1

Conflict in the Abstract Sciences

> How can a philosophical enquiry be conducted without a perpetual *petitio principii*?
>
> Frank Ramsey, *The Foundations of Mathematics*, 1931

CONFLICT RESOLUTION

1905 was an intellectually eventful year. It saw the birth of Russell's "On Denoting" and Einstein's special theory of relativity, to say nothing of the founding of the Bloomsbury Group and the appearance of *The Psychopathology of Everyday Life* and the Binet Test. Relativity theory was attended by conflict right from the beginning, and barely a year passed before disconforming experimental evidence was unearthed.[1] In one of the century's more alluring examples of a theory's resistance of empirical discouragement, relativity hung on until, in 1914–16, it received experimental confirmation strong enough to annul the Kaufmann deviations.[2] While the new physics was awaiting empirical respectability, the foundations of geometry occasioned considerable contention. Frege and Hilbert saw things differently. They clashed over the nature and function of the geometric axioms. Frege saw the axioms as a reflections of conditions necessary for spatial experience, and so as synthetic propositions known a priori. For Hilbert, axioms are the theoretical constructions of the geometer, epistemically secure if consistent. On Hilbert's view, whether a geometric axiom strikes us as a priori true, or, for that matter, as a priori false, is a fact about us, not about geometry intrinsically. Axiom sets are consistent specifications of mathematically possible spaces, whose physical realization, or not, tells neither for nor against the axioms.

We have here two historically important cases of scientific disagreement in the twentieth century. Anyone interested in the dynamics of conflict resolution in the sciences will at once see the two cases as importantly different. The Einstein-Kaufmann conflict was eventually settled. The Frege-Hilbert conflict just went on and on, and ended without resolution, on Frege's death in 1925.

The conflict resolution theorist is bound to make something of this difference and to offer an account of it. On the face of it, he has not far to go for an answer. Relativity theory triumphed in the end on the strength of its *empirical adequacy*.[3] The dispute between Einstein and Kaufmann was settled by Nature. The intractability of the standoff between Frege and Hilbert is similarly explained, but in the opposite direction, so to speak. In this case, empirical adequacy was not an applicable or appropriate resolution device. Theirs was a dispute with regard to which Nature had nothing to offer.

In a rough and ready way, theories divide into those for which the criterion of empirical adequacy is a legitimate standard, if not always a fulfilled one, and those for which the standard is made inappropriate by subject matter and method. This distinction I mean to mark by saying that theories that are properly held to the condition of empirical adequacy are *empirical theories*, whereas those that are not are *abstract theories*.[4] Rough as it is, our present distinction is consequential in a way that I shall try to take the measure of. Empirical theories have inbuilt procedures for conflict resolution – as with the Einstein-Kaufmann dispute – however complex and indirect they may be. Collectively these mechanisms are a theory's empirical check. Abstract theories, such as the epistemology of geometry, lack these mechanisms for conflict resolution, and it is this that makes them methodologically interesting. Among empirical theorists there is a philosophically naive but utterly entrenched inclination to suppose that a theory's empirical check is also a *reality* check for it; that a theory is objectively right in its claims to the extent that it "checks out" empirically. Abstract theories lack an empirical theory's way of negotiating its reality check.[5] On the face of it, this matters. We are left to ask whether abstract theories have reality checks and, if so, what they are and how we come to recognize them. If not, how can the principles and laws of such theories count as true?

Some readers will not much like the putative dualism of the empirical and abstract. Perhaps these skeptics will have been persuaded by Quine's arguments, which for their influence and their artistry demand a certain tarrying over here. Quine is a radicalizer of Duhem's comparatively modest holism about physics. In Quine's hands, the confirmation due to *any* theory applies to it whole and entire rather than sentence by sentence. Confirmation goes global, attaching to individual sentences honorifically, in a mode of attribution that, save for the honorific, would be the ancient fallacy of division.

Mathematics is indispensable to science. Seen in Quine's way, mathematics is essential to a theory's implication of its observation categoricals. Observation categoricals are sentences such as "When it snows, it's cold." They are the "direct expression of inductive expectation," the first intimation of a theory's laws (Quine, 1995, p. 25). This should make us curious about whether its indispensability to theories having empirical checkpoints is sufficient to pass on the status of empirical to mathematics itself, as relativity theory was thought to do for Riemann's geometry. Quine is affirmatively minded.

He asks – rhetorically – whether there is any epistemological advantage in treating the mathematics of a globally confirmed theory differently from what its confirmation requires for the theory itself. Although mathematics lacks empirical *content*,[6] Quine finds no good reason to contrive, for scientifically useful mathematics, a separate epistemology. It is not just that mathematical epistemologies have had a bad track record (as witness, the unhappy careers of synthetic apriority and reductive analyticity); it is also a matter of methodological economics. Why should a scientific theory have two epistemologies – one for the empirical part, the other for its mathematical part – when one could be made to do across the board?

Quine also supposes that the same can be said for a theory's meaning. It is often said that the rejection of verificationism has long been a centerpiece of Quine's philosophy. Thinking so is a serious misapprehension. Quine is an unwavering verificationist. Meaningfulness is conferred by confirmation; not, as we see, sentence by sentence, but on whole theories. Thus, Quine's brand of verification encompasses what is sometimes called "semantic holism," and his complaint against Carnap and other positivists is a complaint not against the verificationism of their semantics but against its atomism; its supposed application to sentences one by one. What, then, of those individual sentences? Do they acquire their meaningfulness from the confirmation conferred on the theories in which they occur? If so, is the achievement of local meaningfulness also honorific, as we supposed in the case of local confirmation? If so, then semantic holism is a dislocater of classical logic. If a theory's sentences are meaningful one by one only in an honorific sense, then they are true or false only honorifically, too – which makes the Bivalence law of classical logic false. Not so for Quine, of course, who rejects any notion of meaning linked to the suggestion that the bivalence of a sentence requires it to have a propositional context. Still, on reflection, we might think better of honorificizing our inferences in *sensu diviso* and plump for more straightforward deductions. In the case of confirmation in isolation, we could say that a sentence is actually, not honorifically, confirmed by its membership in the set of derivations of a confirmed theory. In the logico-semantic case, we could likewise say that a sentence is actually, not honorifically, meaningful by its membership in the set of sentences of a meaningful theory. We would appear to be wrong each time. Equivocation looms. In its application to a theory, "confirmed" means something like "stands in such-and-so relation R to the available evidence," whereas in its application to sentences, "confirmed" means "is derivable in a confirmed theory." Thus it is a non sequitur – the fallacy of division – to infer the confirmation of sentences from the confirmation of the theories in which they are derived. That is to say, any such inference is the fallacy of division *if holism is true*. Whatever the relation R to which a confirmed theory stands to the available evidence, and to which it owes its confirmation, holism insists that it is not *that* relation that any of a confirmed theory's assertions bears to the available evidence. It is the same way with attributions of

meaningfulness. When applied to a theory, "meaningful" means "confirmable," that is, "could come to stand in relation R to evidence that becomes available." As applied to sentences, "meaningful" means "is asserted or denied by a theory that might bear R to evidence that becomes available," a relation in which if holism is true sentences cannot stand one by one. This leaves the semantic holist painfully positioned. He can have his truth-valued sentences either honorifically or actually, but at a cost either way. If honorifically, he must – short of Quine's semantic skepticism – reconcile himself to the loss of classical logic. If nonhonorifically, the price is worse; it is the fallacy of division.

Perhaps the dilemma could be slipped if the requisite ambiguities were noted. Then the inferences,

1. T is a confirmed theory
2. Φ is derivable in T
3. Therefore Φ is a confirmed sentence

and

a. T is a meaningful theory
b. Φ is a sentence of T
c. Therefore Φ is a meaningful sentence

would duck the charge of equivocation if the terminal "confirmed" expressed something different from the initial "confirmed," and likewise for "meaningful." But unless we have antecedent knowledge of the sense of these terminals, we shall not know what these inferences convey, never mind whether the conveyance is valid. We could venture that in line (a) "meaningful" means "verifiable," and suppose that in its recurrence in line (c) it means "has a truth value." There is something to be said for this line of thought, since even if "truth-valued" does not appear to follow from "meaningful," it may appear to follow from "verifiable," from which *on the verificationist account* "meaningful" itself follows. The transitivity of following from takes care of the rest. If this is our solution, it is consequential well beyond our interest in the derivation of (c) from (a) and (b). It gives us grounds for thinking that verificationism is not a theory about meaningfulness after all, or, to say the same thing more circumspectly, that it is an account of meaningfulness in a technical and neologistic sense of the term.

If we have found a way to reconcile ourselves to the validity of the derivation of (c) from (a), and (b), I confess that I am at a loss about the move to (3) from (1) and (2). I am unable to contrive an interpretation of "confirmed" in (3) that leaves any chance of the derivation's validity. Perhaps it is just a failure of imagination. I do not, in any case, propose to attempt to bring our discussion of holism to a final solution.

The general specification of R is, of course, not an open-and-shut affair; neither is the tightness of the fit of evidence to confirmations that R affords an

easy thing to describe. Proxy functions are part of this problem. As we saw in the Prologue, "a set of sentences can be reinterpreted in any one-to-one way, in respect of the things referred to, without falsifying any of the sentences" (1995, p. 72), and so "if we transform the range of objects of our science in any one-to-one fashion, by reinterpreting our terms and predicates as applying to new objects instead of the old ones, the entire evidential support of our science will remain undisturbed" (1992, p. 8).

Those who do not mind the intended dualism between empirical and abstract theories will welcome the difficulties in which semantic holism finds itself. Perhaps they will go even further, insisting that precisely where the dualism is most sharply edged it does not matter whether semantic holism is true. Its edges are sharpest in the higher reaches of mathematics and pure logic, and it is there that holism – which if plausible at all is plausible for scientifically applicable mathematics – quickly becomes implausible for its attempt to snare inapplicable mathematics as well. Quine himself asks "about the higher reaches of set theory itself and kindred domains which there is no thought or hope of applying in natural science" (1992, pp. 5–6). His answer resembles the stand he takes against a special epistemology for mathematics: It is uneconomical to contrive a special semantics for the higher reaches, whose sentences "are couched in the same vocabulary and grammar as applicable mathematics" (1992, pp. 5–6). Special accommodation would involve "an absurdly awkward gerrymandering of our grammar" (1992, pp. 5–6).

For those who are still not drawn to our dualism, Craig's Theorem beckons attractively (1953). The theorem asserts that for any theory in which a partition exists on empirical and theoretical terms, theorems containing theoretical terms reduce without relevant loss to theorems containing empirical terms only. Thus, in principle, empirical terms are all the terms required for the adequacy of any theory containing theoretical terms as well. There is no effective means of finding a purely empirical reducer for any such mixed theory. Craig's Theorem requires the prior specification of the mixed theory in order that the existence of the pure theory can be proved in the abstract. Therein lies a distinction resembling the one I am seeking to invoke. An abstract theory *modulo* Craig's Theorem is a theory requiring such prior specification.

Ramsey sentences offer the same appearance of relief from dualism. They are Ramsey's way of eliminating reference to theoretical entities in science. Ramsey sentences arise from term-containing sentences by displacement of terms with individual variables and concomitant binding by way of the existential quantifier. Applied to a theory's every theoretical term-containing sentence, Ramsification lays bare the topic-neutral structure of the theory (Ramsey, 1931). Ramsification anticipates Quine on proxy functions, a move that extends a thesis about the reference of theoretical terms to a thesis about the reference of all terms. Dualism is avoided right enough, but it is *term*-dualism (which is what I do not want) rather than *theory*-dualism (which is what I do want).

An abstract science is a discipline that makes its enquiries and reaches its conclusions without the benefit or discipline of empirical checkpoints. It is sometimes contended that the definition is empty, since no science or discipline worthy of the name fails to engage the empirical check, however indirectly. Even the upper reaches of mathematics, it is said, make contact with the empirical by virtue of the indispensability of some branches of mathematics to the hard sciences.

I am unconvinced by this argument, but it does not matter. My conception of abstractness is a practical one. When a set theorist or a topologist or logician announces his axioms, produces his arguments, and draws out his theorems, he rarely, if ever, does so with improvements to physics in mind, and he never allows his conclusions to be judged by their amity toward empirical science, even if in the fullness of time such amity proves to have existed (consider, for example, the surprising applicability of category theory to the methodology of mathematical physics, or the fact that the permanent stoppage of the heart cannot be explained fully without a theorem from topology). This is abstractness at the level of praxis, but it is abstractness enough for the purposes of this book.

The general question is "How do abstract theorists go about their business without the comforts of empirical checkpoints?" The particular question is "How do abstract theories resolve their differences, especially their heartfelt differences about basic things?" The particular question is important in a way that the general question is not, important as it is otherwise. Pressing the particular question of conflict resolution strategies is an efficient way of unmasking bad answers to the first, more general, question.

Our two questions bear on a third. Can the abstract theorist do his business and resolve his quarrels in ways that preserve realist assumptions; that is, in ways that allow him to think that how well he does his business and how well he settles his disputes will be a matter of how close he gets to the objective facts of the matter at hand?

It is easy to see that two methodologies dominate the abstract sciences. One I shall call the *method of intuitions*. The other is the *method of costs and benefits*. On the face of it, the method of intuitions is tailor-made for scientific and philosophical realism. The cost-benefit methodology is more a creature of prudence, an exercise in doxastic economics, so to speak. It delivers the goods for realism, if at all, in a much less obvious and less direct way.

A Medical Analogy

I do not, as I say, intend to pursue the distinction between empirical and abstract theories to its philosophical finality. Imperfectly drawn as it may be, and philosophically questionable as it may also be in the abstract, in practice it is a distinction too attractive not to make use of. In this book, I shall be concerned with disagreements that arise in abstract theories such as logic, set theory, formal semantics, and certain of the normative disciplines. The first

task is to specify the dialectical structure of disagreements of the sort that I wish to examine. For this a medical metaphor is an inviting way of proceeding. In medical practice, when an injury or an illness befalls,

* **symptoms** present themselves.

There follows

* a **diagnosis**

and then,

* some **triage**.

Thereupon

* a **treatment** is proposed

in light of which

* a **prognosis** is made.

It is much the same way with conflict in abstract theories. If we take, as an example, the sound and fury that attend the classical theorem known as *ex falso quodlibet* – that if a contradiction is provable then every sentence is provable – our medical figure applies as follows.

Symptoms. In the metatheory of classical propositional logic and in modal systems such as Lewis' S5 *ex falso* is provable.

Diagnosis. A great many theorists are agreed that the derivability of *ex falso* is paradoxical, at least in the sense of being sharply *counterintuitive*.

Triage. Depending on what is made of the verdict of paradox, a number of possibilities present themselves. Triage is a way of answering the question, "How bad is it?" Historically, answers range all the way from "It is not bad at all; *ex falso* is counterintuitive only in a weak sense; it is only a surprise," to "It is very bad. It is counterinuitive in a sense strong enough to convict any theory in which *ex falso* is derivable of the derivation of a falsehood." An even stronger finding is possible: *ex falso quodlibet* violates the very meaning of "is derivable" and "implies," and so is not just false but semantically or conceptually false, hence necessarily so (Anderson and Belnap, 1975, ch. 1).

Treatment. Depending on the results of triage, treatment can range all the way from none to a decision to change one's logic in ways that block *ex falso*. Historically, proponents of systems of strict implication opted for the first treatment-option. For others, such as paraconsistent logicians – relevant logicians being prominent among them – the required treatment is the displacement of the

"classical" treatment of implication by some or other deviant variation, such as the relevant system R of Anderson and Belnap (1975, pp. 249–391).

Prognosis. Where treatment is deemed unnecessary there is no cause for prognosis. For those who opt for treatment, there should be some thought as to how to answer the question, "How will the patient now fare?" If one is a relevant logician, there will be a disposition to argue that not only will the patient benefit from the expulsion of a false theorem, but that in its restored state the patient will do a better job in giving a realistic account of rules of deductive inference, for example.

As conceived of by theorists such as Russell at the turn of the twentieth century, set theory threw up an interesting symptom. The symptom was the derivability in intuitive set theory of the Russell Paradox, which demonstrates the existence of a set that is a member of itself if and only if it is not a member of itself. The diagnosis, again, was paradox, and by a broadly accepted triage the paradox was very bad news indeed, since, with the aid of the law of Excluded Middle, it implies an explicit contradiction in which the Russell set both is and is not a member of itself. In the years since 1902, nearly all theorists have agreed on at least the general type of treatment required. The consensus was that intuitive set theory would have to be replaced by a new theory constructed in ways to avert a Russell Paradox. Prognoses varied depending on how close the analyst was to the symptomatic event of 1902. First-generation postparadox theorists took comfort in the presumed consistency of set theories such as ZF (Zermelo-Fraenkel), ZFC (Zermelo-Fraenkel with Choice), and NBG (von Neumann-Bernays-Gödel), but they also were disposed to think of the mechanisms for the exclusion of the Russell set as artificial, *ad hoc*, and counterintuitive. Later generations came to see ZF, or some or other spinoff of the cumulative hierarchy, as capturing the ordinary concept of set – as natural as breathing almost.

Our two problem cases touch on and, so to say, infect one another. What to make of the Russell Paradox hinges in no mean way on what a contradiction implies, hence on whether *ex falso* is true. As a matter of contingent history, opinion has clustered around the position that because *ex falso* is true the Russell Paradox is bad enough to require the replacement of the old set theory with something new and different enough to prevent paradox from reobtruding. Here is a position in which when a theory T collides with classical logic we change T; we do not change logic. It is well to note, however, that in principle the reverse strategy is also available: *Retain* T and *change* logic. Such is the position of paraconsistent logicians, logicians who see *ex falso* as false, and for whom the presumed coincidence between a theory's negation inconsistency and its absolute inconsistency is a mistake. A paraconsistent theory is both inconsistent and not; it is negation-inconsistent and yet absolutely consistent. Beyond these fundamentals, paraconsistentists fan out in two main, and irreconcilable, directions. There are those for whom the negation-inconsistency of

a theory T is bad enough, short of implying omniderivability, to call for a successor theory T*. Relevant logicians typify this first sort of paraconsistentist. When faced with paradox in a theory T they are *comprehensive* revisionists, changing *both* logic and T alike. Paraconsistentists of a less meddlesome stripe try to hold the line at a change of logic only. It is more easily said than done, of course. The main idea amounts to a bold new policy for the management of negation-inconsistency, its triage and its treatment. What is proposed is that negation-inconsistency is not so bad after all, certainly not bad enough for surgical removal. Under any such policy, set theory will continue to be done with the old inconsistency left in. But it will not be the old set theory. New or old, what a theory of sets is able to prove depends on what it takes sets to be, and on the implication relation that it embeds. A logic in which negation-inconsistency does not imply absolute inconsistency is a logic different enough from classical logic to produce, in the application of its proof structures even to the old axioms on sets, theorems quite different from those authorized by the old theory, the theory got by applying classical proof procedures to the same axioms.

Paraconsistentists of this second stripe likewise come in two variations, weak and strong. The weak paraconsistentist sees a distinction between inconsistency and contradiction. Say what you like about inconsistency, it is not as bad as contradiction, which is very bad. A theory is inconsistent in the sense presently intended if and only if it is negation-inconsistent, that is, for some sentence Φ both it and its negation $\ulcorner\neg\Phi\urcorner$ are derivable. A theory contains a contradiction if and only if, for some Φ, $\ulcorner\Phi \wedge \neg\Phi\urcorner$ is derivable. Among paraconsistentists of this weak breed there is something to be said for suspension of the Adjunction law, which proves the conjunction of arbitrary pairs of theorems. Those who opt for the cancellation of Adjunction can block outright contradiction, but they tend to vary in their treatment of inconsistency, a matter which I take up in Chapter 3. More radical are paraconsistentists of dialethic stripe. "Dialethic" comes from the Greek words for "two" and "truth." It conveys a tolerance for the truth of contradictory pairs of propositions. Equivalently, it allows in selective cases for concurrent possession of both truth values. Dialethic logic may first have been a gleam in the eye of Heraclitus and – however tacitly and half-bakedly – it has tried to hold the coat of Philosophical idealism in certain of its variations, as witness the *Greater* and *Lesser Logic* of Hegel.

It would be handy to have names for the various ways of being a generic paraconsistentist. I reserve the terms "relevant logician" and "relevantist" for paraconsistantists of the first stripe, that is, for those whose treatment of paradox calls for across-the-board change to the paradoxical theory and its underlying logic alike. Weak paraconsistentists and strong paraconsistentists, or dialethists, agree on a policy for a theory's inconsistency, namely, that it need not destroy the theory even if left in, but they fall out over contradictions, with the dialethist allowing that, on occasion, even they might be true.

Our medical metaphor also can be put to use in the case of a third paradox, the so-called Tarski paradox, but that belongs in truth to Eubulides (thought credited by St. Paul to Epimenides). Consider these statements.

(1) is not true
(2) is a statement (i.e., a bivalent sentence).

(1) is true if and only if it is not. Here, too, the symptoms are the demonstration of something paradoxical. Diagnosis reveals a contradiction, since with the aid of Excluded Middle, the express contradiction "(1) is true *and* (1) is not true" is derivable. As before, most triagists agree that contradiction is a serious problem, certainly serious enough to justify even rather radical steps to evade the paradox.[7] Accordingly most treatments involve – or are represented as involving – the gerrymandering of language in ways that prevent paradoxical recurrence. Prognosticators are hopeful, by and large. Although the Tarski Paradox puts natural language out of business, paradox-free formalized languages are available, either in fact or in principle, to do the serious business of science.

Conflicts in the abstract sciences owe something of their dialectical flavor to our medical metaphor. Theorists can disagree in their diagnoses, in their triagic and treatment judgments, and in their prognoses. The Liar Paradox illustrates diagnostic disagreement. Some theorists think that the Liar *proof* is defective. For them there is no paradox, and if there is trouble anywhere near at hand, they tend to see it in the assertion that the Liar sentence is indeed a statement. Rival reactions to the Russell Paradox and *ex falso* are not typically diagnostic. For the most part, theorists agree that there is something genuinely paradoxical under foot. In the case of *ex falso*, there is substantial disagreement at the level of triage, with judgments ranging from "not at all bad" to "not all that bad" to "horrible." Beyond that, contentions ramify noticeably. Strictists (so called after Lewis's systems of strict implication) require no more by way of treatment than the reassurance of a supplementary proof, revealing that *ex falso's* triagic worst is "not all that bad." Among those of harsher triagic judgment, contentions and alarums cluster around treatment options, and to a lesser extent around prognostication. With set theory we see a different contention space: Broad symptomic agreement (there is a paradox here); broad diagnostic agreement (the paradox proves a contradiction); a solid if not perfect consensus about triage (a bad problem); and a flourishing dissensus about treatment, both as regards *what* should be treated, and by what *means*.

The historical record reveals, for both the Liar and the Russell Paradoxes, diagnoses and triages more dire than those we have examined so far. Concerning the latter, Frege and Russell saw the paradox as a proof of the inconsistency of the concept of set. Tarski thought that the Liar established the inconsistency of the concept of truth;[8] or in greater strictness, as I have suggested, that it showed the inconsistency of the concept of statement, that is bivalent