e.g. Pearce and Rantala (1984, pp. 171f).

[4] Cf. Adams (1959, pp. 261f) and Sneed (1971, pp. 231f); more general remarks are made by Stegmüller (1979, pp. 42f).

[5] A protection against excessive humbuggery with arbitrary reduction functions $\mathscr{F}$ is provided by the formal requirements that $\mathscr{F}$ is to map M_p^0 *onto* M_p' and that this M_p' be definable as the class of *all* structures of type τ'. – Mormann (1984, pp. 45–48) indicates a more interesting structuralist explication of the intuitive requirement.

[6] Cf. Mayr (1976, pp. 286f), Tuomela (1978, pp. 220, 226) and particularly Hoering (1984, pp. 37–39). Hoering's (1984, pp. 35f) discussion of Eberle's (1971) syntactic approach is a warning that the statement view is exposed to quite the same dangers.

[7] Formulations are varying here: Adams (1959, p. 260) uses the term "define", Sneed (1971, p. 217) "correspondence", and Stegmüller (1976, p. 128) "transpose".

[8] Adams (1959, p. 261) says "derive", Sneed (1971, p. 220) "deduce", and Stegmüller (1976, p. 128) "map".

[9] The credit goes to Mayr for having pointed out the fact that Sneed (1971, p. 229) and Stegmüller (1973, p. 151) originally did *not* propose (C2^s), but

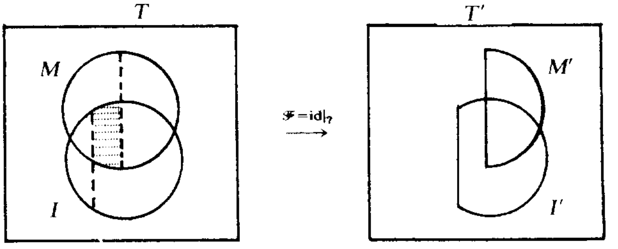
$$(\mathrm{C2}^{s*}) \qquad \forall \emptyset \neq N' \subseteq M' \exists N \subseteq M (\emptyset \neq \mathscr{F}[N] \subseteq N')$$

as a criterion for the reduction of theories ("theory elements" as they were to call it later). The conditions that N' and $\mathscr{F}[N]$ be non-empty presumably are intended by Sneed and Stegmüller and have been added by me. Otherwise, (C2^{s*}) would be completely trivial (simply take $N \subseteq M \backslash M_p^0$). Mayr's theorems 2.7 and 2.8 (1976, pp. 284, 286), according to which (C2^{s*}) would be stronger than (C2^s), are not correct; however, the proofs show that (C2^{s*}) is equivalent to $\mathscr{F}[M] \supseteq M'$ – a curious result, in the light of the following discussion! In Sneed (1976, pp. 137f) and Stegmüller (1979, p. 96) you can find (more complicated versions of) (C2^s) as a criterion. I wonder why neither Sneed nor Stegmüller commented on their – certainly considerable – change of mind (or on Mayr's pertinent observations).

[10] Tuomela (1978, pp. 220f) thinks that the concept of reduction can only be analyzed by the concept of explanation. In contrast, Kamlah (1985, pp. 124f) is of the opinion that (C4) points to an alternative explicandum.

[11] For motivation, see Mayr (1976, pp. 279f, 287).

[12]



Sketches like these are very useful in testing criteria (but note that in general $M_p^0 \subsetneq M_p$).

[13] Yet, it is conceivable that the rule of autodetermination (cf. Stegmüller 1973, pp. 224–231) is effective in the construction of specializations. Then (C6) definitely speaks against the concept of reduction according to $(C1^s)$ and $(C2^s)$.

[14] The substitution of $(C1^{s'})$ for $(C1^s)$ is not efficient if $(C2^s)$ is the second requirement; an $\mathcal{F}$ satisfying $(C1^s)$ and $(C2^s)$ could be turned into a reduction function satisfying $(C1^{s'})$ by restricting the domain of $\mathcal{F}$ to $\mathcal{F}^{-1}[I']$. But cf. note 5.

[15] Balzer (1982, p. 298) uses $(C2^s)$ and $(C2^{s'})$ not in competition, but in combination. The identity $\mathcal{F}[M] = M'$ is also implied by the introduction of $\mathcal{F}$ in the papers of Pearce and Rantala (cf. note 3), who now regard this situation as characteristic for only one type of reduction they call "embedding".

[16] Pearce (1981, p. 25; 1982, p. 314) and Balzer (1982, p. 220) proceed in a more abstract manner and call $\mathcal{F}$ itself a translation of T' into T iff for every T'-sentence ψ, there is a translation of ψ (relative to $\mathcal{F}$).

[17] This quotation already can almost literally be found in Kemeny and Oppenheim (1956, p. 13), who, however, thank Hempel for his clarifying remarks in a footnote to the pertinent paragraph.

[18] It is easy to show that this condition – let us call it $(C9^{s'})$ – is equivalent to $(C9^s)$:

> $(C9^s) \Rightarrow (C9^{s'})$: Let $x \in \mathcal{F}^{-1}[I' \backslash M']$, i.e., $\mathcal{F}(x) \in I' \backslash M' = \mathscr{C}(\mathscr{C}I' \cup M')$, hence, by $(C9^s)$, $x \notin \mathscr{C}I \cup M$, i.e., $x \in I \backslash M$.
>
> $(C9^{s'}) \Rightarrow (C9^s)$: Let $x' \in \mathcal{F}[\mathscr{C}I \cup M]$, i.e., there is an $x \in \mathscr{C}I \cup M$ with $\mathcal{F}(x) = x'$, hence $\mathcal{F}^{-1}(x') \nsubseteq I \backslash M$, hence, by $(C9^{s'})$, $x' \notin I' \backslash M'$, i.e., $x' \in \mathscr{C}I' \cup M'$.

[19] $(C1^{s''})$ is intuitively required by Balzer and Sneed (1977, pp. 202f), too; in the subsequent definition, however, they choose $(C1^s)$. – Kamlah (1985, pp. 136–138) needs $(C1^{s''})$ as a premise to show the equivalence of $(C9^s)$ with the condition $I' \cap \mathcal{F}[M] \subseteq M'$ – let us call it $(C9^{s''})$. As an illustration how very frightening-looking theorems and derivations shrink to moderate proportions in the simple theory model, here is a small proof of this equivalence:

> $(C9^s) \Rightarrow (C9^{s''})$: Let $x' \in I' \cap \mathcal{F}[M]$; hence, there is an $x \in M$ with $\mathcal{F}(x) = x'$, and thus, by $(C9^s)$, $x' \in \mathscr{C}I' \cup M'$ and, since $x' \in I'$, also $x' \in M'$.
>
> $(C9^{s''}) \Rightarrow (C9^s)$: Let $x' \in \mathcal{F}[\mathscr{C}I \cup M]$; hence, there is an $x \in \mathscr{C}I \cup M$ with $\mathcal{F}(x) = x'$; if $x \in \mathscr{C}I$, then, by $(C1^{s''})$, also $x' = \mathcal{F}(x) \in \mathscr{C}I'$; if on the other hand $x \in M$, then, by $(C9^{s''})$, $x' = \mathcal{F}(x) \in \mathscr{C}I' \cup M'$.

[20] If ($C2^s$) or ($C9^s$) (but not ($C2^{s'}$)) is to be the second condition, one can get ($C1^{s''}$) from ($C1^s$) by taking $\mathscr{F}|_I$ instead of $\mathscr{F}$. But cf. note 5.

[21] For instance, this is already ensured if there is a translation τ from T' into T (relative to $\mathscr{F}$) and a T'-sentence ψ with $\|\tau(\psi)\| = I$, i.e., roughly speaking, if the class of intended applications of T (within M_p^0) is definable in T'.

[22] Sneed's formulation of 1971 presupposes even more, namely that $\mathscr{F}^{-1}$ is a function.

[23] I don't believe that (A′) can be justified by the argument mentioned in note 21, since the class of T-models (within M_p^0) presumably cannot be defined in T'. But (A′) sometimes may be guaranteed by the formal trick of choosing (with the help of the axiom of choice) ***one*** $x(x') \in I$ with $\mathscr{F}(x(x')) = x'$ for every $x' \in I'$ (such an $x(x')$ exists by ($C1^s$)), and then using $\mathscr{F}|_{\{x(x'):\, x' \in I\}}$ instead of $\mathscr{F}$. But cf. note 5.

[24] It seems to me, however, that (b) is fractionally better. Considering the class of potential models ***not*** explained by T, (a) suggests this class to be $\mathscr{C}M$, whereas (b) suggests it to be the class of anomalies $I \backslash M$. It is the latter that hits the common usage more precisely.

[25] This separation also is removed in Niiniluoto's (1980, pp. 26f) minimal transformation of the non-statement view into the statement view, replacing a structure $\langle M, I \rangle$ by the statement "Every I is an M".

[26] But cf. note 5.

[27] Hoering's (1984, pp. 47f) suspicion that the "requirements of definability and derivability" ((C1) and (C2)) must sometimes be weakened in order to avoid inconsistencies is certainly not relevant here. He mentions approximations and undefinabilities of theoretical functions as possible reasons, aspects I have left out of consideration in this paper.

[28] For instance, you do not learn anything substantial about reductions from the reflections in Sneed (1971, pp. 224f), but only about some technical problems of his sophisticated theory model. – However, Mormann thinks that his inconsistency (cf. theorem 2(b)) can be overcome just by a massive employment of constraints. Thus he modifies the structuralist readings of (C8) and (C5). But I suspect that if even one single criterion is modified then consequently every other criterion must be modified too – e.g., with reference to $M \cap C$ for a maximal constraint C (cf. Mormann 1984, p. 30) – and therefore the problems will reappear in an analogous way.

[29] Pearce's reference to Stegmüller's (C8) is, as seen above, not quite accurate. Besides, I wasn't able to verify Pearce's (1982, pp. 309, 323) contention that Mayr bases his criticism of the Adams–Sneed concept on that very criterion (C8). Hoering's (1984, pp. 38f) recapitulation of Pearce (in fact, it is a recapitulation of Pearce's recapitualtion of Sneed – Pearce himself favours different criteria) could be mentioned here or in the next note; his first "theorem" is certainly not correct.

[30] I have mentioned Stegmüller (1973, p. 146), Sneed (1976, p. 139), Mayr (1976, pp. 276, 284, 286f), and Kamlah (1985, p. 140).

[31] Checking the central passages on meaning variance and incommensurability in (the earlier) Feyerabend's work, you will find good reasons for saying that *the* source of (Feyerabend's) incommensurability is inconsistency. This has been paid far too little attention up to now, I think. Cf. Feyerabend (1962, pp. 57, 59, 74f, 81f) and (1963, p. 30).

[32] As a first information, see Makinson (1985). A comprehensive treatise by Gärdenfors (1987), summing up and developing all pertinent work further, is in preparation.

[33] Cf. Glymour (1970), esp. p. 341: "Inter-theoretical explanation is an exercise in the presentation of counterfactuals."

REFERENCES

Adams, E.: 1959, 'The Foundations of Rigid Body Mechanics and the Derivation of Its Laws from Those of Particle Mechanics', in L. Henkin, P. Suppes and A. Tarski (eds.), *The Axiomatic Method*, North-Holland, Amsterdam, pp. 250–265.

Balzer, W.: 1982, *Empirische Theorien: Modelle – Strukturen – Beispiele*, Vieweg, Braunschweig, Wiesbaden.

Balzer, W. and J. D. Sneed: 1977, 1978, 'Generalized Net Structures of Empirical Theories I, II', *Studia Logica* **36**, **37**, 195–211, 167–194.

Eberle, R. A.: 1971, 'Replacing one Theory by Another under Preservation of a Given Feature', *Philosophy of Science* **38**, 486–501.

Feyerabend, P. K.: 1962, 'Explanation, Reduction, and Empiricism', in H. Feigl and G. Maxwell (eds.), *Minnesota Studies in the Philosophy of Science*, Vol. III, *Scientific Explanation, Space and Time*, University of Minnesota Press, Minneapolis, pp. 28–97.

Feyerabend, P. K.: 1963, 'How to Be a Good Empiricist – A Plea for Tolerance in Matters Epistemological', in B. Baumrin (ed.), *Philosophy of Science*, The Delaware Seminar, Vol. 2 (1962–1963), Interscience Publishers, New York, London, Sidney, pp. 3–39.

Gärdenfors, P.: 1987, *Rational Changes of Belief*, Bradford Books, MIT Press, Cambridge, Mass.

Glymour, C.: 1970, 'On Some Patterns of Reduction', *Philosophy of Science* **37**, 340–353.

Hempel, C. G.: 1965, *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*, Free Press, New York, London.

Hoering, W.: 1984, 'Anomalies of Reduction', in W. Balzer, D. Pearce and H.-J. Schmidt (eds.), *Reduction in Science*, Reidel, Dordrecht, Boston, Lancaster, pp. 33–50.

Kamlah, A.: 1985, 'On Reduction of Theories', *Erkenntnis* **22**, 119–142.

Kemeny, J. G. and P. Oppenheim: 1956, 'On Reduction', *Philosophical Studies* **7**, 6–19.

Makinson, D.: 1985, 'How to Give It Up: A Survey of Some Formal Aspects of the Logic of Theory Change', *Synthese* **62**, 347–363.

Mayr, D.: 1976, 'Investigations of the Concept of Reduction I', *Erkenntnis* **10**, 275–294.

Mormann, T.: 1984, 'Strukturalistische Reduktionsbeziehungen als Morphismen von Constraintkategorien', *mimeo*, Bielefeld.

Nagel, E.: 1949, 'The Meaning of Reduction in Natural Sciences', in R. C. Stauffer (ed.), *Science and Civilization*, University of Wisconsin Press, Madison, pp. 99–145.

Niiniluoto, I.: 1980: 'The Growth of Theories: Comments on the Structuralist Approach', in J. Hintikka, D. Gruender and E. Agazzi (eds.), *Theory Change, Ancient Axiomatics, and Galileo's Methodology*, Reidel, Dordrecht, pp. 3–47.

Pearce, D.: 1981, 'Is There Any Justification for a Nonstatement View of Theories?', *Synthese* **46**, 1–39.

Pearce, D.: 1982, 'Logical Properties of the Structuralist Concept of Reduction', *Erkenntnis* **18**, 307–333.

Pearce, D. and V. Rantala: 1983, 'New Foundations for Metascience', *Synthese* **56**, 1–26.

Pearce, D. and V. Rantala: 1984, 'Limiting Case Correspondence between Physical Theories', in W. Balzer, D. Pearce and H.-J. Schmidt (eds.), *Reduction in Science*, Reidel, Dordrecht, Boston, Lancaster, pp. 153–185.

Sneed, J. D.: 1971, *The Logical Structure of Mathematical Physics*, Reidel, Dordrecht.

Sneed, J. D.: 1976, 'Philosophical Problems in the Empirical Science of Science: A Formal Approach', *Erkenntnis* **10**, 115–146.

Stegmüller, W.: 1973, *Probleme und Resultate der Wissenschaftstheorie und Analytischen Philosophie*. Band II, *Theorie und Erfahrung*, 2. Halbband, *Theorienstrukturen und Theoriendynamik*, Springer, Berlin, Heidelberg, New York. (English 1976, as *The Structure and Dynamics of Theories*, Springer, New York, Heidelberg, Berlin.)

Stegmüller, W.: 1979, *The Structuralist View of Theories*, A Possible Analogue of the Bourbaki Programme in Physical Science, Springer, Berlin, Heidelberg, New York.

Tuomela, R.: 1978, 'On the Structuralist Approach to the Dynamics of Theories', *Synthese* **39**, 211–231.

Manuscript received 8 April 1986

Kellerstr. 11
D-8000 München 80
F.R.G.