This is a version of the paper published in the Journal of Philosophy vol.95, no.2 (February, 1998), 53-72.

Collapse of the New Wave*

There is a movement in the philosophy of mind recently dubbed "the new wave," one built upon "a novel general account of theory reduction in science." [1] Advocates hail Paul Churchland and Clifford Hooker, [2] claiming that their model enjoys a number of theoretical virtues over its competitors. While discounting these advertised virtues, I argue that the new wave model has but one genuinely novel element, and it should be rejected. Moreover, once its ontological consequences are duly noted, the reductive part the model collapses into the classical theory developed within the logical empiricist tradition, so that it still falls prey to standard anti-reductionist argument. The new wave is thus powerless to erode anti-reductionist sentiment and destined to recede into the distant past, leaving the conceptual landscape largely unchanged.

I. THE MODEL

New wave reduction involves three theories: the basic reducing theory T_B, the original reduced theory T_R, and a corrected image T_R*. With an eye to ideal cases, Paul Churchland describes it thus:

> [A] successful reduction ideally has the outcome that, under the term mapping effected by the correspondence rules, the central principles of T_R (those of semantic and systematic importance) are mapped onto general sentences of T_B that are theorems of T_B. Call the set of such sentences T_R*. This set is the image of T_R within T_B. [3]

Emphasizing a key aspect that will loom large in later discussion, Churchland stipulates that the corrected T_R* must be part of the basic theory, being "general sentences of T_B." He also refers to a "mapping" with "correspondence rules" between this base-level T_R* and the original reduced T_R. Yet, given the nature of T_R*, these rules no longer have a familiar classical function. As Churchland goes on to say:

> First, on the account given above it is not the reduced theory,T_R, that is deduced from the principles of T_B, as some other accounts have it. What is deduced from T_B is rather the set T_R*, an equipotent image of T_R within

> the idiom of T_B. Second, it is important to appreciate that cross-theoretic identity claims, even if they are justly made, are not part of the reduction proper, and they are not essential to the function it performs (ibid., p.83).

So reduction remains deduction. But, unlike the classical account, cross-theoretic identity claims are "not part of the reduction proper," since it is not the original T_R but rather T_R* that is deduced from T_B, an image specified in the basic vocabulary. Hence, inter-theoretic rules are inessential both to what is deduced and how. Instead, their primary function is to show how T_R* can capture the explanatory role of T_R. As Churchland puts it:

> The correspondence rules play no part whatever in the deduction. They show up only later, and not necessarily as material-mode statements, but as mere ordered pairs: <Ax,Jx>, <Bx,Kx>, <Cx,Lx>, <Dx,Mx>. Their function is to indicate which term substitutions in the image T_R* will yield the principles of T_R. The older theory, accordingly, is never deduced; it is just the target of a relevantly adequate mimicry. [4]

Clifford Hooker expresses the emerging picture more formally in this way:

> Within T_B construct an analog, T_R*, of T_R under certain conditions C_R such that T_B and C_R entails T_R* and argue that the analog relation, AR, between T_R and T_R* warrants claiming (some kind of) reduction relation R, between T_R and T_B. Thus $(T_B \bullet C_R \rightarrow T_R^*) \bullet (T_R^* \text{ AR } T_R)$ warrants $(T_B \text{ R } T_R)$. [5]

Of course the analogue relation is not always strong, and the terms of T_R will not always map smoothly onto the terms of T_R*. In Churchland's words: "we must be prepared to count reducibility as a matter of degree. Like translation, which may be faithful or lame, reduction may be smooth, or bumpy, or anywhere in between." [6] So there is a continuum of strong to weak analogies between reduced and reducing theories, one that carries important metaphysical consequences. In the smooth case, though the deduction does not contain cross-theoretic identity claims, nevertheless they are justified by the success of the overall reduction. Churchland is explicit:

> On this view, full-fledged identity statements are licensed by the comparative smoothness of the relevant reduction (i.e., the limiting

> assumptions or boundary conditions on T_B are not wildly counterfactual, all or most of T_R's principles find close analogues in T_R*, etc.). This smoothness permits the comfortable assimilation of the old ontology with the new and thus allows the old theory to retain all or most of its ontological integrity. It is smooth intertheoretic reductions that motivate and sustain statements of cross-theoretic identity, not the other way around. [7]

On the other hand, when the analogy is weak and the particular case is not smooth, then the ontology of the original theory is replaced, not reduced. As Patricia Churchland says:

> The evolving unifications seen in science therefore encompass not only smooth reductions with cross-theoretic identifications but also rather 'bumpy' reductions where cross-theoretic identifications are problematic and involve revision of the old theory's concepts, and outright elimination with no cross-theoretic identifications at all. [8]

Or again, mindful of the metaphysical consequences, John Bickle describes it thus:

> [T]his alternative account of intertheoretic reduction recognizes a spectrum of possible outcomes, and insists that the appropriate ontological consequences depend upon where on this spectrum a given case falls. We can lay out this spectrum (informally) as follows: at the left-most endpoint lie the perfectly smooth reduction pairs, where T_R* is the exactly equipotent isomorphic image of T_R ... At the right-most endpoint lie the extremely bumpy or replacement cases. And separating these two extremes is a continuous spectrum of cases, approximating more or less closely one of the two endpoints. [9]

So, putting together the various points, new wave reduction involves four main features:

> (i) NW CONSTRUCTION -- the basic reducing T_B, not the original reduced T_R, supplies the conceptual resources for constructing the corrected T_R*;
> (ii) NW DEDUCTION -- the corrected T_R*, not the original reduced T_R, is deduced from the basic reducing T_B;
> (iii) NW RELATION -- there is a required analogical relation, not bridge laws, between the reduced T_R and the corrected T_R*;
> (iv) NW CONTINUUM -- there is a continuum of strong to weak analogies between the reduced T_R and the corrected T_R*, with the strong relations justifying retention and the weak relations justifying replacement of the ontology of T_R.

Viewed as a process, one can summarize the account in two stages: first, the intra-theoretic deduction, which focuses on conditions (i) and (ii); second, the inter-theoretic mapping, which focuses on conditions (iii) and (iv). Once the second stage is complete, meaning once the analogue relation between T_R* and T_R has been established, then one can make a claim about the reduction of the original T_R by virtue of the deduction of its analogue T_R*.

II. AFFINITIES WITH LOCAL & APPROXIMATE REDUCTION

Certainly the new wave features have no echo in the classical model. There were bridge laws between the reduced and reducing theories, not mere analogies; the bridge laws played an essential role in the derivation of the original reduced theory, not the derivation of a corrected image; and even if a corrected image was implied, it need not have been constructed from the basic reducing theory.

Yet the new wave model is not without precedent. For example, condition (ii) on new wave deduction is implied by previous accounts of domain-specific reduction developed by Jaegwon Kim and David Lewis according to which a corrected T_R* is deduced from the basic reducing T_B, where the former is a domain-specific version of a more general and otherwise by the classical view reductively recalcitrant T_R. [10]

Even closer in spirit and ideology, both conditions (ii) and (iii) are explicit in Kenneth Schaffner's earlier model of approximate reduction. [11] Specifically, like condition (ii), Schaffner stipulates "that T_R* be derivable from T_B," a theory which "corrects" the original T_R; and like condition (iii), he stipulates that the "relation between T_R and T_R* should be one of strong analogy," the predictions of T_R* being "close" to those of the original (op. cit., p.144). Also, the new wave continuum of condition (iv) is easily generated from Schaffner's account by loosening his strong analogy into a spectrum of strong to weak relations of the same kind. Moreover, there can be no difference if one reserves the term 'reduction' for cases at the retentive end of continuum where, by new wave lights, the analogy is strong (involving "close analogues" and "equipotent images"). Consonant with this well-established usage, Hooker preserves the distinction between 'reduction' and 'replacement' by applying the former exclusively to "cases which fall sufficiently far towards the retention end of the retention/ replacement

spectrum." [12] So too does Patricia Churchland, who employs the same distinction from the opposite direction, saying that disanalogous cases which require massive correction are "better described as having been displaced outright." [13]

So, on the matter of conditions (ii) through (iv), the Schaffner and Churchland-Hooker models are virtually indistinguishable. Condition (i) on new wave construction is the genuinely novel element -- that the basic reducing T_B and not the original reduced T_R must supply the conceptual resources for the corrected image T_R*. Only with this constraint can the older Schaffner conditions appear in a new light, turning what might have been a corrected heterogeneous deduction of T_R* from T_B into a purely intra-theoretic one, or, contrawise, turning what might have been an intra-theoretic analogy between T_R* and T_R into a purely heterogeneous one. [14]

III. A MATTER OF DISPLACEMENT

Proponents of the new wave claim their model exhibits a number of theoretical virtues. The first concerns displacement of the original theory. Paul Churchland says of the two-stage process:

> The intra-theoretic deduction (of T_R* within T_B), and the inter-theoretic mapping (of T_R into T_R*), constitute a fell-swoop demonstration that the older theory can be displaced wholesale by the new, without significant explanatory or predictive loss. [15]

More precisely, the new and corrected T_R* has the resources to <u>mimic</u> the explanatory role of the original T_R in an ideally smooth case, and <u>better</u> the explanatory role of T_R in a rough and disanalogous case, either way demonstrating the old's replaceability by the new. And surely the point is intriguing, since, contrary to appearances, if Churchland is correct then the new wave model does not yield a continuum from retention to replacement after all. No, it is <u>theory replacement across the board</u>. [16]

Yet there are problems. First, even if the original "theory" can always be displaced, not so for the theory's ontology. For in a smooth and strongly analogous case

the ontology of T_R is retained, since claims about identity are sustained. So on this side of the spectrum there is replacement only in a weak and attenuated sense, a mere exchange of labels with referents and roles largely preserved.

Second, even this weak sense of replacement appears suspect. Taking theories as sets of sentences, for example, T_R should be individuated in such a way as to contain all the sentences adequate to express the objects and properties in its domain. But this will include T_R* in ideally smooth cases, since, ex hypothesi, T_R* is an equipotent image that preserves the ontology of the original. Indeed, Hooker indicates that T_R* is the original T_R in cases of "perfect retention," which, either way, makes replacement impossible. [17] (It follows, moreover, that there is no uniform theory of reduction, since at this endpoint of retention T_R* (= T_R) is not specified within T_B -- assuming T_R is not -- contrary to what the new wave generally requires.)

Third, and ideal cases to one side, some argue that replacement only makes sense in the intra-level case where theories compete for the same logical space and target the same explanandum phenomena. [18] If so, a higher-level T_R could not be eliminated by a purely base-level T_R*. The novel new wave constraint would thus show itself not a help but a hindrance to replacement.

Fourth, it is no accident that Churchland refers to a "wholesale" replacement of T_R, since the corrected T_R* must be specified entirely within T_B, excluding any contribution from the original T_R. [19] But, on the contrary, theories are seldom if ever replaced whole and entire. To cite a well-known case, Copernican theory retained much of the Ptolemaic scheme, including all the heavenly bodies and the same celestial mechanics, only slightly fewer epicylces and eccentrics. [20] As the new wave insists for reduction, so too for replacement -- it is a matter of degree.

Fifth, and finally, other models of reduction offer a similar benefit, for example, the views developed by Schaffner and Kim, since they also traffic in a corrected theory T_R* that can better the explanatory role of the original T_R, thus establishing T_R's displaceability. [21] In fact, they offer a more convincing displacement if that phenomenon most plausibly occurs when T_R and T_R* are intra-level competitors, or when the

replacement is partial rather than wholesale. For, unlike the new wave, these other views allow T_R* to be specified by and thus retain the conceptual resources of an original higher-level T_R.

IV. REDUCING THE FALSE BY THE TRUE

So consider a second virtue. Proponents of the new wave claim their model can account for the otherwise puzzling case where the original reduced theory is false. Compare the classical view. If T_R is deduced from T_B, and if T_R is false, then it seems to follow that T_B must be false. As Churchland explains:

> Difficulties with this view begin to emerge with the observation that most reduced theories turn out to be, strictly speaking and in a variety of respects, <u>false</u>. (Real gases don't really obey PV = μRT, as in classical thermodynamics; the planets don't really move in ellipses, as in Keplerian astronomy; the acceleration of falling bodies isn't really uniform, as in Galilean dynamics, etc.) If reduction is deduction, modus tollens would thus require that the premises of the reducing theories (statistical thermodynamics in the first case, Newtonian dynamics in the second and third) be somehow false as well, contrary to their assumed truth. [22]

On the new wave model, however, one deduces a corrected version T_R* which can be sufficiently unlike the original and incorrect T_R so as to yield a difference in truth value. In Churchland's words: "on the liberal account we are here embracing it is clear that a true theory may reduce a false one." [23]

But matters are far from clear. First, and again, there is nothing particularly "new wavish" about the present virtue, since it has nothing to do with the novel constraint on theory construction. Rather, it concerns the condition on <u>deduction</u>, that T_R* and not the incorrect T_R be derived from T_B. Accordingly, Schaffner's account yields the same result, as does Kim's and any other view that deduces a corrected, approximate, fragmented, or domain-specific counterpart whose truth value can differ from the original.

Even the classical model may yield the same result, since the original T_R is not deduced from T_B alone, but from T_B in conjunction with the pertinent connecting principles and boundary conditions. Any falsehood implied by the derivation could then be located in these other elements, not in the basic T_B (hypothetical connections and

idealized conditions still permit valid deduction, their usefulness, perhaps, being judged by the closeness of the adjoining domains, the amount of falsehood in the reduced theory, and the like). [24] Churchland mentions this type of response, but adds:

> This defense will not deal with all cases of falsity, however, since in some cases the reduced theory is so radically false that some or all of its ontology must be rejected entirely, and the "correspondence rules" connecting that ontology to the newer ontology therefore display a problematic status. [25]

This is not convincing, however; for if the original theory is "radically false," then by the classical model and well-established usage it is simply not a target for reduction. In the old parlance, theories of this kind are subject to "replacement," where this contrasts with reduction in its definitive role as the provider of ontological and explanatory unification. Put in a different way, since radically false theories do not even approximate the ontology of the supposedly true reducing theory, there will be no correspondence rules to affect the classical derivation and thus unify the two domains. Contra Churchland's final remark, the rules do not then "display a problematic status." No, in radically false cases they simply do not exist! So the point remains -- at least for moderately false cases the classical derivation goes through, the false by the true, hand-in-hand with the necessary <u>façon de parler</u>.

Finally, even granting that the classical theory should accord T_R and T_B the same truth value, this may not be an untoward result. For suppose a scientific theory is counted true only when all constituent propositions are true, and false when at least one proposition is false (a trivial result on the assumption that a theory is defined as the logical conjunction of all constituent propositions). If so, then the reduced and reducing theories will have the same truth value -- both will be false. That is, barring the philosopher's ideal science or God's omniscient perspective, every basic reducing theory is likely to contain at least one false proposition. Indeed, past history will testify to the same, that basic reducing theories are false, strictly speaking, and including the very examples Churchland cites, like the now outdated Newtonian dynamics reducing Keplerian and Galilean theories. Hence the desire to reduce the false by the true will go unfulfilled. Error is the way of the world. [26]

V. THE EXPLANATORY BURDEN OF CONNECTING PRINCIPLES

So consider a third supposed virtue. John Bickle claims that, given the new wave model, there is no need to explain the status of bridge laws between reduced and reducing theories. Referring to the deduction of T_R* rather than the original T_R Bickle says:

> This contrast has a key consequence: it eliminates what was a vexing problem for the orthodox empiricist account of reduction, namely that of specifying the logical and ontological status of the bridge laws, especially in reductions that imply or reflect significant ontological change (op. cit., p.58).

Bickle repeats the claim elsewhere, saying: "One advantage of the H-C [Hooker-Churchland] account is that it avoids having to specify the logical status of cross-theoretic identity statements, a problem that becomes especially pressing for the bridge-law account when a reduction entails significant correction to the reduced theory." [27]

Yet Bickle's remark that the problem is especially acute for cases that reflect "significant ontological change" echoes Churchland's confusion about the status of correspondence rules when the original theory is radically false. On the classical view, significant ontological change calls for theory replacement, not reduction, meaning that there will be no system of bridge laws to affect the classical derivation, and hence no question about their logical or ontological status. [28] Moreover, when there is no significant change, then the status of the bridge laws is infamously clear -- they are identity statements, the considered opinion came to be, an opinion originally offered to solve (among other things) the very problem at issue. [29] Surely the new wave cannot complain, since their model is also committed to cross-theoretic identities as a consequence of relatively smooth cases at the retentive end of the reductive spectrum.

But the important point, for purposes of explaining bridge laws, is that the tables can be turned. For a defender of the new wave inherits a weightier burden on grounds that he or she has <u>more</u> inter-theoretic principles to explain, not just those implicated by ontologically retentive cases, but an inter-theoretic mapping across the entire spectrum from smooth reduction to bumpy replacement. That is, the pertinent new wave mapping also encompasses "correspondence rules" or "ordered pairs" in rough and disanalogous

cases, their function, again, being "to indicate which term substitutions in the image T_R* will yield [mimic or better?] the principles of T_R." [30] Nor is this inter-theoretic mapping entirely unproblematic, seeing that, in nonretentive cases, T_B (and thus its subset T_R*) are often radically incommensurate with respect to the original T_R. How, then, can radically incommensurate terms be mapped upon or substituted one for another? Is this an idealization, the classical façon de parler ? [31] The alleged sins of the old are repeated tenfold by the new.

VI. CO-EVOLUTION & NEW WAVE CONSTRUCTION

I have examined three central new wave virtues and found them either too common or too objectionable. Another acclaimed virtue will be discussed shortly. But I now want to present some positive arguments against. The first draws upon William Wimsatt's observations concerning "co-evolutionary development" between reduced and reducing theories, a doctrine which members of the new wave wholeheartedly embrace. As Patricia Churchland describes it:

> [T]heories at distinct theoretical levels often co-evolve [citing Wimsatt], as each informs and corrects the other, and if a theory at one stage of its history cannot reduce a likely candidate at a higher level, it may grow and mature so that eventually it does succeed in the reductive goal. In the meantime the discoveries and problems of each theory may suggest modifications, developments, and experiments for the other, and thus the two evolve towards a reductive consummation. [32]

As Hooker and Bickle also emphasize, this co-evolutionary feedback is not unidirectional, or exclusively from low to high levels. [33] There is a top-down flow of information and influence. For example, 'cistron' is a corrected image of the Mendelian gene (a term in T_R*, and hence a term supposedly formulated within the idiom of T_B). Yet it was not created from molecular genetics (T_B) ex nihilo, but from the pressure of the original Mendelian theory (T_R) to find a structure with the function of a gene. So co-evolved terms within T_B or rather its subset T_R* are by their very nature dually constrained by the rationales and conceptual resources grounded at both levels. In a word,

they are theoretical hybrids, mirror images of the inter-theoretic correspondence rules within classical reduction, differing from them only by the cover of a single term. The moral is that, because of the natural ebb and flow between levels of scientific language and scheme, the basic reducing theory becomes permeated with high level concepts and concerns.

Now the problem is straightforward. The new wave constraint on theory construction stipulates that the basic T_B and not the original T_R must supply the conceptual resources for constructing the corrected image T_R*. Yet this seems flatly contradicted by the fact that, once co-evolution has run its natural course, T_R* has become a mutual product of T_B and T_R. How, then, is T_R* specified "within the idiom of T_B" in any meaningful sense that excludes T_R?

The answer is not readily apparent. But it is clear what form the answer must take. Namely, the new wave advocate must discount the historical contribution made by the upper-level theory. But such a move will not register any deep semantic fact about the terms in question if historical properties carry some weight, as they do (directly or indirectly) by considerations about the social surround, remote baptismal groundings, linguistic divisions of labor, externalism about conventions, and naturally selected biofunctions. [34]

Moreover, the new wave cannot appeal to conceptual or inferential role as a way of partitioning off the contribution of the original upper-level theory, [35] since the model guarantees current inferential links between T_R* and T_R, those secured by the inter-theoretic mapping in the second stage of the new wave reductive process, both inferences that yield identity and substitutions that indicate mimicry. Also, a special class of these inferential links will block the otherwise quite natural suggestion that T_R* counts within the idiom of T_B, exclusively, on grounds that it can be deduced from T_B. For, likewise, T_R* can be deduced from T_R in cases of perfect retention, since they are for all intents and purposes the very same theory. Similarly, T_R* can be deduced from T_R in less than perfect but still strongly analogous cases, only now in conjunction with more generous boundary conditions, hypothetical assumptions, and convenient fictions (remember how the classical theory can deduce the false by the true).

VII. HISTORY, METHODOLOGY, & THE INITIAL COLLAPSE

But suppose there is a way to reconcile new wave theory construction with co-evolutionary development, which is to say, some plausible account of term-to-theory individuation which grants T_B sole propriety over T_R*. Even so, the distinction between constructing T_R* out of T_B versus T_R can only appear superficial from the perspective of the actual historical development of the sciences, seeing that cross-theoretic evolution guarantees a healthy interplay between the two levels. More important, new wave theory construction can only appear unduly restrictive from the vantage point of reductionist ideology and methodology, since it rules out strategies that may require aid from above, from the level of concepts in T_R.

Consider local reduction, whereby a corrected T_R* can be generated out of an original and more general T_R either by "relativizing" the original types to a restricted domain or by viewing those general types as "fragmenting" into narrower ones. [36] Either way, it is an operation upon original higher-level types. The new concepts thus owe their existence to the old, if not entirely, at least in conjunction with familiar methodological pressures from the basic reducing theory (human pain is a modification of pain, constructed directly out of general psychology, but with an eye to finding some stable nonvariable structure within neuroscience).

The upshot is this: on the worst-case scenario new wave construction is flatly contradicted by co-evolutionary facts; on the best-case scenario it is historically shallow and methodologically restrictive. Sound reason, therefore, enjoins that it must be rejected. Once rejected, however, the new wave picture takes on a familiar look; for the corrected T_R* may now be specified within the idiom of T_R, making the deduction of T_R* inter-theoretic in nature, and bringing in train all the hybrid correspondence rules that a corrected classical derivation must provide. More pointedly, since the constraint on theory construction is the only genuinely novel element in the new wave model, and that the three remaining conditions are either directly implied or easily generated from Schaffner's account, then the entire project will collapse into its predecessor's. 1980s new wave reduction and 1960s revolutionary approximation are one.

VIII. BRIDGE LAWS & THE FINAL COLLAPSE

Yet surely the most important of all virtues is the capacity of the model to resist anti-reductionist criticism. To that end all defenders of the new wave unite in holding that their view enables them to counter arguments which originally led to the demise of the classical theory. Without this all would be for naught. Qua reductionism, the entire project would fail.

So consider those arguments which reject psychological and function-to-structure bridge laws, chiefly Hilary Putnam and Jerry Fodor's multiple realizability argument, and also Donald Davidson's normativity argument, since they brought about widespread anti-reductionist sentiment with regard to psychology and the functionally specified domains of biology. [37]

Now observe the following new wave strategy. Responding to Davidsonian arguments, John Bickle claims: "the impossibility of psychophysical laws is irrelevant to the new thesis of mind-brain reductionism and the novel account of intertheoretic reduction underwriting it." [38] Why? Bickle asks and then answers the question:

> Having now presented the account of intertheoretic reduction adopted by proponents of the new mind-brain reductionism, what can we conclude about the Davidsonian objection based upon the impossibility of psychophysical laws? It fails, and quite conclusively. For the lack of cross-theoretic laws is of no consequence to whether an H-C reduction is possible, <u>since an H-C reduction nowhere requires bridge laws</u>. [39]

Generalizing, then, whether it be Davidson's normativity considerations or Putnam and Fodor's more influential multiple realizability argument, all fail because they challenge reductionism on the point of bridge laws between the reduced and reducing theories -- a palpable hit when directed against the classical account, but one that falls wide of the mark when directed against the new wave. For "H-C reduction nowhere requires bridge laws."

But this is mistaken. For new wave reduction includes an inter-theoretic mapping in its second stage, which, when carried out for strongly analogous cases, justifies cross-theoretic identities. As Churchland says: "a successful reduction of the ideal sort described provides an excellent reason for asserting the relevant cross-theoretic identities,

the best reason one can have." [40] Yet property identity guarantees nomic coextension. So bridge laws exist within the new wave account, being directly implied by retentive cases.

Nor will it do, as a response, to insist on a distinction between the "reduction proper" versus its "consequences," confining bridge laws to the latter. [41] Consequences are consequences, and to deny them is like a smuggler caught in the act whose only defense is: "I meant there was no contraband <u>on my person</u> !" Declared or no, up front or trailing behind in tow, the goods are there. And so the problem is clear:

> (i) If a case falls at the retentive end of the new wave continuum, then cross-theoretic property identities exist between reduced and reducing theories.
> (ii) If cross-theoretic property identities exist between reduced and reducing theories, then biconditional bridge laws exist between reduced and reducing theories.
> (iii) Therefore if a case falls at the retentive end of the new wave continuum, then biconditional bridge laws exist between reduced and reducing theories.

The conclusion is inescapable: the retentive end of the new wave spectrum will collapse into the classical account of reduction, replete with its inter-theoretic bridge laws, and subject to all the slings and arrows of anti-reductionist argument against them. Moreover, if some of these arguments are sound, then the new wave is forced into a familiar position. [42] That is, to return full circle:

> (iv) It is <u>not</u> the case that biconditional bridge laws exist between intentional psychology or functionally construed theories in biology and more basic physical theories.
> (v) Therefore it is <u>not</u> the case that intentional psychology or functionally construed theories in biology will fall at the retentive end of the new wave continuum.

There is no type reduction. Rather, a defender of the new wave must locate all such theories at the replacement end of the new wave continuum. [43] Surprisingly, then, matters remain much as they were before the new wave arrived upon the shore. In the absence of some other account, one must either accept anti-reductionism or embrace type eliminativism.

IX. CONCLUDING POSTCRIPT

The astute reader will have gathered that the new wave model has collapsed on several fronts. By rejecting its constraint on theory construction, the model collapses into approximate reduction. By observing bridge laws implied by its inter-theoretic mapping conjoined with their ontological consequences, the retentive end of the model collapses into classical reduction. And by accepting standard arguments against those bridge laws, the retentive end of the model collapses, period.

Let me add, in conclusion, a few remarks about certain newer, new wave proposals that might be deemed relevant. In particular, Churchland no longer speaks of theories in terms of sentences or sets thereof, but rather as vectors through connectionist phase space. [44] And Bickle has recently moved to a nonsentential account of theories understood in terms of model-theoretic structures. [45] Therefore one might claim that these *non*sentential accounts will enable the new wave to avoid inter-theoretic correspondence rules and thus burn the bridge with classical reduction. [46] But not so. Traditional correspondence rules are not avoided, only relocated.

No one wishes to deny the existence of public language sentences, or their use by scientists when announcing, demonstrating, and publishing their theories (whatever the ultimate nature of theories might be). Accordingly, William Bechtel proposes what amounts to a two-factor theory, supplementing internal neurocomputation with publicly manifested deductions and diagrams and "various external representational systems." [47] Paul Churchland, too, accommodates these plain facts of institutional science, though he now accords them a secondary role within a broader scheme of theoretical activity that encompasses animals, protohumans, children, and nonscientific lay persons. [48]

In fact, the public domain of sententialist epistemology seems necessary for the preservation of the new wave model, given certain assumptions often associated with those like Churchland who are both eliminativists and connectionists about the mind. Specifically, in light of such views, the condition that T_R* be deduced from T_B cannot be internalized within neurocomputational-level systems, since prototype activation in a

connectionist net does not involve the sententially-inspired rule-based deduction (ironically, traditional syntactic-driven architectures fare better on this score).

Worse still, there seems to be no plausible neurocomputational corollary for the one novel new wave constraint that T_R* be constructed out of the idiom of T_B rather than the original T_R. Would it mean that, as a matter of brute fact, all human brains are so constituted that T_R*'s activation vector is always a subsection of T_B's, and that T_R's activation vector never overlaps either on T_B or its subsection T_R*? That should depend entirely upon how the theories become calibrated within each individual, which is to say, upon the vicissitudes of the several input histories cum variable connections and weights.

So, belabored but important, the public language of science remains. Yet, also, advocates of the new wave have not retreated from the position that their model delivers important ontological consequences, namely, property identities in strongly analogous cases. Hence the problem surfaces again. Property identities conjoined with the public language expressions of the hypothesized nonsentential vector spaces/sets of models T_R and T_B (or its subset T_R*) will yield nomic coextension for the predicates used in those public language expressions. The result is a set of intertheoretic bridge laws, and such is the classical model, in full sententialist dress. From this newest perspective, old bridging principles emerge at the meta-level, within public language descriptions which all new wavers employ and to which all new wavers are unavoidably committed (whereof one must speak, one cannot remain silent).

To end with a fitting metaphor, the "collapse of the wave function" in quantum mechanics refers to the fact that a particular wave can be treated in a distinctly un-wave-like fashion, in terms of classical physics as a particle with specific location and velocity. Similarly, here the "collapse of the new wave" refers to the fact that a particular new-wave proposal can be treated in a distinctly un-wave-like fashion, in terms of classical reduction with bridge laws between reduced and reducing theories.

NOTES

* My thanks to Charles Carr, Jerry Fodor, Thomas Grimes, Terence Horgan, Jaegwon Kim, John Post, and Kent Staley for their helpful comments on an earlier draft of this paper. I owe John Bickle a special debt of gratitude for many discussions and friendly disagreements about the new wave.

1 John Bickle (1996, p.57).

2 The model and its virtues are outlined in Churchland (1979, 1985) and Hooker (1981).

3 Churchland (1979, p.81, with a slight change in the variables, henceforth). It should be noted that Churchland no longer speaks of theories in terms of public language sentences, but prefers connectionist phase spaces. More on this in section IX.

4 Churchland (1985, p.10).

5 Hooker (1981, Part I, p.49).

6 Churchland (1979, p.84).

7 Churchland (1985, p.11). He then defines "property" reduction as follows: "A property F, postulated by the older theory or conceptual framework T_R, is reduced to a property G in some new theory T_B just in case (1) T_B reduces T_R; (2) F and G are correspondence-rule paired in the reduction; and (3) the reduction is sufficiently smooth to sustain the ontology of T_R, and thus to sustain the identity claim, 'F-ness = G-ness' (loc. cit.)."

8 Patricia Churchland (1986, p.284).

9 Bickle (1992a, p.417).

10 Kim (1972, p.190, 1989, 1992) and Lewis (1969, pp.24-25, 1980).

11 Schaffner (1967). See also Wimsatt (1976, p.217).

12 Hooker (1981, Part I, p.45). The rationale for Hooker is that reduction must still achieve some measure of explanatory and ontological unification.

13 Patricia Churchland (1986, p.311). See also her remarks about what the reducing theory must explain (ibid., p.283), as well as the distinction between "bumpy reduction" and "outright displacement" (p.284), all of which presuppose the well-established usage in question. Yet compare Paul Churchland (1979, p.84) and Bickle (1996, pp.64-5), who describe nonborderline, nonretentive cases as "bumpy reductions." By standard usage, however, calling outright replacement a "bumpy reduction" is slightly perverse -- like calling one's divorced status a "bumpy marriage."

14 I say "might have been" since Schaffner does not require that T_R* be constructed out of the resources of a higher-level T_R. Instead, his model is defined to include "homogeneous" cases (1967, p.144).

15 Paul Churchland (1985, p.11). In his earlier work, similar remarks were followed by the disclaimer: "Displacement, of course, need not actually take place, however much considerations of unity and simplicity might demand it. Familiarity, entrenchment, convenience, and continuity may together counsel a less puritan course" (1979, p.82). So it is a form of displacement <u>in principle</u>. Even so, Churchland implies that it <u>can</u> take place, and certainly "familiarity, entrenchment, convenience, and continuity" do not accord any explanatory privilege to the original reduced theory. On the contrary, in

ideally smooth cases the corrected theory "will perform all the same predictive and explanatory functions" of the original (loc. cit.), and in less than ideal cases it will be "superior" (ibid., p.83).

16 Curiously, Bickle (1996, pp.64-7) chastises Fodor for claiming that reduction will eliminate (i.e., displace) the special science theory, failing to note that Churchland made this an explicit platform of the new wave model. Nor is Bickle's plea for the preservation of special sciences based upon any essential explanatory function that would be lost in the displacement. Instead, like Churchland's concession to the impure demands of "familiarity, entrenchment, convenience, and continuity" (see fn.15 above), Bickle is content to say only that: "Reducing theories are typically more general in scope than reduced ones, and so are often much more unwieldy for use in actual scientific practice" (ibid., pp.66-7). There remains, then, a substantive disagreement -- the new wave believes that sciences like psychology can be displaced without loss of explanatory power, retained only for the sake of convenience, while anti-reductionists deny it.

17 Hooker (1981, Part II, p.203), Specifically, he says: "where perfect retention fails $T_R^* \neq T_R$" (loc. cit.). Note, too, though Hooker conjectures elsewhere that "the retention extreme of the retention/replacement continuum goes unoccupied" (Part I, p.45), that is a contingent matter, depending upon how actual theories may fare. So the model itself does not guarantee displacement, not in principle, not at the extreme endpoint of retention where $T_R = T_R^*$.

18 See Wimsatt (1976, pp.215, 222-23) and McCauley (1986, 1996).

19 That one must take the new wave constraint in this exclusionary sense is clear, otherwise the deduction of T_R^* from T_B would not be intra-theoretic in nature, specifically, in derivations where elements of a higher-level T_R remain in T_R^*.

20 This is Copernicus' considered view, found in the later books of the De Revolutionibus. See Thomas Kuhn (1957). Kent Staley has pointed out to me that the historical fact of partial replacement was emphasized long ago by William Whewell (1851).

21 Accordingly, I have argued that Kim's model is best viewed as a form of eliminativism via the original higher-level theory. See Endicott (1993, pp. 306-10).

22 Churchland (1985, p.9). Note that Churchland comes dangerously close to a fallacy of equivocation -- modus tollens needs "truth," Newtonian dynamics needs "assumed truth" (or "assumed truth at the time proposed," since it is now known to make false predictions when velocities near the speed of light, etc.).

23 Churchland (1979, p.84). Jerry Fodor has pointed out to me in corrspondence that whether the model allows T_R to be false while T_B is true depends entirely on the (unstated) constraints which govern the inter-theoretic mapping. E.g., if biconditional bridge laws are in the offing (as they will be at the retentive end of the continuum, see sec. VIII), then the same trouble affects the new wave account.

24 This only forces one to reconstrue the nature of the deduction as a kind of "transformation" versus a "sound argument" with true premises. See Nickels (1973).

25 Churchland (1985, pp.9-10). The new wave may complain that the classical model thus has very little it can subsume. See Churchland (1979, p.84). But two points should

be made. First, this is a different issue. The classical account can reduce the false by the true, only not for an (allegedly) acceptable range of cases. Second, one might challenge the assumption that the classical model does not subsume an acceptable range of cases. Often "textbook" cases of reduction turn out, upon reflection, to be cases better described as replacement.

26 Terry Horgan has suggested to me that, while my arguments are technically correct, the spirit of "reducing the false by the true" might be better served by saying that new wave reduction can eliminate one source of error -- falsity in the higher-level theory -- even if there are other sources of error in the reducing theory. Yet Horgan notes that this advantage is not exclusive to new waveism, which is my central point. I add, too, deep reservations about the underlying epistemic principle. For the one domain is not inherently more error-ridden than the other. So, e.g., it is not like the (already suspect) empiricist scruple whereby one chooses "fallible observation" over "more fallible commitment to unobservables." And if there is no reason to favor low-level error, then impartiality would demand either that one eliminate all sources or retain both levels instrumentally without commitment to their truth.

27 Bickle (1992b, p.223). See also his (1992c, p.54).

28 What Bickle could argue is that the status of bridge laws is problematic when there are "minor" ontological changes. For in that case reduction is still in the offing, not replacement, and yet the classical resources of strict identity via bridge laws might be incompatible with these differences. Yet, in my view, such arguments ultimately conflate differences in meaning with differences at the level of reference and ontology (a difference in ontology versus a difference in description of the same ontology). In any case, the new wave is saddled with the same strict identities in ontologically retentive cases. See sec. VIII.

29 See Lawrence Sklar (1967); Robert Causey (1977, chaps.4 and 5); and a nice summary in Kim (1996, pp.212-16).

30 Churchland (1985, p.10). There is no question that Churchland means to include rough and disanalogous cases. For he just cited the "radically false" cases that make any deduction of the original problematic (1985, pp.10-11), and says "what cases like these invite us to give up is the idea that what gets deduced in a reduction is the theory to be reduced" (ibid., p.10). Hence he presents the intra-theoretic deduction, and then the inter-theoretic mapping with the pertinent remarks quoted in the text.

31 Churchland also misses the problem. He says "it is possible on this account for a theory to reduce even an incommensurable competitor," and discusses the reduction of classical mechanics "CM" by the special theory of relativity "STR" (1985, p.85). It is possible, he says, in spite of the difference in how mass is conceived, because "it is easy to deduce within STR ... a highly convincing image of CM" (loc. cit.), which is to say, because of the first stage of the new wave reductive process where one avoids deducing the incommensurable CM. But this fails to address the "mapping" or "substitution" in the second stage between that same allegedly incommensurable CM and the base-level counterpart CM* constructed entirely from STR.

32 Patricia Churchland (1986, p.264). The original source is Wimsatt (1976, pp.230-37). See also Hooker (1981, Part I, p.48). However, contra Churchland, there is no reason to

assume the theories will always co-evolve towards "reductive consummation." They may drift further apart or settle into a stable equilibrium. See McCauley (1996).

33 See Hooker (1981, Part III, pp.513-14); Bickle (1996, p.76).

34 See, e.g., Kripke (1972), Putnam (1975), and Millikan (1984). In fact, aside from current conceptual role, there is probably just one leading semantic theory that ignores historical factors -- a covariational theory that depends upon counterfactual relations and not actual causes. See Fodor (1994, pp.115-19). But even this is no ally to the present partitioning strategy mentioned in the text, since, for all cases at the retentive end of the new wave continuum, the properties counterfactually related to T_R and T_R* will be the same, and thus the terms will have identical semantic content.

35 See Harman (1974) and Field (1977). Paul Churchland has been committed to something like this account when speaking with the vulgar (e.g., 1979, pp. 52-80). For his later state-space semantics, see Churchland (1986). For detractors, see Fodor and Lepore (1992, 1996).

36 See Causey (1977, pp.147-49); Kim (1989, p.273).

37 For the multiple realizability argument, see Putnam (1967), Fodor (1974), and my summary in Endicott (1996). For the normativity argument, see Davidson (1985, esp. pp.245,249) and Kim (1985).

38 Bickle (1992b, p.218).

39 Bickle (ibid., p.224, italics mine; see also 1996, pp.58-59).

40 Churchland (1979, p.83; also 1985, p.11; Hooker, 1981, Part I, pp.45-6, Part II).

41 Bickle has suggested something like this in correspondence, telling me: "I only acknowledge [new wave conditions] (i) and (ii) as part of the intertheoretic reduction relation proper. (iii) and (iv) are attempts to relate intertheoretic results to ontological conclusions." And, granted, Churchland also speaks of a "reduction proper." Still, Churchland did include the intertheoretic mapping and its ontological conclusions in the reduction (1979, p.81; 1985, p.11, etc.). Indeed, he counts "reducibility as a matter of degree," which can only mean the inter-theoretic relations of conditions (iii) and (iv), surely not the intra-theoretic deduction of (i) and (ii). And, regardless, if we have a reduction of the original T_R at all, it must be tied to the relation between T_R and its image T_R* which conditions (ii) and (iv) address, otherwise the deduction of T_R* from T_B would be completely irrelevant to the status of T_R.

42 For a defense of the Putnam-Fodor genre of arguments, see Endicott (1989, 1993), Horgan (1993, forthcoming), and Kincaid (1990).

43 Of course this is not to deny other cases of smooth reduction where type-identity is preserved, e.g., physical optics, Kepler's laws of planetary motion, or a domain-restricted reduction thermodynamics. Also, my remarks concern the retention versus replacement of types. I have not addressed Hooker's function-to-structure token reductions (1981, Part III, pp.504-5; also Bickle (1992c, pp.55-56). I feel there is no pressing need to address this issue here, since the historical debate over reduction has always concerned types and not tokens, the antireductionists being the (mere) token physicalists.

[44] The deep reason is that Churchland has moved from a deductive nomological account of explanation to a psychological theory of prototype understanding grounded in the neurocomputational framework. See his (1989) and (1990). Similarly, Hooker has moved to a "regulatory systems" view of science which complements the connectionist paradigm. See Hooker et. al (1992) and Hooker (1995).

[45] See Bickle (forthcoming); but cf. Churchland's (1990) less than enthusiastic view of the model-theoretic approach to the nature of theories.

[46] So Bickle tells me that "no bridge laws or correspondence rules exist anywhere in my [forthcoming Psychoneural Reduction, not the earlier] new wave account of intertheoretic reduction. How could they? Bridge laws and correspondence rules are sentential items, and my account of theory structure and intertheoretic relations nowhere employs sentences. Theories are sets of models with a certain set-theoretic structure." And Bickle is right. If narrowly focused on the models and their set-theoretic conditions, then no sentences can be found. But, as I argue in the text, taking a wider view of scientists and new wavers discoursing about the nonsentential items, coupled with new wave ontological commitments for those items in retentive cases, reveals and old and familiar, i.e., classical story.

[47] William Bechtel (1996, p.126).

[48] Churchland (1996, pp.266-7).

BIBLIOGRAPHY

Bechtel, W. 1996: "What Should a Connectionist Philosophy of Science Look Like?," in Robert McCauley, ed., The Churchlands and Their Critics. Oxford: Blackwell, 1996, pp.121-44.

Bickle, J. 1992a: "Revisionary Physicalism," Biology and Philosophy 7, 411-30.

_____ 1992b: "Mental Anomaly and the New Mind-Brain Reduction," Philosophy of Science 59: 217-30.

_____ 1992c: "Multiple Realizability and Psychophysical Reduction," Behavior and Philosophy 20: 47-58.

_____ 1996: "New Wave Psychophysical Reduction and the Methodological Caveats," Philosophy and Phenomenological Research 56 , 57-78

_____ forthcoming: Psychoneural Reduction: The New Wave. Cambridge MA: MIT Press.

Causey, R. 1977: Unity of Science. Dordrecht, Holland: D. Reidel Publishing.

Churchland, P.M. 1979: Scientific Realism and the Plasticity of Mind. London: Cambridge University Press.

_____ 1985: "Reduction, Qualia, and the Direct Introspection of Brain States," Journal of Philosophy 82: 8-28.

_____ 1986: "Some Reductive Strategies in Cognitive Neurobiology," rpt. in A Neurocomputational Perspective: The Nature of Mind and the Structure of Science. Cambridge MA: MIT Press, 77-110.

_____ 1989: "On the Nature of Explanation: A PDP Approach," rpt. in Neurocomputational Perspective, 197-230.

_____ 1990: “On the Nature of Theories: A Neurocomputational Perspective,” rpt. in Neurocomputational Perspective, pp.153-96.
_____ 1996: “Bechtel on the Proper Form of a Connectionist Philosophy of Science,” The Churchlands and Their Critics, pp.265-70.
Churchland, P.S. 1986: Neurophilosophy: Toward a Unified Science of the Mind/Brain. Cambridge MA: M.I.T. Press.
Davidson, D. 1985: “Replies,” in B Vermazen and M. Hintikka, eds., Essays on Davidson: Actions and Events. NY: Oxford.
Endicott, R. 1989: "On Physical Multiple Realization," Pacific Philosophical Quarterly 70: 212-24.
_____ 1993: “Species-Specific Properties and More Narrow Reductive Strategies,” Erkenntnis 38: 303-21.
_____ 1996: "Multiple Realization," Encyclopedia of Philosophy, Supplement, 365-66.
Field, H. 1977: “Logic, Meaning, and Conceptual Role,” Journal of Philosophy 74: 379-409
Fodor, J. 1974: "Special Sciences," rpt. in Representations: Philosophical Essays on the Foundations of Cognitive Science. Cambridge MA: M.I.T. Press, 1981, 127-45.
_____ 1994: The Elm and The Expert: Mentalese and its Semantics. Cambridge MA: MIT Press.
Fodor, J. and Lepore, E. 1992: Holism: A Shopper’s Guide. Oxford: Blackwell.
_____ 1996: “Paul Churchland and State Space Semantics,” in McCauley, ed., The Churchlands and Their Critics. Oxford: Blackwell, 145-62.
Harman, G. 1974: “Meaning and Semantics,” in Milton Munitz and Peter Unger, eds., Semantics and Philosophy. N.Y.: New York University Press, 1-16.
Hooker, C. 1981: “Towards a General Theory of Reduction. Part I: Historical and Scientific Setting. Part II: Identity in Reduction. Part III: Cross-Categorical Reduction,” Dialogue 20: 38-59, 201-236, 496-529.
_____ 1995: Reason, Regulation, and Realism: Toward a Regulatory Systems Theory of Reason and Evolutionary Epistemology. Albany, NY: SUNY Press.
Hooker, C., Penfold, H., and Evans, R., 1992: “Control, Connectionism, and Cognition: Toward a New Regulatory Paradigm,” British Journal for the Philosophy of Science 43: 517-36
Horgan, T. 1993: “Nonreductive Materialism and the Explanatory Autonomy of Psychology,” in S. Wagner and R. Warner, eds., Naturalism: A Critical Appraisal. Notre Dame, IN: University of Notre Dame Press, 295-320.
_____ forthcoming: “Multiple Reference, Multiple Realization, and the Reduction of Mind,” in F. Siebert and B. Preyer, eds, Reality and Humean Supervenience: Essays on the Philosophy of David Lewis.
Kim, J. 1972: “Phenomenal Properties, Psychophysical Laws, and the Identity Theory,” Monist 56: 177-92.
_____ 1985: “Psychophysical Laws,” in E. Lepore & B. McLaughlin, eds., Actions and Events: Perspectives on the Philosophy of Donald Davidson. Oxford: Basil Blackwell, 369-86.
_____ 1992: “Multiple Realization and the Metaphysics of Reduction,” rpt. in Supervenience and Mind. London: Cambridge University Press, 1993, 309-35.
_____ 1996: Philosophy of Mind. Boulder, CO: Westview Press.

Kincaid, H. 1990: “Molecular Biology and the Unity of Science,” Philosophy of Science 57: 575-93.
Kripke, S. 1972: Naming and Necessity. Cambridge MA: Harvard University Press, 1980.
Kuhn, T. 1957: The Copernican Revolution: Planetary Astronomy in the Development of Western Thought. Cambridge MA: Harvard University Press.
Lewis, D. 1969: “Review of *Art, Mind, and Religion*,” Journal of Philosophy 66: 22-35.
_____ 1980: “Mad Pain and Martian Pain,” in Ned Block, ed., Readings in Philosophy of Psychology, vol.1. Cambridge MA: Harvard University Press, 216-22.
McCauley, R. 1986: “Intertheoretic Relations and the Future of Folk Psychology,” Philosophy of Science 53: 179-99.
_____ 1996: “Explanatory Pluralism and the Co-evolution of Theories in Science,” in McCauley, ed., The Churchlands and Their Critics, 17-47.
Millikan, R. 1984: Language, Thought, and Other Biological Categories. Cambridge MA: MIT Press.
Nickles, T. 1973: “Two Concepts of Inter-Theoretic Reduction,” Journal of Philosophy 70: 181-201.
Putnam, H. 1967: “The Nature of Mental States,” rpt. in Mind, Language and Reality: Philosophical Papers, vol.2. London: Cambridge University Press, 1975, 429-40.
_____ 1975: “The Meaning of ‘Meaning’,” in Mind, Language and Reality, 215-71.
Schaffner, K. 1967: "Approaches to Reduction," Philosophy of Science 34, pp.137-47.
Sklar, L. 1967: "Types of Inter-Theoretic Reduction," British Journal for the Philosophy of Science 18: 109-24.
Whewell, W. 1851: “Of the Transformation of Hypotheses in the History of Science,” Transactions of the Cambridge Philosophical Society 9: 139-47, rpt. in Robert E. Butts, ed., William Whewell: Theory of Scientific Method. Indianapolis, IN: Hackett Publishing Co.,1989, 251-62.
Wimsatt, W. 1976: "Reductionism, Levels of Organization, and the Mind-Body Problem," in G. Globus, G. Maxwell, and I. Savodnik, eds., Consciousness and the Brain: A Scientific and Philosophical Inquiry. New York: Plenum, 199-267.