

Chapter One

THE DECLINE OF EXPLICIT ESSENTIALISM

Upon examining "Moore's and Russell's principal argument for the reality of universals, in order to determine whether any spark of life remains in it"—"Is it truly dead, or only neglected?"—A. Donagan concludes that it is still very much alive. And surely this is what the history of philosophy would lead one to expect: philosophical positions—except for scientific theories classified as philosophical, e.g. Aristotle's theory of the ethereal spheres—rarely, if ever, perish. Being criticism-resistant, as it were, they remain to be taken up by anyone attracted to them; which is one of the great enigmas of philosophy.¹

From just explicit formulations, it might appear that essentialism, after a bright Platonic dawn, medieval high noon, and modern decline, has finally entered its twilight time and is fading from view. At best, this picture is partial. A better metaphor, in view of recent developments, might be an ocean whose successive waves surged ever higher as the tide of essentialism rose and now lap ever lower as the tide runs out, but which leave the sand damp with their passing. The waves are successive essentialistic theories; the dampness is the essentialistic practice that endures after the waves have subsided. In the present chapter I shall speak of theoretical developments within the present century which illustrate this metaphor of successive, retreating waves. In subsequent chapters I shall turn to varied forms of philosophical practice which attest the lingering power of essentialistic thinking.

A Theoretical Sampling

In the first part of the century, broad essentialistic claims were still common. Listen, for example, to Sir David Ross:

The essence of the theory of Ideas lay in the conscious recognition of the fact that there is a class of entities, for which the best name is probably "universals," that are entirely different from sensible things. Any use of lan-

guage involves the recognition, either conscious or unconscious, of the fact that there are such entities; for every word used, except proper names—every abstract noun, every general noun, every adjective, every verb, even every pronoun and every preposition—is a name for something of which there are or may be instances. The first step towards the conscious recognition of this class of entities was, if we may believe Aristotle, taken by Socrates when he concentrated on the search for definitions; to ask for the meaning of a general word was a step from the mere use of such a word towards the recognition of universals as a distinct class of entities.²

Ross here goes beyond Socrates or Plato. Even Aquinas might have hesitated to speak with such assurance of “any use of language” and “every word used,” attentive as he was to analogous concepts.

In its breadth, Ross’s view echoes Russell’s in *The Problems of Philosophy*. “We succeed in avoiding all notice of universals as such,” Russell wrote, “until the study of philosophy forces them upon our attention.”³ Plato pioneered their exploration; the theory Russell advocated was, he acknowledged, “largely Plato’s, with merely such modifications as time has shown to be necessary.”⁴ Developments since Spinoza have shown the need to generalize Plato’s essentialism. “When we examine common words, we find that, broadly speaking, proper names stand for particulars, while other substantives, adjectives, prepositions, and verbs stand for universals.”⁵ The latter categories merit special emphasis. “Even among philosophers, we may say, broadly, that only those universals which are named by adjectives or substantives have been much or often recognized, while those named by verbs and prepositions have been usually overlooked. This omission has had a very great effect upon philosophy.”⁶

Talk of “universals,” like talk of “essences,” can have a stronger or weaker sense; it may or may not express the kind of unifying, simplifying, one-thing-only conception that concerns us in this study. The same is true of “general ideas,” the term Moore used to express a similarly comprehensive viewpoint. There can be little doubt, however, about the genuineness of Moore’s essentialism. His treatment of color concepts is particularly revealing:

This character wh. we express by “is a shade of blue,” is, of course, something which is common to all shades of blue—something which they have “in common.” Some people seem loth to admit that they have anything “in

common." And of course this character is not "in common" to both of 2 blue shades, in the sense that it is a part or constituent of both . . . Obviously this character also is not identical with any shade which possesses it, nor yet with any other shade of colour that we see. It is not similar in shade to any shade that we see. So that, if it is "seen" at all, it is only in a completely different sense.⁷

"All the shades we see occupy some position in the colour octahedron; but 'blue,' in the sense in which many of the shades in the octahedron are 'blue,' occupies no position in it: therefore it is not seen."⁸

Moore was not thinking of a hidden cause. "Consider yellow," he wrote. "We may try to define it, by describing its physical equivalent; we may state what kind of light-vibrations must stimulate the normal eye, in order that we may perceive it. But a moment's reflection is sufficient to shew that those light-vibrations are not themselves what we mean by yellow. They are not what we perceive."⁹

The like holds for good. "It may be true that all things which are good are also something else, just as it is true that all things which are yellow produce a certain kind of vibration in the light . . . But far too many philosophers have thought that when they named those other properties they were actually defining good; that these properties, in fact, were simply not 'other,' but absolutely and entirely the same with goodness."¹⁰ "My point is that 'good' is a simple notion, just as 'yellow' is a simple notion; that, just as you cannot, by any manner of means, explain to any one who does not already know it, what yellow is, so you cannot explain what good is."¹¹

Were it suggested that neither yellow nor good is simple—that there are different shades of yellow and different species of good—doubtless Moore would respond in the way just seen: as yellow or blue is invisible and distinct from visible hues, so good is a nonnatural entity, distinct from any natural, descriptive properties. Specific goods may vary, but not *the good*. Not the essence they all share.

Various indications¹² suggest that Russell understood "universals" as essentialistically as Moore did "general ideas." Scheler's remarks bespeak still more clearly the unitary, non-disjunctive nature of the essences he espoused. For Scheler, "that a man or a deed is 'noble' or 'base,' 'courageous' or 'cowardly,' 'innocent' or 'guilty,' 'good' or 'evil,' is not made certain for us by constant characteristics which can be discerned in such things and events; nor

do such values *consist* in such characteristics. In certain circumstances a *single* deed or a *single* person is all that we need to grasp the *essence* of the value in question."¹³ And what holds for value concepts holds for all.

An essence, or whatness, is in this sense *as such* neither universal nor particular. The essence red, for example, is given in the universal concept as well as in each perceivable nuance of this color. The differences between universal and particular meanings come about only in relation to the objects in which an essence comes to the fore. Thus, an essence becomes *universal* if it comes to the fore in a plurality of otherwise different objects as an identical essence: in all and everything that "has" or "bears" this essence. The essence can, on the other hand, also constitute the nature of an *individual thing* without ceasing to be such an *essence*.¹⁴

In *Ideas*, Husserl used other illustrations to express a similar viewpoint:

An individual object is not simply and quite generally an individual, a "this-there" something unique; but being constituted thus and thus "*in itself*" it has *its own proper mode of being*, its own supply of *essential* predicables which must qualify it (*qua* "Being as it is in itself"), if other secondary relative determinations are to qualify it also. Thus, for example, every tone in and for itself has an essential nature, and at the limit the universal meaning-essence "tone in general," or rather the acoustic in general—understood in the pure sense of a phase or aspect intuitively derivable from the individual tone (either in its singleness, or through comparison with others as a "common element").¹⁵

"An instance of the essence 'colour' and an instance of the essence 'sound' are intuitively 'present,' and indeed *as* instancing their own essences."¹⁶

Husserl went further and noted that we can recognize such essences even if the instantiating samples are not real ones.

If in the play of fancy we bring spatial shapes of one sort or another to birth, melodies, social happenings, and so forth, or live through fictitious acts of everyday life, of satisfaction or dissatisfaction, of volition and the like,

we can through "ideation" secure from this source primordial and even on occasion adequate insight into pure essences in manifold variety: essences, it may be, of spatial shape *in general*, of melody *as such*, of social happening *as such*, and so forth, or of the shape, melody, etc., of the relevant special *type*. It is a matter of indifference in this connexion whether such things have ever been given in actual experience or not.¹⁷

Wittgenstein's Critique

A characteristic feature of twentieth-century views like those just sampled and of similar views in earlier centuries was their slight attention to language and to the customary use of the terms the authors employed. When, for example, Russell, Moore, or Scheler alluded to the various shades of some color—red, white, yellow, blue—and alleged an essence present in all members of the class, he did not state explicitly that the class in question was the one picked out by the standard employment of that particular English term (*red, white, yellow, blue*) or by its German equivalent (*rot, weiss, gelb, blau*). So long as one does not look beyond one's native tongue and its near neighbors and envisage a different color cartography (joining red and pink, say, or red and orange, under a single predicate), this silence about usage may appear relatively unproblematic. More evidently questionable is Scheler's assertion that "in certain circumstances a *single* deed or a *single* person is all that we need to grasp the *essence* of the value in question."

If an essence, to qualify as an essence of X, must be present in all X's, then there is no telling from any single member of the class that the essence present in it is the essence of X. Perhaps Scheler meant merely that a person can become acquainted with an essence through acquaintance with any instance in which the essence is present. (One could, for instance, become acquainted with the essence of nobility, if it has one, by observing a single noble deed.) However, the further difficulty would remain that the use of terms like *noble, courageous, innocent, and good* reveals much diversity from speaker to speaker, in both intension and extension.

So the question arises: What relation is the purported essence supposed to have to the current or past use of the term applied to it? What relationship holds between the essence of X and the use of the word X? Must the essence be found in all the things so named, or just in some of them? If just in some, in which ones, and

why? Or does usage matter? If not, why speak of the essence of X rather than of Y or Z?

These queries take on a special urgency when Husserl asserts that the "play of fancy," not restricted to real cases, may reveal the essence of melody *as such* or of social happening *as such*. Does it make no difference whether, in order to characterize the nonlinguistic phenomenon Husserl alludes to, one speaks of "*melody as such*" or "*tune as such*" or "*melodic line as such*"—or for that matter of "*music as such*" or "*sound as such*"? Does it make no difference whether the expression *melody* or *social happening* has ever been employed or has been applied to such a phenomenon (as his closing remark would seem to suggest)? If it does make a difference, should not actual samples and actual usage be consulted to ascertain whether this or that aspect of the contemplated fiction instantiates a given essence? What does the expression "as such" signify if it has no connection with familiar applications of the terms Husserl employs (*melody, social happening*)? And if there is a connection, what is its character? How does the essence of X relate to the word X?¹⁸

Essentialists before Wittgenstein seldom addressed considerations like these, and when they did, their explanations did little or nothing to clarify the status of the "essences" they alleged.¹⁹ Moore, for example, said he was concerned with "that which is meant by 'good'"²⁰—with "that object or idea, which I hold, rightly or wrongly, that the word is generally used to stand for"²¹—yet declared, "At the same time I am not anxious to discuss whether I am right in thinking that it is so used."²² The good, we can imagine Moore explaining, would still be the good even if called by some other name, just as fish would still be fish if called figs, or bison would still be bison if called candelabra. This sounds reasonable enough, except that Moore has taken away with one hand what he gave with the other. To verify what he says about the good, we must examine what, if anything, the word *good* customarily designates, and see whether he accurately describes it; yet we are also advised that we need not concern ourselves with whether the word does in fact stand for the entity Moore says it does. Perhaps no finer illustration can be found of what Wittgenstein had in mind when he wrote: "The essential thing about metaphysics: it obliterates the distinction between factual and conceptual investigations."²³ Not merely ignores: obliterates, systematically obfuscates.²⁴

Other remarks of Moore sound more sensible. "I should, indeed, be foolish," he wrote, "if I tried to use [the word *good*] for something which it did not usually denote; if, for instance, I were

to announce that, whenever I used the word 'good,' I must be understood to be thinking of that object which is usually denoted by the word 'table.'"²⁵ It would be similarly misleading to speak of "all X's" and mean only some of the things so designated; or to speak of "the essence of X's" and mean a trait found in some X's but not all, or in all X's but also in some Y's and Z's; or to speak of what the word X denotes or stands for when the word does not typically denote or stand for anything but has a different function.

For reasons like these, Wittgenstein advised that when philosophers employ some term and "try to grasp the *essence* of the thing, one must always ask oneself: is the word ever actually used in this way in the language-game which is its original home? What *we* do is to bring words back from their metaphysical to their everyday use."²⁶ Not that everyday use is sacrosanct;²⁷ but if another use seems more opportune—practically or theoretically, contextually or generally—let it be introduced as a new word-use, not paraded as an essence.

To recognize the actual, everyday use of words, Wittgenstein advised that we *look*, not *think*.²⁸ His meaning can be gathered from preceding samples. An invisible blue distinct from all shades of blue, for example, is not seen with the eyes but with the mind. It arises from the thought that varied blues—cobalt, turquoise, aquamarine—are, after all, all instances of blueness. They share this trait, have it in common. So it must be somehow perceptible in each instance.

Wittgenstein recognized that there is such a thing as spotting what is common and abstracting it from varied instances. "Suppose I shew someone various multi-coloured pictures, and say: 'The colour you see in all these is called "yellow ochre" . . . Then he can look *at*, can point *to*, the common thing.' But "compare this case: I shew him samples of different shades of blue and say: 'The colour that is common to all these is what I call "blue".'²⁹ Now what can be looked at or pointed to save the varied hues of blue? And don't say, "There *must* be something common, or they would not be called 'blue,'" "but *look and see* whether there is anything common to all."³⁰

With regard to "games" Wittgenstein reached the well-known verdict that nothing common to all games makes us name them so, "but similarities, relationships, and a whole series of them at that."³¹ To this "complicated network of similarities, overlapping and criss-crossing," he applied the label "family-resemblances," explaining that "the various resemblances between members of a family: build, features, colour of eyes, gait, temperament, etc., overlap and criss-cross in the same way."³² Blues, however, do not

relate in this way. Neither do the members of many other classes.³³ Wittgenstein recognized a variety of conceptual structures besides the family-resemblance configuration. But a single-essence, all-and-only structure, permitting the statement of sufficient and necessary conditions of application, did not figure among them. Still, he did not demonstrate the impossibility of such a structure—did not even attempt to. And for all a few samples like *blue* and *game* demonstrate, essences might be common. So it is not immediately evident why Wittgenstein's critique had the impact it did or merited so much attention.

One reason emerges from the quotations with which we began our sampling of essentialistic thinking. When Russell and Ross alleged the omnipresence of essences and universals, for nouns, verbs, adjectives, adverbs, and even prepositions, their claim did not result from empirical inquiry. They had not perused samples of each class of expression, in representative languages, and found that time after time these samples revealed such essentialistic uniformity. No, they supposed—a priori—that it had to be so. The significance of Wittgenstein's jejune examples derives from the challenge they pose to this underlying assumption.³⁴ If color predicates, proposed as paradigmatic by essentialists like Russell, Moore, and Scheler, reveal no essence when scrutinized a posteriori, and neither does a "univocal" concept as typical as *game*,³⁵ perhaps essences are rare or nonexistent. Indeed, the all-or-nothing essentialistic mode of thought would tend to reinforce this negative conclusion. According to Moore, if there are any essences, there are tremendous numbers of them;³⁶ a *modus tollens* argument would therefore suggest that if there are not tremendous numbers, there are not any.

Wittgenstein's impact also stemmed from his stress on the neglected linguistic parameters of essentialism. If an essence does not correlate with the word an essentialist uses to label it, then he should label it differently. If it does and must correlate with the label, then its existence is not just a fact of nature but also a fact of language. The *concept X* reveals the existence or nonexistence of an essence of X.³⁷ But the concept belongs to a language; and that language is the common property and common creation of a vast and varied population, whose employment of expressions, on various occasions and for varied purposes, cannot be expected to manifest the austere uniformity required to establish or perpetuate essences. When children, housewives, mechanics, poets, journalists, and mystics are the speakers, a similarity from shade to shade of some color, or a family resemblance between various

games, may suffice for the common appellation "blue" or "game." Indeed, in metaphysically licentious moments such folk may speak of "feeling blue," "singing the blues," "war games," "a war of nerves," or "the game of politics." Hence, when viewed as a matter of linguistic usage, and not of metaphysics, the essentialistic structure of language, thought, and world appears extremely dubious.³⁸

Carnapian "Explication"

Vagaries of usage have not deterred scientifically-minded theorists. "There is still the opportunity," writes Ian McGreal, "which Wittgenstein tended to ignore or overlook, of constructing a definition of 'games' by the use of which one could give a definitive (but creative) answer to a question as to the essence of games. One might take advantage of the ambiguity of such a term as 'amusing,' or one might, by stipulation, rule out many of those cases now conventionally called 'games.'"³⁹ In science, notes Joseph Margolis, "'family resemblances' between different kinds of energy, for example, [have] had to give way gradually to an empirically adequate definition of the necessary and sufficient properties of energy."⁴⁰ Philosophers, it would seem, may follow suit. However, when they have attempted to do so, they have generally been as unclear as traditional metaphysicians concerning the nature of their enterprise. The intended relation of their theories or definitions to the concepts or terms they employed has remained obscure.⁴¹ Often, when they have offered something by way of explanation for their partial or total disregard of usage, they have characterized their undertaking as Carnapian "explication."

"If there is one concept that might be said to provide a key to Carnap's philosophy," observes Peter Achinstein, "it is *explication*, a concept that has had considerable influence on many philosophers."⁴² "By an explication," writes Carnap, "I understand the replacement of a pre-scientific inexact concept (which I call 'explicandum') by an exact concept ('explicatum'), which frequently belongs to the scientific language . . . Although explications are often given also by scientists, it seems to me particularly characteristic of philosophical work that a great part of it is devoted to proposing and discussing explications of certain basic, general concepts."⁴³ At least the proposals look more plausible viewed this way. For in explication "the only essential requirement is that the explicatum be more precise than the explicandum."⁴⁴ Hence, "the interpretation which we shall adopt . . . deviates deliberately from

the meaning of descriptions in the ordinary language. Generally speaking, it is not required that an explicatum have, as nearly as possible, the same meaning as the explicandum."⁴⁵ Rather, it suffices that the explicandum satisfy the following four requirements, "to a sufficient degree":

1. The explicatum is to be *similar to the explicandum* in such a way that, in most cases in which the explicandum has so far been used, the explicatum can be used; however, close similarity is not required, and considerable differences are permitted.

2. The characterization of the explicatum, that is, the rules of its use (for instance, in the form of a definition), is to be given in an *exact* form, so as to introduce the explicatum into a well-connected system of scientific concepts.

3. The explicatum is to be a *fruitful* concept, that is, useful for the formulation of many universal statements (empirical laws in the case of a nonlogical concept, logical theorems in the case of a logical concept).

4. The explicatum should be as *simple* as possible; this means as simple as the more important requirements (1), (2), and (3) permit.⁴⁶

"Philosophers, scientists, and mathematicians," Carnap adds, "make explications very frequently. But they do not often discuss explicitly the general rules which they follow implicitly."⁴⁷ In philosophy, one frequently has the impression that the tag *explication* serves as a warrant to do business as usual in the essentialistic manner. With the constraints of standard usage loosened, the theorist can continue to equate his explicandum with some single explicatum, and do so with an appearance of scientific rigor. But with the rules of the game left ill-defined,⁴⁸ or the rules once stated thereafter ignored,⁴⁹ the reality of the enterprise contradicts the appearances. In fairness to Carnap it should be said that most of the explications that have failed have not been genuinely Carnapian; they have not adhered to the guidelines just cited. However, the fog surrounding the explicative enterprise,⁵⁰ the obscurity of individual specimens, and the tendency to conceive explication essentialistically are traceable in part to Carnap's own characterizations.

First, the very name *explication* suggests something more than merely stipulating a handy sense for a term. Yet basically that is all explication, as stated, amounts to. Consider a simple case. The interpretation of "if . . . then" in the sense of material implication

satisfies all four of Carnap's conditions and has been cited by Carnap and others as an instance of explication.⁵¹ Yet clearly such a sharpening does not explain the existing concept, as the term *explication* suggests. The process is the familiar one of stipulating or specifying the sense an expression will have within a given context, work, or discipline. One sense among several gets singled out for its utility.⁵²

Second, as an exception to the general failure to clarify the rules of explication, Carnap cites the "good explicit formulation . . . given by Karl Menger in connection with his explication of the concept of dimension":

He states the following requirements. The explicatum "must include all entities which are always denoted and must exclude all entities which are never denoted" by the explicandum. The explication "should extend the use of the word by dealing with objects not known or not dealt with in ordinary language. With regard to such entities, a definition [explication] cannot help being arbitrary." The explication "must yield many consequences," theorems possessing "generality and simplicity" and connecting the explicatum with concepts of other theories.⁵³

As reported by Carnap, these stipulations of Menger look inconsistent; in any case, they clearly do not agree with Carnap's own four requirements. Carnap's first prescription says nothing about extending the range of the explicandum nor, more essentialistically, about including all entities which are always denoted and excluding all which are never denoted. He stipulates more flexibly. For in his own conception no explicatory formula—no inclusion or exclusion—is to be arbitrary; each is to be fruitful, and fruitfulness may require narrowing in one instance, stretching in another.

Third, given the looseness of Carnap's conditions and their pragmatic motivation, it is not clear why any specific set, stated any specific way, should qualify as *the* requirements for worthwhile explication. He cites increased precision as "the only essential requirement"; yet often that can be achieved without any attention to fruitfulness. On the other hand, greater fruitfulness can sometimes be had without increased precision; a sense may simply be narrowed or extended. More realistic than Carnap's requirements would be a list of desiderata to consider when stipulating senses; and such a list would have to be much longer than his. It would need to speak, for example, of the "explication's" context, of its

recipients, and of how it is conveyed to them.⁵⁴ A once-and-for-all listing like Carnap's, suggesting that all stipulations of sense should satisfy all these criteria,⁵⁵ ignores context and to that extent is imperfectly pragmatic. It betrays a lingering essentialism.

Fourth, the listing of all four conditions to be met, simultaneously, suggests that there can be only one successful explicatum for any explicandum. But this is not the case, as Carnap is aware. "In a problem of explication," he writes,

the datum, viz., the explicandum, is not given in exact terms; if it were, no explication would be necessary. Since the datum is inexact, the problem itself is not stated in exact terms; and yet we are asked to give an exact solution. This is one of the puzzling peculiarities of explication. It follows that, if a solution for a problem of explication is proposed, we cannot decide in an exact way whether it is right or wrong. Strictly speaking, the question whether the solution is right or wrong makes no good sense because there is no clear-cut answer.⁵⁶

The imprecision lies more in Carnap's prescriptions than in the data to which they are applied. All four of his conditions may, on occasion, conflict with one another, and he assigns no respective weights to them, save to prescribe that the first three conditions take precedence, individually, over the fourth condition of simplicity. His stipulations give no indication whether an explication that is more precise but less fruitful should be preferred to one that is more fruitful but less precise; whether one that departs notably from standard usage but is fruitful should be preferred to one that is closer to usage but somewhat less fruitful; and so forth.

Within this collective imprecision further indefiniteness appears with regard to the individual conditions. The fourth requirement, of simplicity, may be variously interpreted, and the different interpretations will yield different verdicts.⁵⁷ For example, does an explicatum employing fewer but more complex concepts count as more or less simple than one employing more numerous but less complex concepts?⁵⁸ In the first requirement, uncertainty arises from the fact that proximity to existing usage varies with the degree of precision with which the explicandum is identified (for example, as the concept *meaning*, or the concept *meaning* as applied to expressions, or the concept *meaning* as applied to complete utterances, or the concept *meaning* as applied to complete empirical utterances).⁵⁹

These and other reasons account for my conjecture that the

imprecision Carnap recognized lies more in his account of explication than in the explicanda. But whatever the nature or source of the indefiniteness, the upshot is the same: If explication is understood as Carnap defined it, no explicatum can be proposed as the uniquely correct one.

Hence it comes as no surprise that the Carnapian wave was followed by a more clearly essentialistic wavelet. Treating Carnap's account as an explication of explication, Joseph Hanna faulted both Carnap's choice of explicandum and his formulation of the explicatum, and proposed instead a more restricted, uniform account. With regard to the explicandum, he suggested: "Although Carnap frequently speaks of explicating concepts, it seems more perspicacious to follow Quine and view the method as applying fundamentally to linguistic terms (i.e. predicates, operators, names, etc.) and as applying only derivatively to concepts. Concepts are mysterious entities and it is best to avoid them when the same point can be made by referring to predicates."⁶⁰ Carnap's explicatum, which does not require close similarity, likewise needs tightening. "Carnap appears to be assimilating all cases in which presystematic terms are adapted for scientific use to cases of explication."⁶¹ But such assimilation looks doubly objectionable. "In the first place, this use of 'explication' does not seem to be consistent with the method of analysis actually employed by other 'ideal-language' philosophers (the school of which Carnap is spiritual father)."⁶² Hanna acknowledged that "it would be difficult to substantiate this claim because of the conflicting and not very detailed remarks that various 'ideal-language' philosophers have addressed to this issue," then proceeded to "a second, though related, objection to Carnap's assimilation of all cases of 'scientific reconstruction' to explication," namely the objection "that the resulting notion is vague, and it is not apparent how this vagueness is to be remedied."⁶³

Hanna's own presentation supposedly avoids both these difficulties when, for instance, it prescribes that the explicatum should "agree with" the explicandum "in those cases where our intuitions concerning the meaning of the latter term are not defective. In the remaining cases, the explicatum serves to educate our intuitions or re-educate our intuitions, as the case may be. In short, the explicatum systematically fills in the gaps and corrects the inconsistencies of the explicandum."⁶⁴

Hanna's account looks still more defective than Carnap's. For one thing, substituting the word *meaning* for the word *concept* (as in the preceding quotation) hardly removes all vagueness from the

discussion. And if the test of adequacy is agreement with the term to be explicated, why are Carnap's stipulations criticized for their disagreement with the "method of analysis" employed by kindred thinkers rather than for their disagreement with these thinkers' or others' use of the term *explication*? It seems odder still that Hanna's own explication of the term *explication* is not based in any empirical sampling of usage, and that, instead, the explicandum is identified as "the brand of philosophic analysis variously called explication, rational reconstruction, formal analysis, etc."⁶⁵

With regard to the word *explication*, Hanna briefly noted that in its presystematic usage it has a variety of senses, then explained that only one of these senses was the primary concern of his paper.⁶⁶ It was the sense *characterized* by his explicatum and also, it appears, the sense *identified* by that characterization. Hence his explication looks indefeasible, and if indefeasible, uninteresting (save, perhaps, for other, nondescriptive purposes). Where agreement with the explicandum is the aim of explication or the criterion of success, the explicatum cannot also serve to pinpoint the explicandum it "agrees with." Where, on the contrary, fruitfulness figures as at least a co-determinant, the test of agreement must be loosened, as in Carnap's treatment.

Had Hanna been genuinely concerned about close descriptive equivalence (closer than Carnap prescribed), he would have attended more carefully to the use of the term to be described. Had he attended more carefully to the vagaries of that term's employment, I think he would have recognized that such messy linguistic data really did not interest him. His proposal amounts to a stipulation, for the sake of greater precision; and though the stipulation may be less precise than Hanna imagined, so far as agreement with familiar usage is concerned it does look preferable to Carnap's. To apply the label *explication* to a mere conceptual revision such as that of the concept *fish* does indeed sound odd. But a stipulated sense of *explication* neither reveals nor generates an essence of explication. And with regard to the type of analysis he commended, Hanna's own sample—the explication of *explication*—does not look promising. It offers no reason to surmise that other samples may succeed any better in satisfying essentialistic aspirations.

Whatever Hanna's motivation may have been, his account responds to two basic essentialistic urges. First, an essence must be uniform and sharp, whereas Carnap's characterization admitted considerable variety, with no clear limit; the explicatum would just need to resemble the explicandum to some unstated extent. Second, essential definition must capture an existing essence, not

merely stipulate a useful meaning for a term, whereas Carnap's prescriptions pointed in the latter direction. Had they done so still more clearly, they would have been less apt to perpetuate essentialistic expectations.

Essences and "Rigid Designators"

The doctrine of "rigid designation" advanced in the 70's by Saul Kripke and Hilary Putnam gave rise to another surge of essentialistic thinking. Putnam explained the doctrine by means of the term *water*. "Kripke calls a designator 'rigid' (in a given sentence) if (in that sentence) it refers to the same individual in every possible world in which the designator designates. If we extend the notion of rigidity [from proper names] to substance names, then we may express Kripke's theory and mine by saying that the term 'water' is *rigid*."⁶⁷ For, "once we have discovered that water (in the actual world) is H_2O , *nothing counts as a possible world in which water isn't H_2O* ."⁶⁸ Thus water is necessarily H_2O ; H_2O constitutes its essence. The word *water* rigidly designates H_2O regardless of what superficial properties the H_2O may or may not have. In Putnam's opinion, similar, essentialistic rigidity characterizes not only other natural-kind terms like *gold*, *lemon*, *tiger*, and *acid*,⁶⁹ but also "the great majority of all nouns," for instance "the names of artifacts—words like 'pencil,' 'chair,' 'bottle,' etc.," and "other parts of speech as well."⁷⁰ In each instance, the extension of the term is fixed by the "important physical properties"⁷¹ or "hidden structure"⁷² revealed, or still to be revealed, by science.⁷³

This alleged fact of language Putnam sought to demonstrate through various thought experiments, of which his Twin-Earth fantasy is the best known. Let us imagine, he suggested, that somewhere a planet exists which exactly duplicates ours, except that the liquid there called "water," which in all superficial respects resembles what we call "water," does not consist of H_2O but rather of XYZ (an abbreviated formula for something very long and complicated). Having further spelled out this hypothesis, Putnam considered it evident that the word *water* does not have the same meaning in both places;⁷⁴ that in our sense of the word their stuff is not water;⁷⁵ and that we would not and should not call it water once the discrepancy was noted.⁷⁶ The Twin-Earth scenario thus highlights the essential meaning of the word *water* and, more generally, the functioning of similar expressions.⁷⁷

What Putnam took as evident has not seemed so to others. We

do in fact extend words to specimens with different structures. And the plausibility of Putnam's surmise appears to derive from some unwitting sleight of hand.⁷⁸ As he himself noted elsewhere, hidden structure looks important because of its connection with surface features (specific gravity, fluidity, freezing point, taste, characteristic feel, etc.).⁷⁹ But his Twin-Earth fiction cuts this connection; the difference in microstructure makes no perceptible difference between the two liquids. Anyone who noted this further difference, and not just the one Putnam focused on, might not attach the customary importance to hidden structure. For some strange reason, the structural dissimilarity between H₂O and XYZ would have even fewer consequences than that between H₂O (ordinary water) and D₂O (heavy water). The term *water* might therefore be extended to Putnam's Twin-Earth liquid as it actually was, for instance, to heavy water. Indeed, suppose we alter Putnam's fictional hypothesis and imagine that Twin-Earth "water" was all D₂O. Would Putnam still refuse to call it "water"? Would other people?

It may be suggested that with this counter-example, the war is lost but the battle, at least, may be won. Putnam may be wrong in his essentialism but right about Twin-Earth water. For the similarity between D₂O and H₂O is internal, not external, whereas that in his example is external. However, once this much flexibility is recognized and similarity replaces sameness, there is no reason to suppose that similarity will function more rigidly here than on other occasions.

In a previous work I proposed a "principle of relative similarity" (PRS for short), according to which "for a statement of fact, or informative utterance, to be true it suffices that its use of terms resemble more closely the established uses of those terms than it does those of rival, incompatible terms."⁸¹ Calling the Twin-Earth liquid "water" would satisfy this condition; *water* would conform with current usage more closely than any other existing name. Not only would the appellation satisfy the principle, but it would agree with the facts of usage on which the principle rests—with our familiar use of terms in making assertions and with our familiar use of *true* in assessing the assertions. It would resemble countless other instances in which a term has been extended from familiar paradigms to new ones on the basis of relative similarity.⁸²

Often, even this basis is lacking: we speak, for instance, of "deep sorrow," "thin excuses," "sharp regrets," and "tender mercies." Some may wish to characterize such extensions as "figurative" or "metaphorical"; nonetheless, we do make them, readily and repeatedly. Hence to predict that we would not extend a term like

water even where extensive similarity obtains—more extensive than for any other term available—would reveal more about the persons who make the prediction than about speakers of the language at large. The doctrine of rigid designation not only arises from a selective sampling of cases but also demands a selective sampling of speakers. Doubtless it would contain more truth, and Kripke's and Putnam's intuitions would agree better with those of the populace as a whole, were everyone as scientifically, essentialistically minded as are these philosophers of science.⁸³

The importance of a carefully restricted population within selectively scientific settings, for Putnam's doctrine to have a semblance of realism, appears, for instance, when he writes: "It is logically possible (although empirically unlikely, perhaps) that a species of fruit biologically unrelated to lemons might be indistinguishable from lemons in taste and appearance. In such a case, there would be two possibilities: (1) to call them *lemons*, and thus let 'lemon' be a word for any one of a number of natural kinds; or (2) to say that they are not lemons (which is what, I suspect, biologists would decide to do)." This closing reference to biologists elicits three queries: (1) Does the doctrine of rigid designation apply just to scientific speakers, in scientific contexts, or to speakers generally? (2) If the former, does it surmise what scientific speakers would "decide to do" merely in the sense of what meaning-stipulations they might introduce, or in the sense of what they would assert ("They are not lemons") without stipulating new meanings? (3) If the latter, would their assertions be correct, or would they reveal the kind of conceptual intolerance Carnap,⁸⁴ Achinstein,⁸⁵ and others have criticized?

The doctrine of rigid designators seems to suggest the wrong answers to all three of these queries. And even with regard to scientific speakers in scientific contexts its claims appear unrealistic. Listing "three strategies that might be attempted for identifying privileged sameness relations between the members of a species"—strategies "based, respectively, on intrinsic properties of the individuals, on reproductive isolation of a group of individuals, and on evolutionary descent of a group of individuals"—John Dupré observes: "There are many sameness relations that serve to distinguish classes of organisms in ways that are relevant to various concerns . . . none of these relations is privileged."⁸⁷ Similar flexibility has been noted in other sciences.⁸⁸ For example,

In the context of investigating electronic energy levels in the atom, one might define "electron" as a term referring

to charged particles orbiting about the atomic nucleus. In other situations definitions of this sort would be different, e.g., "particles emitted from a cathode," "fundamental units of electricity," "particles shared by atoms to form chemical bonds." Any or all of these might be considered "definitional" characteristics in a given context, though in others they might constitute hypotheses to be subjected to experimental verification.⁸⁹

Suppose, by way of comparison, that we envisage an empiricist rather than a scientific theory of rigid designation, and apply it to the term *chess*. Chess, a theorist might say, requires a board and pieces. Such is the essence of chess; such are its components in any possible world. Granted, speakers might loosely extend the term *chess* to cover chess by mail or chess in one's head or chess played with a computer; in fact, they have thus twisted "fibre on fibre," despite the absence of board and pieces. However, these variants are not chess in the same sense of the word; they are not genuine instances of *chess*. (Here the word is uttered with emphasis, so as to carry conviction or convey an impression of rigor.) Such a theory would tell us more about the fixed mind-set of the theorist than about the fixity of the language or its terms.⁹⁰ And the theory's vogue, if it found a following, would reveal the current prevalence of similarly essentialistic thinking.⁹¹

Variations

The Kripke-Putnam theory has had a mixed reception. Most commentators have reacted negatively;⁹² but some have limited their criticism to Putnam's extension of the theory beyond natural-kind words,⁹³ while others have argued for a more rigorous or plausible version of the theory. Robert Hollinger has written that rather than assert conceptual rigidity, he "would allow for the possibility—which Putnam does not seriously consider or rule out—that the laws which define a term like 'gold' are such that all and only gold (or at least all gold) conform to them. Such laws would specify empirically necessary and sufficient laws and properties, which would be scientifically defended, and would characterize natural kinds. This possibility is needed to defend essentialism."⁹⁴ Hollinger does not allege that scientists have discovered such laws. "My only claim is that it is a possibility that they might, and that if they did, essentialism would be true. If there are to be

'real essences,' such laws must in principle be discoverable, and the view of science based on the view that they can be must be a regulative ideal if essentialism is to be defensible."⁹⁵

If Hollinger's essentialism looks more plausible than the version it replaces, that is because it makes such generous concessions that it barely merits the label "essentialism." Abandoning conceptual necessity and substituting contingent, empirical necessity marks one retreat from traditional essentialism; waiving the requirement that only members of the class—things of the named kind—can possess the essence constitutes another. Even so, if there are to be essences of the type envisaged, nature will no doubt need some help from definitions. The resulting essences may reflect the essentialistic preferences of those who speak of them as much as they reflect the uniformity of nature.

Consider isotopes. In opposition to a view he labels "Wittgensteinian," Hollinger rightly argues that no new meaning-stipulation was required—no "more or less arbitrary linguistic decision"—when scientists announced "that most, if not all, elements are really mixtures of isotopes, and always have been."⁹⁶ Wittgenstein would raise no objection here, nor would PRS. For, as Hollinger notes, "isotopes resemble their twins in all chemical respects."⁹⁷ Similar flexibility, natural and reasonable, might countenance many another diversification of inner structure without change of name, as for instance in the case of heavy water. But at this point the would-be essentialist faces a dilemma. Either the resulting "essences" will be disjunctive, in which case they will resemble still less the uniform realities traditionally labeled essences, or some linguistic adjustment will need to be made: the name will have to be accorded to just one kind of water, hydrogen, or the like, or be restricted to some aspect of the inner structure that is common to all varieties of the substance. From this dilemma one can sense why Hollinger proposes no actual samples as Kripke and Putnam did, but merely a schema for possible instantiation: "I will . . . construe schema[ta] such as 'Gold is an element with properties $P_1 \dots P_n$ (which conforms to laws $l_1 \dots l_n$)' as empirical/theoretical identification, analogous to 'Water is an aggregation of H_2O molecules.'⁹⁸

Hollinger would apparently grasp the second, more restrictive horn of the dilemma, and assert, more specifically, that real essences, if found, would make our linguistic decisions for us. They would determine what was gold, what was water, and so forth. For the only correct linguistic route is the scientific one, and the only correct, or optimal, way of doing science is the essentialistic way.⁹⁹

Thus the adoption of essentialistic classifications would be neither arbitrary nor conventional.

One way of bringing out this point is by noting the process-product ambiguity in the term "classification." Classification may be the activity of collecting things into groups, or the product of such activities. (I shall call these, respectively, "classification-1" and "classification-2.") Now, as an activity, classification-1 can be either natural or unnatural. Classification-1 is natural only if it serves as a means—ideally the best or only means—of achieving certain ends . . . I claim that there are some ends, e.g., the scientific goals of explanation, prediction, systematization, which can be best achieved—perhaps only achieved—by classifying-1 things into natural kinds, since doing so allows us to gain the most knowledge. If this is so, the fact that our interests put constraints on admissible classifications in no way shows that classifications-1, much less classifications-2, are conventional or arbitrary, or invented rather than discovered.¹⁰⁰

Hollinger's thinking is thus doubly essentialistic: natural kinds have essences, because science has a single, ideal essence. "On this view, in which the conventional element is assumed to be at a minimum, the discovery of real essences amounts, in principle, to the discovery of the ultimate science which is, for Peirce, the ideal limit of scientific inquiry. Thus, if the Peircean ideal is legitimate, so is essentialism."¹⁰¹

To define science, which presently explains so much, in terms of an ideal which has no present instances, sounds suspiciously Platonic. Still, it opens interesting avenues of reflection. Does this contrast between actual science and the alleged ideal attest to present immaturity and incompleteness, or does it bespeak other possibilities than Hollinger envisions? May science continue to enunciate general explanatory laws in terms of bodies, movements, charges, and the like, or must it eventually pass to more specific laws, with regard to tigers, gold, quarks, and other natural kinds, and cluster them in rigid, essentialistic definitions?

It might appear that were specific laws added to the present general laws the result would be "the most knowledge." In support of this conclusion, Hollinger proposes that "the concepts 'essential property,' 'natural kind,' and 'natural necessity' play an explanatory role in theory,"¹⁰² and offers the following illustrative schema: