36. Reichenbach, Hans. *Experience and Prediction*. Chicago: University of Chicago Press, 1938.
37. Scheffler, Israel. *The Anatomy of Inquiry*. New York: Knopf, 1963.
38. Schlick, Moritz. *Allgemeine Erkenntnislehre*. 2nd ed. Berlin: Springer, 1925 (1st ed. 1918).
39. Smart, J. J. C. *Between Science and Philosophy*. New York: Random House, 1968.
40. Woodger, Joseph Henry. *The Techniques of Theory Construction*. Chicago: University of Chicago Press, 1939.

PAUL K. FEYERABEND

Against Method: Outline of an Anarchistic Theory of Knowledge

What is all this commotion good for? The most it can achieve is to ruin one's peace of mind. There one has one's little rooms. Everything in them is known, has been added, one item after another, has become loved, and well esteemed. Need I fear that the clock will breathe fire into my face or that the bird will emerge from its cage and greedily attack the dog? No. The clock strikes six when it is six like it has been six for three thousand years. This is what I call order. This is what one loves, this is what one can identify with. CARL STERNHEIM, *Die Hose*

Preface

The following essay has been written in the conviction that anarchism, while perhaps not the most attractive political philosophy, is certainly an excellent foundation for epistemology, and for the philosophy of science.

The reason is not difficult to find.

"History generally, and the history of revolutions in particular, is always richer in content, more varied, more manysided, more lively and 'subtle' than even" the best historian and the best methodologist can imagine.[1] * "Accidents and conjunctures, and curious juxtapositions of events" [2] are the very substance of history, and the "complexity of human change and the unpredictable character of the ultimate consequences of any given act or decision of men" [3] its most conspicuous feature. Are we really to believe that a bunch of rather naive and simple-minded rules will be capable of explaining such a "maze of interactions"? [4] And is it not clear that a person who participates in a complex process of this kind will succeed only if he is a ruthless opportunist, and capable of quickly changing from one method to another?

This is indeed the lesson that has been drawn by intelligent and

AUTHOR'S NOTE: For support of research I am indebted to the National Science Foundation.

* The notes for this essay begin on p. 94.

thoughtful observers. "From this [character of the historical process]," writes Lenin, continuing the passage just quoted, "follow two very important practical conclusions: first, that in order to fulfill its task, the revolutionary class [i.e., the class of those who want to change either a part of society, such as science, or society as a whole] must be able to master all forms and sides of social activity [it must be able to understand, and to apply not only one particular methodology, but any methodology, and any variation thereof it can imagine], without exception; second, [it] must be ready to pass from one to another in the quickest and most unexpected manner." [5] "The external conditions," writes Einstein, "which are set for [the scientist] by the facts of experience do not permit him to let himself be too much restricted in the construction of his conceptual world by the adherence to an epistemological system. He therefore must appear to the systematic epistemologist as a type of unscrupulous opportunist . . ." [6]

The difference between epistemological (political, theological) theory and scientific (political, religious) practice that emerges from these quotations is usually formulated as a difference between "certain and infallible" (or, at any rate, clear, systematic, and objective) rules, or standards, and "our fallible and uncertain faculties [which] depart from them and fall into error." [7] Science as it should be, third-world science,[8] agrees with the proscribed rules. Science as we actually find it in history is a combination of such rules and of error. It follows that the scientist who works in a particular historical situation must learn how to recognize error and how to live with it, always keeping in mind that he himself is liable to add fresh error at any stage of the investigation. He needs a theory of error in addition to the "certain and infallible" rules which define the "approach to the truth."

Now error, being an expression of the idiosyncrasies of an individual thinker, observer, even of an individual measuring instrument, depends on circumstances, on the particular phenomena or theories one wants to analyze, and it develops in highly unexpected ways. Error is itself a historical phenomenon. A theory of error will therefore contain rules of thumb, useful hints, heuristic suggestions rather than general laws, and it will relate these hints and these suggestions to historical episodes so that one sees in detail how some of them have led some people to success in some situations. It will develop the imagination of the student without ever providing him with cut-and-dried prescriptions and

procedures. It will be more a collection of stories than a theory in the proper sense and it will contain a sizable amount of aimless gossip from which everyone may choose what fits in with his intentions. Good books on the art of recognizing and avoiding error will have much in common with good books on the art of singing, or boxing, or making love. Such books consider the great variety of character, of vocal (muscular, glandular, emotional) equipment, of personal idiosyncrasies, and they pay attention to the fact that each element of this variety may develop in most unexpected directions (a woman's voice may bloom forth after her first abortion). They contain numerous rules of thumb, useful hints, and they leave it to the reader to choose what fits his case. Clearly the reader will not be able to make the correct choice unless he has already some knowledge of vocal (muscular, emotional) matters and this knowledge he can acquire only by throwing himself into the process of learning and hoping for the best. In the case of singing he must start using his organs, his throat, his brain, his diaphragm, his buttocks before he really knows how to use them, and he must learn from their reactions the way of learning most appropriate to him. And this is true of all learning: choosing a certain way the student, or the "mature scientist," creates a situation as yet unknown to him from which he must learn how best to approach situations of this kind. This is not as paradoxical as it sounds as long as we keep our options open and as long as we refuse to settle for a particular method, including a particular set of rules, without having examined alternatives. "Let people emancipate themselves," says Bakunin, "and they will instruct themselves of their own accord." [9] In the case of science the necessary tact can be developed only by direct participation (where "participation" means something different for different individuals) or, if such direct participation cannot be had, or seems undesirable, from a study of past episodes in the history of the subject. Considering their great and difficult complexity these episodes must be approached with a novelist's love for character and for detail, or with a gossip columnist's love for scandal and for surprising turns; they must be approached with insight into the positive function of strength as well as of weakness, of intelligence as well as of stupidity, of love for truth as well as of the will to deceive, of modesty as well as of conceit, rather than with the crude and laughably inadequate instruments of the logician. For nobody can say in abstract terms, without paying attention to idiosyncrasies of person and circumstance, what precisely it was that led to progress in the past, and nobody can say what moves will succeed in the future.

Now it is of course possible to simplify the historical medium in which a scientist works by simplifying its main actors. The history of science, after all, consists not only of facts and conclusions drawn therefrom. It consists also of ideas, interpretations of facts, problems created by a clash of interpretations, actions of scientists, and so on. On closer analysis we even find that there are no "bare facts" at all but that the facts that enter our knowledge are already viewed in a certain way and are therefore essentially ideational. This being the case the history of science will be as complex, as chaotic, as full of error, and as entertaining as the ideas it contains and these ideas in turn will be as complex, as chaotic, as full of error, and as entertaining as are the minds of those who invented them. Conversely, a little brainwashing will go a long way in making the history of science more simple, more uniform, more dull, more "objective," and more accessible to treatment by "certain and infallible" rules: a theory of errors is superfluous when we are dealing with well-trained scientists who are kept in place by an internal slave master called "professional conscience" and who have been convinced that it is good and rewarding to attain, and then to forever keep, one's "professional integrity." [10]

Scientific education as we know it today has precisely this purpose. It has the purpose of carrying out a rationalistic simplification of the process "science" by simplifying its participants. One proceeds as follows. First, a domain of research is defined. Next, the domain is separated from the remainder of history (physics, for example, is separated from metaphysics and from theology) and receives a "logic" of its own.[11] A thorough training in such a logic then conditions those working in the domain so that they may not unwittingly disturb the purity (read: the sterility) that has already been achieved. An essential part of the training is the inhibition of intuitions that might lead to a blurring of boundaries. A person's religion, for example, or his metaphysics, or his sense of humor must not have the slightest connection with his scientific activity. His imagination is restrained[12] and even his language will cease to be his own.[13]

It is obvious that such an education, such a cutting up of domains and of consciousness, cannot be easily reconciled with a humanitarian attitude. It is in conflict "with the cultivation of individuality which [alone]

produces, or can produce well developed human beings";[14] it "maim[s] by compression, like a Chinese lady's foot, every part of human nature which stands out prominently, and tends to make a person markedly dissimilar in outline" [15] from the ideal of rationality that happens to be fashionable with the methodologists.

Now it is precisely such an ideal that finds expression either in "certain and infallible rules" or else in standards which separate what is correct, or rational, or reasonable, or "objective" from what is incorrect, or irrational, or unreasonable, or "subjective." Abandoning the ideal as being unworthy of a free man means abandoning standards and relying on theories of error entirely. Only these theories, these hints, these rules of thumb must now be renamed. Without universally enforced standards of truth and rationality we can no longer speak of universal error. We can only speak of what does, or does not, seem appropriate when viewed from a particular and restricted point of view, different views, temperaments, attitudes giving rise to different judgments and different methods of approach. Such an anarchistic epistemology—for this is what our theories of error now turn out to be—is not only a better means for improving knowledge, or of understanding history. It is also more appropriate for a free man to use than are its rigorous and "scientific" alternatives.

We need not fear that the diminished concern for law and order in science and society that is entailed by the use of anarchistic philosophies will lead to chaos. The human nervous system is too well organized for that.[16] Of course, there may arrive an epoch when it becomes necessary to give reason a temporary advantage and when it is wise to defend its rules to the exclusion of everything else. I do not think we are living in such an epoch today.

When we see that we have arrived at the utmost extent of human [understanding] we sit down contented. HUME[17]

The more solid, well defined, and splendid the edifice erected by the understanding, the more restless the urge of life . . . to escape from it into freedom. [Appearing as] reason it is negative and dialectical, for it dissolves into nothing the detailed determinations of the understanding. HEGEL[18]

Although science taken as whole is a nuisance, one can still learn from it. BENN[19]

1. Introduction; The Limits of Argument

The idea of a method that contains firm, unchanging, and absolutely binding principles for conducting the business of science gets into considerable difficulty when confronted with the results of historical research. We find, then, that there is not a single rule, however plausible, and however firmly grounded in epistemology, that is not violated at some time or other. It becomes evident that such violations are not accidental events, they are not the results of insufficient knowledge or of inattention which might have been avoided. On the contrary, we see that they are necessary for progress. Indeed, one of the most striking features of recent discussions in the history and philosophy of science is the realization that developments such as the Copernican Revolutions, or the rise of atomism in antiquity and recently (kinetic theory; dispersion theory; stereochemistry; quantum theory), or the gradual emergence of the wave theory of light occurred either because some thinkers decided not to be bound by certain "obvious" methodological rules or because they unwittingly broke them.[20]

This liberal practice, I repeat, is not just a *fact* of the history of science. It is not merely a manifestation of human inconstancy and ignorance. It is reasonable and *absolutely necessary* for the growth of knowledge. More specifically, the following can be shown: considering any rule, however "fundamental," there are always circumstances when it is advisable not only to ignore the rule, but to adopt its opposite. For example, there are circumstances when it is advisable to introduce, elaborate, and defend ad hoc hypotheses, or hypotheses which contradict well-established and generally accepted experimental results, or hypotheses whose content is smaller than the content of the existing and empirically adequate alternatives, or self-inconsistent hypotheses, and so on.[21]

There are even circumstances—and they occur rather frequently—when argument loses its forward-looking aspect and becomes a hindrance to progress. Nobody wants to assert[22] that the teaching of *small children* is exclusively a matter of argument (though argument may enter into it and should enter into it to a larger extent than is customary[23]), and almost everyone now agrees that what looks like a result of reason—the mastery of a language, the existence of a richly articulated perceptual world,[24] logical ability—is due partly to indoctrination, partly to a process of *growth* that proceeds with the force of natural law. And where arguments do seem to have an effect this must often be ascribed to their *physical repetition* rather than to their semantic content.[25] This much having been admitted, we must also concede the possibility of non-argumentative growth in the adult as well as in (the theoretical parts of)

institutions such as science, religion, and prostitution. We certainly cannot take it for granted that what is possible for a small child—to acquire new modes of behavior on the slightest provocation, to slide into them without any noticeable effort—is beyond the reach of his elders. One should expect that catastrophic changes of the physical environment, wars, the breakdown of encompassing systems of morality, political revolutions, will transform adult reaction patterns, too, including important patterns of argumentation.[26] This may again be an entirely natural process and rational argument may but increase the mental tension that precedes and causes the behavioral outburst.

Now, if there are events, not necessarily arguments, which cause us to adopt new standards, including new and more complex forms of argumentation, will it then not be up to the defenders of the status quo to provide, not just arguments, but also contrary causes? (Virtue without terror is ineffective, says Robespierre.) And if the old forms of argumentation turn out to be too weak a cause, must not these defenders either give up or resort to stronger and more "irrational" means? (It is very difficult, and perhaps entirely impossible, to combat the effects of brainwashing by argument.) Even the most puritanical rationalist will then be forced to stop reasoning and to use, say, *propaganda* and coercion, not because some of his reasons have ceased to be valid, but because the *psychological conditions* which make them effective, and capable of influencing others, have disappeared. And what is the use of an argument that leaves people unmoved?[27]

Of course, the problem never arises quite in this form. The teaching of standards never consists in merely putting them before the mind of the student and making them as clear as possible. The standards are supposed to have maximal *causal* efficacy as well. This makes it very difficult to distinguish between the *logical* force and the *material* effect of an argument. Just as a well-trained pet will obey his master no matter how great the confusion he finds himself in and no matter how urgent the need to adopt new patterns of behavior, in the very same way a well-trained rationalist will obey the mental image of his master, he will conform to the standards of argumentation he has learned, he will adhere to these standards no matter how great the difficulty he finds himself in, and he will be quite unable to discover that what he regards as the "voice of reason" is but a *causal aftereffect* of the training he has received. We see here very clearly how the appeal to "reason" works. At first sight this appeal seems to be to some ideas which convince a man instead of pushing him. But conviction cannot remain an ethereal state; it is supposed to lead to action. It is supposed to lead to the appropriate action, and it is supposed to sustain this action as long as necessary. What is the force that upholds such a development? It is the causal efficacy of the standards to which appeal was made and this causal efficacy in turn is but an effect of training, as we have seen. It follows that appeal to argument either has no content at all, and can be made to agree with any procedure,[28] or else will often have a conservative function: it will set limits to what is about to become a natural way of behavior.[29] In the latter case, however, the appeal is nothing but a concealed political maneuver. This becomes very clear when a rationalist wants to restore an earlier point of view. Basing his argument on natural habits of reasoning which either have become extinct or have no point of attack in the new situation, such a champion of "rationality" must first restore the earlier material and psychological conditions. This, however, involves him in "a struggle of interests and forces, not of argument." [30]

That interests, forces, propaganda, brainwashing techniques play a much greater role in the growth of our knowledge and, a fortiori, of science than is commonly believed can also be seen from an analysis of the relation between idea and action. One often takes it for granted that a clear and distinct understanding of new ideas precedes and should precede any formulation and any institutional expression of them. (An investigation starts with a problem, says Popper.) First, we have an idea, or a problem; then we act, i.e., either speak, or build, or destroy.[31] This is certainly not the way in which small children develop. They use words, they combine them, they play with them until they grasp a meaning that so far has been beyond their reach. And the initial playful activity is an essential presupposition of the final act of understanding.[32] There is no reason why this mechanism should cease to function in the adult. On the contrary, we must expect, for example, that the idea of liberty could be made clear only by means of the very same actions which were supposed to create liberty. Creation of a *thing*, and creation plus full understanding of a correct *idea* of the thing, very often are parts of one and the same indivisible process and they cannot be separated without bringing the process to a standstill. The process itself is not guided by a well-defined program; it cannot be guided by such a program for it contains the conditions of the realization of programs. It is rather guided by a

vague urge, by a "passion" (Kierkegaard). The passion gives rise to specific behavior which in turn creates the circumstances and the ideas necessary for analyzing and explaining the whole development, for making it "rational." [33]

The development of the Copernican point of view from Galileo up to the twentieth century is a perfect example of the situation we want to describe. We start with a strong belief that runs counter to contemporary reason. The belief spreads and finds support from other beliefs which are equally unreasonable, if not more so (law of inertia; telescope). Research now gets deflected in new directions, new kinds of instruments are built, "evidence" is related to theories in new ways until there arises a new ideology that is rich enough to provide independent arguments for any particular part of it and mobile enough to find such arguments whenever they seem to be required. Today we can say that Galileo was on the right track, for his persistent pursuit of what once seemed to be a silly cosmology created the material needed for the defense of this cosmology against those of us who accept a view only if it is told in a certain way and who trust it only if it contains certain magical phrases, called "observational reports." [34] And this is not an exception—it is the normal case: theories become clear and "reasonable" only after incoherent parts of them have been used for a long time. Such unreasonable, nonsensical, unmethodical foreplay thus turns out to be an unavoidable precondition of clarity and of empirical success.[35]

Trying to describe developments of this kind in a general way, we are of course obliged to appeal to the existing forms of speech which do not take them into account and which must be distorted, misused, and beaten into new patterns in order to fit unforeseen situations (without a constant misuse of language there cannot be any discovery and any progress). "Moreover, since the traditional categories are the gospel of everyday thinking (including ordinary scientific thinking) and of everyday practice, [such an attempt at understanding] in effect presents rules and forms of false thinking and action—false, that is, from the standpoint of [scientific] commonsense." [36] This is how *dialectical* thinking arises as a form of thought that "dissolves into nothing the detailed determinations of the understanding." [37]

It is clear, then, that the idea of a fixed method, or of a fixed (theory of) rationality, arises from too naive a view of man and of his social surroundings. To those who look at the rich material provided by history, and who are not intent on impoverishing it in order to please their lower instincts, their craving for intellectual security as it is provided, for example, by clarity and precision, to such people it will seem that there is only one principle that can be defended under all circumstances, and in *all* stages of human development. It is the principle: *anything goes*.[38]

This abstract principle (which is the one and only principle of our anarchistic methodology) must now be elucidated, and explained in concrete detail.

2. Counterinduction I: Theories

It was said that when considering any rule, however fundamental or "necessary for science," one can imagine circumstances when it is advisable not only to ignore the rule, but to adopt its opposite. Let us apply this claim to the rule that "experience," or "the facts," or "experimental results," or whatever words are being used to describe the "hard" elements of our testing procedures, measure the success of a theory, so that agreement between the theory and "the data" is regarded as favoring the theory (or as leaving the situation unchanged), while disagreement endangers or perhaps even eliminates it. This rule is an essential part of all theories of induction, including even some theories of corroboration. Taking the opposite view, I suggest introducing, elaborating, and propagating hypotheses which are inconsistent either with well-established theories or with well-established *facts*. Or, as I shall express myself: *I suggest proceeding counterinductively in addition to proceeding inductively.*

There is no need to discuss the first part of the suggestion which favors hypotheses inconsistent with well-established *theories*. The main argument has already been published elsewhere.[39] It may be summarized by saying that evidence that is relevant for the test of a theory T can often be unearthed only with the help of an incompatible alternative theory T'. Thus, the advice to postpone alternatives until the first refutation has occurred means putting the cart before the horse. In this connection, I also advised increasing empirical contents with the help of a *principle of proliferation*: invent and elaborate theories which are inconsistent with the accepted point of view, even if the latter should happen to be highly confirmed and generally accepted. Considering the arguments just summarized, such a principle would seem to be an essential part of any critical empiricism.[40]

The principle of proliferation is also an essential part of a humanitarian

outlook. Progressive educators have always tried to develop the individuality of their pupils, and to bring to fruition the particular and sometimes quite unique talents and beliefs that each child possesses. But such an education very often seemed to be a futile exercise in daydreaming. For is it not necessary to prepare the young for life? Does this not mean that they must learn one particular set of views to the exclusion of everything else? And, if there should still remain a trace of their youthful gift of imagination, will it not find its proper application in the arts, that is, in a thin domain of dreams that has but little to do with the world we live in? Will this procedure not finally lead to a split between a hated reality and welcome fantasies, science and the arts, careful description and unrestrained self-expression?[41] The argument for proliferation shows that this need not be the case. It is possible to *retain* what one might call the freedom of artistic creation and *to use it to the full*, not just as a road of escape, but as a necessary means for discovering and perhaps even changing the properties of the world we live in. For me this coincidence of the part (individual man) with the whole (the world we live in), of the purely subjective and arbitrary with the objective and lawful, is one of the most important arguments in favor of a pluralistic methodology.[42]

3. Philosophical Background: Mill, Hegel

The idea that a pluralistic methodology is necessary both for the advancement of knowledge and for the development of our individuality has been discussed by J. S. Mill in his admirable essay *On Liberty*. This essay, according to Mill, is "a kind of philosophical text book of a single truth, which the changes progressively taking place in modern society tend to bring out into ever stronger relief: the importance, to man and society, of a large variety in types of character, and of giving full freedom to human nature to expand itself in innumerable and conflicting directions." [43] Such variety is necessary both for the production of "well-developed human beings" (page 258) and for the improvement of civilization. "What has made the European family of nations an improving, instead of a stationary, portion of mankind? Not any superior excellence in them, which, when it exists, exists as the effect, not as the cause, but their remarkable diversity of character and culture. Individuals, classes, nations have been extremely unlike one another: they have struck out a great variety of paths, each leading to something valuable; and although at every period those who traveled in different paths have been intolerant of one another, and each would have thought it an excellent thing if all the rest would have been compelled to travel his road, their attempts to thwart each other's development have rarely had any permanent success, and each has in time endured to receive the good which the others have offered. Europe is, in my judgment, wholly indebted to this plurality of paths for its progressive and many-sided development" (pages 268–269).[44] The benefit to the individual derives from the fact that "[t]he human faculties of perception, judgment, discriminative feeling, mental activity, and even moral preference are exercised only in making a choice . . . [t]he mental and moral, like the muscular, powers are improved only by being used. The faculties are called into no exercise by doing a thing merely because others do it, no more than by believing a thing only because others believe it" (page 252). Choice presupposes alternatives between which to choose; it presupposes a society which contains and encourages "different opinions" (page 249), "antagonistic modes of thought,"[45] as well as "different experiments of living" (page 249), so that the "worth of different modes of life is proved not just in the imagination, but practically" (page 250).[46] "[U]nity of opinion," however, "unless resulting from the fullest and freest comparison of opposite opinions, is not desirable, and diversity not an evil, but a good . . ." (page 249).

This is how proliferation is introduced by Mill. It is not the result of a detailed epistemological analysis, or, what would be worse, of a linguistic examination of the usage of such words as "to know" and "to have evidence for." Nor is proliferation proposed as a solution to *epistemological* problems such as Hume's problem, or the problem of the testability of general statements. (The idea that *experience* might be a basis for our knowledge is at once removed by the remark that "[t]here must be discussion to show how experience is to be interpreted," page 208.) Proliferation is introduced as the solution to a problem of *life*: how can we achieve full consciousness; how can we learn what we are capable of doing; how can we increase our freedom so that we are able to decide, rather than adopt by habit, the manner in which we want to use our talents? Considerations like these were common at a time when the connection between truth and self-expression was still regarded as a problem, and when even the arts were supposed not just to please, but to elevate and to instruct.[47] Today the only question is how science can improve its own resources, no matter what the human effect of its meth-

ods and of its results. For Mill the connection still exists. Scientific method is part of a general theory of man. It receives its rules from this theory and is built up in accordance with our ideas of a worthwhile human existence.

In addition, pluralism is supposed to lead to the *truth*: ". . . the peculiar evil of silencing the expression of an opinion is that it is robbing the human race, posterity as well as the existing generation—those who dissent from the opinion, still more than those who hold it. If the opinion is right, they are deprived of the opportunity of exchanging error for truth; if wrong, they lose, what is almost as great a benefit, the clearer perception and livelier impression of truth produced by its collision with error" (page 205).[48] "The beliefs which we have most warrant for have no safeguard to rest on but a standing invitation to the whole to prove them unfounded" (page 209). If "with every opportunity for contesting it [a certain opinion, or a hypothesis] has not been refuted" (page 207), then we can regard it as better than another opinion that has "not gone through a similar process" (page 208).[49] "If even the Newtonian philosophy were not permitted to be questioned, mankind could not feel as complete assurance of its truth as they now do" (page 209). "So essential is this discipline to a real understanding of moral and human subjects [as well as of natural philosophy—page 208] that, if opponents of all-important truths do not exist, it is indispensable to imagine them and to supply them with the strongest arguments which the most skillful devil's advocate can conjure up" (page 228). There is no harm if such opponents produce positions which sound absurd and eccentric: "Precisely because the tyranny of opinion is such as to make eccentricity a reproach, it is desirable, in order to break through that tyranny, that people should be eccentric" (page 267).[50] Nor should those who "admit the validity of the arguments for free discussion[s] . . . object to their being 'pushed to an extreme' . . . unless the reasons are good for an extreme case, they are not good for any case" (page 210).[51] Thus methodological and humanitarian arguments are intermixed in every part of Mill's essay,[52] and it is on *both* grounds that a pluralistic epistemology is defended, for the natural as well as for the social sciences.[53]

One of the consequences of pluralism and proliferation is that stability of knowledge can no longer be guaranteed. The support a theory receives from observation may be very convincing; its categories and basic principles may appear well founded; the impact of experience itself may be extremely forceful. Yet there is always the possibility that new forms of thought will arrange matters in a different way and will lead to a transformation even of the most immediate impressions we receive from the world. Considering this possibility, we may say that the long-lasting success of our categories and the omnipresence of a certain point of view is not a sign of excellence or an indication that the truth or part of the truth has at last been found. *It is, rather, the indication of a failure* of reason to find suitable alternatives which might be used to transcend an accidental intermediate stage of our knowledge. This remark leads to an entirely new attitude toward success and stability.

As far as one can see, the aim of all methodologies is to find principles and facts which, if possible, are not subjected to change. Principles which give the impression of stability are, of course, tested. One tries to refute them, at least in some schools. If all attempts at refutation fail, we have a *positive* result, nevertheless: we have succeeded in discovering a new stable feature of the world that surrounds us; we have come a step closer to the truth.

Moreover, the process of refutation itself rests on assumptions which are not further investigated. An instrumentalist will assume that there are stable facts, sensations, everyday situations, classical states of affairs, which do not change, not even as the result of the most revolutionary discovery. A "realist" may admit changes of the observational matter, but he will insist on the separation between subject and object and he will try to restore it wherever research seems to have found fault with it.[54] Believing in an "approach to the truth," he will also have to set limits to the development of concepts. For example, he will have to exclude incommensurable concepts from a series of succeeding theories.[55] This is the traditional attitude, up to, *and including*, Popper's critical rationalism.

As opposed to it, the attitude about to be discussed regards any prolonged stability, either of ideas and impressions which are capable of test or of background knowledge which one is not willing to give up (realism; separation of subject and object; commensurability of concepts), as an indication of *failure*, pure and simple. Any such stability indicates *that we have failed to transcend an accidental stage of knowledge, and that we have failed to rise to a higher stage of consciousness and of understanding*. It is even questionable whether we can still claim to possess knowledge in such a state. As we become familiar with the existing categories and with the alternatives that are being used in the examination

of the received view, our thinking loses its spontaneity until we are reduced to the "bestial and goggle-eyed contemplation of the world around us." [56] "The more solid, well defined, and splendid the edifice erected by the understanding, the more restless the urge of life to escape from it into freedom." [57] Each successful refutation, by opening the way to a new and as yet untried system of categories, temporarily returns to the mind the freedom and spontaneity that is its essential property.[58] But complete freedom is never achieved. For any change, however dramatic, always leads to a new system of fixed categories. Things, processes, states are still separated from each other. The existence of different elements, of a manifold, is still "exaggerated into an opposition by the understanding." [59]

This "evil manner of reflection,[60] to always work with fixed categories," [61] is extended by the customary modes of research to the most widely presupposed and unanalyzed opposition between a subject and an entirely different world of objects.[62] The following assumptions which are important for a methodological realism have been made in this connection: "the object . . . is something finished and perfect that does not need the slightest amount of thought in order to achieve reality while thought itself is . . . something deficient that needs . . . material for its completion[63] and must be soft enough to adapt itself to the material in question." [64] "If thought and appearance do not completely correspond to each other, one has, to start with, a choice: the one or the other may be at fault. [Scientific empiricism] blames thought for not adequately mirroring experience . . ." [65] "These are the ideas which form the core of our customary views concerning the relation between subject and object," [66] and they are responsible for whatever immobility remains in science, even at a time of crisis.

How can this immobility be overcome? How can we obtain insight into the most fundamental assumptions, not only of science and common sense, but of our existence as thinking beings as well? Insight cannot be obtained as long as the assumptions form an unreflected and unchanging part of our life. But, if they are allowed to change, can we then finish the task of criticizing as identically the same persons who started it? Problems like these are raised not only by the abstract question of criticism, but also by more recent discoveries in anthropology, history of science, and methodology. I shall return to them when I discuss incommensurable theories. For the moment, I would like to indicate, very briefly, how certain ideas of Hegel can be used to get a tentative first answer, and thus to make a first step in our attempt to reform the sciences.

Science, common sense, and even the refined common sense of critical rationalism use certain fixed categories ('subject'; 'object'; 'reality'), in addition to the many changing views they contain. They are therefore not fully rational. Full rationality can be obtained by extending criticism to the stable parts also. This presupposes the invention of alternative categories and their application to the whole rich material at our disposal. The categories, and all other stable elements of our knowledge, must be set in motion. "Our task is to make fluid the petrified material which we find, and to relight [*wieder entzuenden*] the concepts contained in this dead stuff." [67] We must "dissolve the opposition of a frozen subjectivity and objectivity and comprehend the origin of the intellectual and real world as a becoming, we must understand their being as a product, as a form of producing." [68] Such dissolving is carried out by reason, which is "the force of the negative absolute, that is, an absolute negation," [69] that "annihilates" [70] science and common sense, and the state of consciousness associated with both. This annihilation is not a conscious act of a scientist who has *decided* to eliminate some basic distinctions in his field. For although he may consciously try to overcome the limitations of a particular stage of knowledge, he may not succeed for want of objective conditions (in his brain, in his social surroundings, in the physical world[71]) favoring his wish.[72] Hegel's general theory of development, his cosmology, as one might call it, gives an account of such conditions.

According to this cosmology, every object, every determinate being, is related to everything else: "a well determined being, a finite entity is one that is related to others; it is a content that stands in the relation of necessity to another content and, in the last resort, to the world. Considering this mutual connectedness of the whole, metaphysics could assert . . . the tautology that the removal of a single grain of dust must cause the collapse of the whole universe." [73] The relation is not external. Every process, object, state, etc., actually *contains* part of the nature of every other process, object, state, etc.[74] *Conceptually* this means that the complete description of an object is self-contradictory. It contains elements which say what the object *is*; these are the elements used in the customary accounts provided by science and by common sense which consider part of its properties only and ascribe the rest to the outside.

And the description also contains other elements which say what the object *is not*. These are the elements which science and common sense put outside the object, attributing them to things which are supposed to be completely separated but which are actually contained in the object under consideration. The result is that "all things are beset by an internal contradiction." [75] This contradiction cannot be eliminated by using different words, for example, by using the terminology of a process and its *modifications*. For the process will again have to be separated, at least in thought, from something other than itself; otherwise it is pure being which is in no way different from pure *nothingness*.[76] It will contain part of what it is separated from, and this part will have to be described by ideas inconsistent with the ideas used for describing the original process, which therefore is bound to contain contradictions also.[77] Hegel has a marvelous talent for making visible the contradictions which arise when we examine a concept in detail, wishing to give a complete account of the state of affairs it describes. "Concepts which usually appear stable, unmoved, dead are analyzed by him and it becomes evident that they move." [78]

Now we come to a second principle of Hegel's cosmology. The motion of concepts is not merely a motion of the *intellect*, which, starting the analysis with certain determinations, moves away from them and posits their negation. It is an *objective* development as well, and it is caused by the fact that every finite (well-determined, limited) object, process, state, etc., has the tendency to emphasize the elements of the other objects present in it, and to become what it is not. The object, "being restless within its own limit," [79] "strives not to be what it is." [80] "Calling things *finite*, we want to say that they are not merely limited . . . but rather that the negative is essential to their nature and to their being . . . Finite things are, but the truth of their being is their *end*.[81] What is finite does not merely change . . . it passes away; nor is this passing away merely possible, so that the finite thing could continue to be, without passing away; quite the contrary, the being of a finite thing consists in its having in itself the seeds of passing away . . . the hour of its birth is the hour of its death." [82] "What is finite, therefore, can be set in motion." [83]

Moving beyond the limit, the object ceases to be what it is and becomes what it is not; it is negated. A third principle of Hegel's cosmology is that the result of the negation is "not a mere nothing; it is a spe-

cial content, for . . . it is the negation of a determined and well defined thing." [84] Conceptually speaking, we arrive at a "new concept which is higher, richer than the concept that preceded it, for it has been enriched by its negation or opposition, contains it as *well as* its negation, being the unity of the original concept and of its opposition." [85] This is an excellent description, for example, of the transition from the Newtonian conception of space to that of Einstein, *provided* we continue using the *unchanged* Newtonian concept.[86] "It is clear that no presentation can be regarded as scientific that does not follow the path and simple rhythm of this method, for this is the path pursued by the things themselves." [87]

Considering that the motion beyond the limit is not arbitrary, but is directed "towards its [i.e., the object's] end" [88] it follows that not *all* the aspects of other things which are present in a certain object are realized in the next stage. Negation, accordingly, "does not mean simply saying No, or declaring a thing to be non-existent, or destroying it in any way one may choose . . . Each kind of thing . . . has its own peculiar manner of becoming negated, and in such a way that a development results from it, and the same holds good for each type of ideas and conceptions . . . *This must be learned like* everything else." [89] What has to be learned, too, is that the "negation of the negation" does not lead further away from the original starting point but that it returns to it.[90] This is an "extremely universal and just on that account extremely far-reaching and important law of development in nature, history and thought; a law which . . . asserts itself in the plant and animal world, in geology, in mathematics,[91] in history, in philosophy." [92] Thus for example "a grain of barley falling under suitable conditions on suitable soil disappears, is negated, and in its place there arises out of it the plant, the negation of the grain . . . This plant grows, blossoms, bears fruit and finally produces other grains of barley, and as soon as these ripen, the stalk dies, is in turn negated. As a result of this negation of the negation, we again have the grain of barley we started with, not singly, but rather in ten, twenty or thirtyfold number . . . and perhaps even qualitatively improved . . ." [93] "It is evident that I say nothing whatever about the particular process of development which, for example, the grain of barley undergoes from its germination up to the dying off of the fruit-bearing plant, when I state that it is the negation of the negation . . . I rather comprise these processes altogether under this one *law* of motion and

just for that reason disregard the peculiarities of each special process. *Dialectics*, however, is nothing else than the science of the general laws of motion and development in nature, human society and thought." [94]

In the foregoing account, concepts and real things have been treated as separate. Similarities and correspondences were noted: each thing contains elements of everything else, it *develops* by turning into these alien elements, it changes, and it finally tries to *return* to itself. The *notion* of each thing, accordingly, contains contradictory elements. It is negated, and it moves in a way corresponding to the movement of the thing. This presentation has one serious disadvantage: "Thought is here described as a mere subjective and formal activity while the world of objects, being situated vis-à-vis thought, is regarded as something fixed and as having independent existence. This dualism . . . is not a true account of things and it is pretty thoughtless to simply take over the said properties of subjectivity and objectivity without asking for their origin . . . Taking a more realistic view we may say that the subject is only a stage in the development of being and essence." [95] The concept, too, is then part of the general development of nature, in a materialistic interpretation of Hegel. "Life," for example, "or organic nature is that phase of nature when the concept appears on the stage; it enters the stage as a blind concept that does not comprehend itself, i.e., does not think." [96] Being part of the *natural behavior*, first of an organism, then of a thinking being, it not only mirrors a nature that "lies entirely outside of it," [97] it is not merely "something subjective and accidental," [98] it is not "merely a concept";[99] it participates in the general nature of all things, i.e., it contains an element of everything else, it has the tendency to be the end result of the development of a specific thing, so that, finally, the concept and this thing become one.[100] "That real things do not agree with the idea ["read: with the total knowledge of man"[101]] constitutes their *finitude*, their *untruth* because of which they are *objects*, each determined in its special sphere by the laws of mechanics, chemistry, or by some external purpose." [102] In this stage "there can be nothing more detrimental and more unworthy of a philosopher than to point, in an entirely vulgar fashion, to some experience that contradicts the idea . . . When something does not correspond to its concept, it must be led up to it" [103] (counterinduction!) until "concept and thing have become one." [104]

To sum up: Knowledge is part of nature and is subjected to its general laws. The laws of dialectics apply to the motion of objects and concepts, as well as to the motion of higher units comprising objects and concepts. According to these general laws, every object participates in every other object and tries to change into its negation. This process cannot be understood by attending to those elements in our subjectivity which are still in relative isolation and whose internal contradictions are not yet revealed. (Most of the customary concepts of science, mathematics, and especially the rigid categories used by our modern axiomaniacs are of this kind.) To understand the process of negation we must attend to those other elements which are fluid, about to turn into their opposites, and which may, therefore, bring about knowledge and truth, "the identity of thing and concept." [105] The identity itself cannot be achieved mechanically, i.e., by arresting some aspect of reality and fiddling about with the remaining aspects, or theories, until agreement is achieved (the aspects one wants to arrest, being in motion, will soon be replaced by dogmatic opinions of them, rigid perceptions included). We must rather proceed dialectically, i.e., by an interaction of concept and fact (observation, experiment, basic statement, etc.) that affects *both* elements. The lesson for methodology is, however, this: Do not work with stable concepts. Do not eliminate counterinduction. Do not be seduced into thinking that you have at last found the correct description of "the facts" when all that has happened is that some new categories have been adapted to some older forms of thought, which are so familiar that we take their outlines to be the outlines of the world itself.

4. Counterinduction II: Experiments, Observations, "Facts"

Considering now the invention, the use, and the elaboration of theories which are inconsistent, not only with other theories, but even with *experiments, facts, observations*, we may start by pointing out that *not a single theory ever agrees with all the known facts in its domain*. And the trouble is not created by rumors, or by the results of sloppy procedure. It is created by experiments and measurements of the highest precision and reliability.

It will be convenient, at this place, to distinguish two different kinds of disagreement between theory and fact: numerical disagreements and qualitative failures.

The first case is quite familiar: a theory makes a certain numerical prediction and the value that is actually obtained differs from the prediction

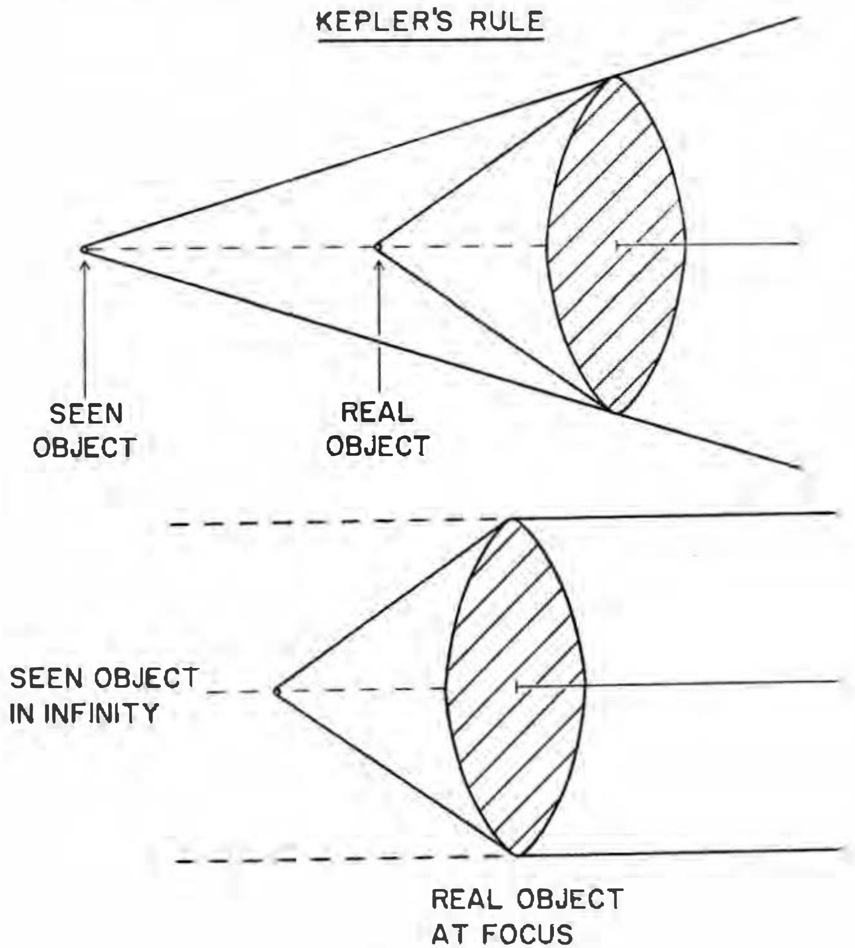
made outside the margin of error. Precision instruments are usually involved here. Numerical disagreements abound in science.

Thus the Copernican view at the time of Galileo was inconsistent with facts so plain and obvious that Galileo had to call it "surely false." [106] "There is no limit to my astonishment," he writes in a later work,[107] "when I reflect that Aristarchus and Copernicus were able to make reason so conquer sense that, in defiance of the latter, the former became mistress of their belief." Newton's theory of gravitation was beset, from the very beginning, by a considerable number of difficulties which were serious enough to provide material for refutations. Even today, and in the nonrelativistic domain, there exist "numerous discrepancies between observation and theory." [108] Bohr's atomic model was introduced and retained in the face of very precise and unshakable contrary evidence.[109] The special theory of relativity was retained, despite D. C. Miller's decisive refutation. (I call the refutation "decisive" because the experiment was, from the point of view of contemporary evidence, at least as well performed as the earlier experiment of Michelson and Morley.[110]) The general theory of relativity, though surprisingly successful in some domains, failed to explain about 10″ in the movement of the nodes of Venus and more than 5″ in the movement of the perihelion of Mars. All these are quantitative difficulties which can be resolved by discovering a better set of numbers but which do not force us to make qualitative adjustments.

The second case, the case of qualitative failures, is less familiar, but of much greater interest. In this case a theory is inconsistent not with a recondite fact that must be unearthed with the help of complex equipment and is known to experts only, but with circumstances which can be noticed with the unaided senses and which are familiar to everyone.

The first and to my mind the most important example of an inconsistency of this kind is Parmenides' theory of the unchanging One. The theory has much in its favor[111] and it plays its role even today, for example in the general theory of relativity.[112] Used in an undeveloped form by Anaximander it led to the insight, repeated by Heisenberg in his theory of elementary particles,[113] that the basic substance, or the basic elements of the universe, cannot obey the same laws as do the visible elements. Zeno's arguments, on the other hand, show the difficulties inherent in the idea of a continuum consisting of isolated elements. Aristotle took these arguments seriously and developed his own theory of the continuum.[114] Yet the idea of a collection of elements remained and continued to be used, despite the quite obvious difficulties, until these difficulties were removed early in the twentieth century.[115]

Another example of a theory with qualitative defects is Newton's theory of colors. According to this theory light consists of rays of different refrangibility which can be separated, reunited, refracted, but which are never changed in their internal constitution, and which have a very small lateral extension in space. Considering that the surface of mirrors is much rougher than is the lateral extension of the rays, the ray theory is found to be inconsistent with the existence of mirror images (as is admitted by Newton himself: Opticks, book II, part III, proposition viii): if light con-



sists of rays, then a mirror should behave like a rough surface, i.e., it should look to us like a wall. Newton retains his theory, eliminating the difficulty with the help of an ad hoc hypothesis: "the reflection of a ray is effected, not by a single point of the reflecting body, but by some power of the body which is evenly diffused all over its surface . . ." [116]

In Newton's case the qualitative discrepancy between theory and fact is removed by an ad hoc hypothesis. In other cases not even this very flimsy maneuver is used. One retains the theory and tries to forget its shortcomings. An example is the attitude toward Kepler's rule according to which an object seen through a lens is perceived at the distance at which the rays traveling from the lens toward the eye intersect (see the first diagram).[117] The rule implies that an object situated at the focus will be seen infinitely far away (see the second diagram). "But on the contrary," writes Barrow, Newton's teacher and predecessor in Cambridge, commenting on this predication, "we are assured by experience that [a point situated close to the focus] appears variously distant, according to the different situations of the eye . . . And it does almost never seem farther off than it would be if it were beheld with the naked eye; but, on the contrary, it does sometime appear much nearer . . . All which does seem repugnant to our principles. But for me," Barrow continues, "neither this nor any other difficulty shall have so great an influence on me, as to make me renounce that which I know to be manifestly agreeable to reason." [118]

Barrow mentions the qualitative difficulties, and he says that he will retain the theory nevertheless. This is not the usual procedure. The usual procedure is to forget about the difficulties, never to talk about them, and to proceed as if the theory were without fault. This attitude is very common today.

Thus classical electrodynamics contains the absurd consequence that the motion of a free particle is self-accelerated.[119] This consequence is little known though it makes it impossible to calculate even the simplest case of a motion in a homogeneous electric field. What one does is to make "an approximation" which neglects effects too small to be noticed but which also eliminates the quite noticeable absurd consequence. Theory plus "approximation" produces a reasonable prediction though the theory itself suffers from qualitative difficulties. The quantum theory of fields which one might want to consult in order to remove the troubles of classical electromagnetic theory has absurdities of its own such as the infinite self energies. The situation is not improved by the remark that these self energies can be corrected by renormalization. They can of course be corrected by this method, and in a consistent manner, but only after redefining certain terms in the calculations with an eye to the results to be achieved. This procedure, which is ad hoc, certainly does not establish the excellence of the theory; it shows that as it stands the theory is either refuted[120] or else woefully incomplete.

Another example of modern physics is quite instructive, for it might have led to an entirely different development of our knowledge concerning the microcosm. Ehrenfest has proved a theorem[121] according to which the classical electron theory of Lorentz taken together with the equipartition principle excludes induced magnetism. The reasoning is exceedingly simple: according to the equipartition principle the probability of a motion is proportional to $\exp[-U/kT]$, where U is the energy of the motion. Now the energy in a constant magnetic field is, according to Lorentz, $= q(E + [vB]) \cdot v$, where q is the charge of the moving parts, E the electric field, B the magnetic field, v the velocity of the moving parts. This magnitude reduces to qE in all cases unless one is prepared to admit the existence of single magnetic poles (given the proper context, this result strongly supports the ideas and the experimental findings of the late Felix Ehrenhaft[122]).

Occasionally it is impossible to survey all the interesting consequences and to discover the absurd results of a theory. This may be due to a deficiency in the existing mathematical methods; it may also be due to the ignorance of those who defend the theory.[123] Under such circumstances the most common procedure is to use an older theory up to a certain point (which is often quite arbitrary) and to add the new theory for calculating refinements. Seen from a methodological point of view the procedure is a veritable nightmare. Let us explain it, using the relativistic calculation of the path of Mercury as an example.

The perihelion of Mercury moves along at a rate of about 5600″ per century. Of this value, 5026″ are geometric, having to do with the movement of the reference system; 575″ are dynamical, due to perturbations in the solar system. Of these perturbations all but the famous 43″ are accounted for by classical celestial mechanics. And the remaining 43″ are accounted for by general relativity. This is how the situation is usually explained.

Now this explanation shows that the premise from which we derive

the 43″ is not the general theory of relativity plus suitable initial conditions. The premise contains classical physics in addition to whatever relativistic assumptions are made. Furthermore, the relativistic calculation, the so-called "Schwarzschild solution," does not deal with the planetary system as it exists in the real world (i.e., our own asymmetric galaxy); it deals with the entirely fictional case of a central symmetrical universe containing a singularity in the middle and nothing else. What are the reasons for employing such an insane conjunction of premises?[124]

One reason, so the customary reply continues, is that we are dealing with approximations. The formulas of classical physics do not appear because relativity is incomplete. Nor is the central symmetrical case used because relativity does not offer anything better. Both schemata flow from the general theory under the special circumstances realized in our planetary system *provided* we omit magnitudes too small to be considered. Hence, we are using the theory of relativity throughout, and we are using it in an adequate manner.

Note, now, how this idea of an approximation differs from the legitimate idea: usually one has a theory, one is able to calculate the particular case one is interested in, one notes that this calculation leads to magnitudes below experimental precision, one omits such magnitudes, and one obtains a vastly simplified formalism. In the present case making the required approximations would mean calculating the n-body problem relativistically, omitting magnitudes smaller than the precision of observation reached, and showing that the theory thus curtailed coincides with classical celestial mechanics as corrected by Schwarzschild. This procedure has not been used by anyone simply because the relativistic n-body problem has as yet withstood solution.[125] There are not even approximate solutions for important problems such as, for example, the problem of stability (a first great stumbling stone for Newton's theory). This being the case, the classical part of the explanans is not only used for convenience, *it is absolutely* necessary. And the approximations made are not a result of relativistic calculation, they are introduced in order to make relativity fit the case. One may properly call them ad *hoc* approximations.

Ad hoc approximations abound in modern mathematical physics. They play a very important part in the quantum theory of fields and they are an essential ingredient of the correspondence principle. At the moment we are not concerned with the reasons for this fact, we are only concerned with its consequences: ad hoc approximations conceal, and even entirely eliminate, qualitative difficulties. They create a false impression of the excellence of our science. It follows that a philosopher who wants to study the adequacy of science as a picture of the world, or who wants to build up a realistic scientific methodology, must look at modern science with special care. In most cases modern science is more opaque and much more deceptive than its sixteenth- and seventeenth-century ancestors have ever been.

As a final example of qualitative difficulties we mention again the heliocentric theory at the time of Galileo. We shall soon have occasion to show that this theory was inadequate both qualitatively and quantitatively, and that it was also philosophically absurd.

To sum up this brief and very incomplete list: Wherever we look, whenever we have a little patience and select our evidence in an unprejudiced manner, we find that theories fail to adequately reproduce certain quantitative results; and that they are *qualitatively* incompetent to a surprising degree.[128] Science gives us theories of high beauty and sophistication. Modern science has developed mathematical structures which exceed anything that has existed so far in coherence and generality. But in order to achieve this miracle all the existing troubles had to be pushed into the relation between theory and fact, and had to be concealed, by ad hoc approximations, and by other procedures.

This being the case—what shall we make of the methodological demand that a theory must be judged by experience and must be rejected if it contradicts accepted basic statements? What attitude shall we adopt toward the various theories of confirmation and corroboration which all rest upon the assumption that theories can be made to completely agree with the known facts and which use the amount of agreement reached as a principle of evaluation? This demand, these theories, are now all quite useless. They are as useless as a medicine that heals a patient only if he is bacteria free. In practice they are never obeyed by anyone. Methodologists may point to the importance of falsifications—but they blithely use falsified theories; they may sermonize how important it is to consider all the relevant evidence, and never mention those big and drastic facts which show that the theories which they admire and accept, the theory of relativity, the quantum theory, are at least as badly off as the older theories which they reject. In practice methodologists slavishly repeat the most recent pronouncements of the top dogs in physics, though in doing

so they must violate some very basic rules of their trade. Is it possible to proceed in a more reasonable manner? Let us see!

According to Hume theories cannot be derived from facts. The demand to admit only those theories which follow from facts leaves us without any theory. Hence, a science *as we know it* can exist only if we drop the demand and revise our methodology.

According to our present results hardly any theory is consistent with the facts. The demand to admit only those theories which are consistent with the available and accepted facts again leaves us without any theory. (I repeat: without *any* theory, for there is not a single theory that is not in some trouble or other.) Hence, a science as we know it can exist only if we drop this demand also and again revise our methodology, now admitting counterinduction in addition to admitting unsupported hypotheses. The right method no longer consists of rules which permit us to choose between theories on the basis of falsifications. It must rather be modified so as to enable us to choose between theories which we have already tested and which are falsified.

To proceed further. Not only are facts and theories in constant disharmony, they are not even as neatly separated as everyone makes them out to be. Methodological rules speak of "theories" and "observations" and "experimental results" as if these were clear and well-defined objects whose properties are easy to evaluate and which are understood in the same sense by all scientists.

However, the material which a scientist actually has at his disposal, his laws, his experimental results, his mathematical techniques, his epistemological prejudices, his attitude toward the absurd consequences of the theories which he accepts, is indeterminate in many ways, it is ambiguous, *and never fully separated from the historical background*. This material is always contaminated by principles which he does not know and which, if known, would be extremely hard to test. Questionable views on cognition, such as the view that our senses, used in normal circumstances, give reliable information about the world, may invade the observation language itself, constituting the observational terms and the distinction between veridical and illusory appearances. As a result observation languages may become tied to older layers of speculation which affect, in this roundabout fashion, even the most progressive methodology. (Example: the absolute space-time frame of classical physics which was codified and consecrated by Kant.) The sensory impression, however simple, always contains a component that expresses the reaction of the perceiving subject and has no objective correlate. This subjective component often merges with the rest, and forms an unstructured whole which must then be subdivided from the outside with the help of counterinductive procedures. (An example of this is the appearance of a fixed star to the naked eye, which contains the subjective effects of irradiation, diffraction, diffusion, restricted by the lateral inhibition of adjacent elements of the retina.) Finally, there are the auxiliary premises which are needed for the derivation of testable conclusions, and which occasionally form entire *auxiliary sciences.*

Consider the case of the Copernican hypothesis, whose invention, defense, and partial vindication run counter to almost every methodological rule one might care to think of today. The auxiliary sciences here contained laws describing the properties and the influence of the terrestrial atmosphere (meteorology); optical laws dealing with the structure of the eye and telescopes, and with the behavior of light; and dynamical laws describing motion in moving systems. Most importantly, however, the auxiliary sciences contained a theory of cognition that postulated a certain simple relation between perceptions and physical objects. Not all these auxiliary disciplines were available in explicit form. Many of them merged with the observation language, and led to the situation described at the beginning of the preceding paragraph.

Consideration of all these circumstances, of observation terms, sensory core, auxiliary sciences, background speculation, suggests that a theory may be inconsistent with the evidence, not because it is not correct, *but because the evidence is contaminated.* The theory is threatened either because the evidence contains unanalyzed sensations which only partly correspond to external processes, or because it is presented in terms of antiquated views, or because it is evaluated with the help of backward auxiliary subjects. The Copernican theory was in trouble for *all* these reasons.

It is this *historico-physiologic character of the evidence,*[127] the fact that it does not merely describe some objective state of affairs, *but* also expresses some *subjective, mythical, and long-forgotten* views concerning this state of affairs, that forces us to take a fresh look at methodology. It shows that it would be extremely imprudent to let the evidence judge our theories directly, and without any further ado. A straightforward and unqualified judgment of theories by "facts" is bound to eliminate ideas

simply because they do not fit into the framework of some older cosmology. Taking experimental results and observations for granted and putting the burden of proof on the theory means taking the observational ideology for granted without having ever examined it. (Note that the experimental results are supposed to have been obtained with the greatest possible care. Hence "taking observations, etc., for granted" means "taking them for granted after the most careful examination of their reliability"—for even the most careful examination of an observation statement does not interfere with the concepts in terms of which it is expressed, or with the structure of the sensory expression.)

Now—how can we possibly examine something we are using all the time? How can we criticize the terms in which we habitually express our observations? Let us see![128]

The first step in our criticism of commonly used concepts is to create a measure of *criticism,* something with which these concepts can be compared. Of course, we shall later want to know a little more about the measure stick itself, for example, we shall want to know whether it is better than, or perhaps not as good as, the material examined. But in order for *this* examination to start there must be a measure stick in the first place. Therefore the first step in our criticism of customary concepts and customary reactions is to step outside the circle and to invent a new conceptual system, a new theory, for example, that clashes with the most carefully established observational results and confounds the most plausible theoretical principles. This step is, again, counterinductive. Counterinduction, therefore, is both a fact—science could not exist without it—and a legitimate and much-needed *move* in the game of science.

5. The Tower Argument Stated: First Steps of Analysis

As a concrete illustration and as a basis for further discussion, I shall now briefly describe the manner in which Galileo defused an important counterargument against the idea of the motion of the earth. I say "defused," and not "refuted," because we are dealing with a changing conceptual system as well as with certain attempts at concealment.

According to the argument which convinced Tycho, and which is used against the motion of the earth in Galileo's own *Trattato della sfera*, observation shows that "heavy bodies . . . falling down from on high, go by a straight and vertical line to the surface of the earth. This is considered an irrefutable argument for the earth being motionless. For if it made the diurnal rotation, a tower from whose top a rock was let fall, being carried by the whirling of the earth, would travel many hundreds of yards to the east in the time the rock would consume in its fall, and the rock ought to strike the earth that distance away from the base of the tower." [129]

In considering the argument, Galileo at once admits the correctness of the sensory content of the observation made, viz. that "heavy bodies . . . falling from a height, go perpendicularly to the surface of the earth." [130] Considering an author (Chiaramonti) who sets out to convert Copernicans by repeatedly mentioning this fact, he says: "I wish that this author would not put himself to such trouble trying to have us understand from our senses that this motion of falling bodies is simple straight motion and no other kind, nor get angry and complain because such a clear, obvious, and manifest thing should be called into question. For in this way he hints at believing that to those who say such motion is not straight at all, but rather circular, it seems they see the stone move visibly in an arc, since he calls upon their senses rather than their reason to clarify the effect. This is not the case, Simplicio; for just as I . . . have never seen nor ever expect to see the rock fall any way but perpendicularly, just so do I believe that it appears to the eyes of everyone else. It is therefore better to put aside the appearance, on which we all agree, and to use the power of reason either to confirm its reality or to reveal its fallacy." [131] The correctness of the observation is not in question. What is in question is its "reality" or "fallacy." What is meant by this expression?

The question is answered by an example that occurs in Galileo's next paragraph, and "from which . . . one may learn how easily anyone may be deceived by simple appearances, or let us say by the impressions of one's senses. This event is the appearance to those who travel along a street by night of being followed by the moon, with steps equal to theirs, when they see it go gliding along the eaves of the roofs. There it looks to them just as would a cat really running along the tiles and putting them behind it; an appearance which, if reason did not intervene, would only too obviously deceive the senses."

In this example we are asked to start with a sensory impression and consider a statement that is forcefully suggested by it. (The suggestion is so strong that it has led to entire systems of belief and rituals as becomes clear from a closer study of the lunar aspects of witchcraft and of other religions.) Now "reason intervenes": the statement suggested by the impression is examined, and one considers other statements in its place. The nature of the impression is not changed a bit by this activity. (This is only approximately true; but we can omit for our present purpose the complications arising from the interaction of impression and proposition.) But it enters new observation statements and plays new, better or worse, parts in our knowledge. What are the reasons and the methods which regulate such exchange?

To start with we must become clear about the nature of the total phenomenon: appearance plus statement. There are not two acts, one, noticing a phenomenon, the other, expressing it with the help of the appropriate statement, *but* only one, viz. saying, in a certain observational situation, "the moon is following me," or "the stone is falling straight down." We may of course abstractly subdivide this process into parts, and we may also try to create a situation where statement and phenomenon seem to be psychologically apart and waiting to be related. (This is rather difficult to achieve and is perhaps entirely impossible.[132]) But under normal circumstances such a division does not occur; describing a familiar situation is, for the speaker, an event in which statement and phenomenon are firmly glued together.

This unity is the result of a process of learning that starts in one's childhood. From our very early days we learn to react to situations with the appropriate responses, linguistic or otherwise. The teaching procedures both shape the 'appearance' or the 'phenomenon' and establish a firm connection with words, so that finally the phenomena seem to speak for themselves, without outside help or extraneous knowledge. They just are what the associated statements assert them to be. The language they 'speak' is of course influenced by the beliefs of earlier generations which have been held for such a long time that they no longer appear as separate principles, but enter the terms of everyday discourse, and, after the prescribed training, seem to emerge from the things themselves.

Now at this point we may want to compare, in our imagination and quite abstractly, the results of the teaching of different languages incorporating different ideologies. We may even want to consciously change some of these ideologies and adapt them to more 'modern' points of view. It is very difficult to say how this will change our situation, unless we make the further assumption that the quality and structure of sensations (perceptions), or at least the quality and structure of those sensations which enter the body of science, are independent of their linguistic expression.

I am very doubtful about even the approximate validity of this assumption which can be refuted by simple examples. And I am sure that we are depriving ourselves of new and surprising discoveries as long as we remain within the limits defined by it. Yet the present essay will remain quite consciously within these limits. (My first task, if I should ever resume writing, would be to explore these limits and to venture beyond them.)

Making the additional simplifying assumption, we can now distinguish between (a) sensations, and (b) those "mental operations which follow so closely upon the senses" [133] and are so firmly connected with their reactions that a separation is difficult to achieve. Considering the origin and the effect of such operations, I shall call them *natural interpretations.*

6. Natural Interpretations

In the history of thought, natural interpretations have been regarded either as *a priori presuppositions* of science or else as *prejudices* which must be removed before any serious examination can proceed. The first view is that of Kant, and, in a very different manner and on the basis of very different talents, that of some contemporary linguistic philosophers. The second view is due to Bacon (who had, however, predecessors, such as the Greek skeptics).

Galileo is one of those rare thinkers who neither wants to forever *retain* natural interpretations nor wants to altogether *eliminate* them. Wholesale judgments of this kind are quite alien to his way of thinking. He insists upon *critical discussion* to decide which natural interpretations can be kept and which must be replaced. This is not always clear from his writings. Quite the contrary, the methods of reminiscence, to which he appeals so freely, are designed to create the impression that nothing has changed and that we continue expressing our observations in old and familiar ways. Yet his attitude is relatively easy to ascertain: natural interpretations are necessary. The senses alone, without the help of reason, cannot give us a true account of nature. What is needed for arriving at such a true account are "the . . . senses, accompanied *by* reasoning." [134] Moreover, in the arguments dealing with the motion of the earth, it is this reasoning, it is the connotation of the observation terms, and not the message of the senses or the appearance, that causes trouble. "It is therefore better to put aside the appearance, on which we all agree, and to use the power of reason either to confirm [its] reality or to reveal [its]

fallacy." [135] "To confirm the reality or reveal the fallacy of appearances" means, however, to examine the validity of those natural interpretations which are so intimately connected with the appearances that we no longer regard them as separate assumptions. I now turn to the first natural interpretation implicit in the argument from falling stones.

According to Copernicus the motion of a falling stone should be "mixed straight-and-circular." [136] By the "motion of the stone" is meant, not just its motion relative to some visible mark in the visual field of the observer, or its *observed motion*, but rather its motion in the solar system, or in (absolute) space, or its *real motion*. The familiar facts appealed to in the argument assert a different kind of motion, a simple vertical motion. This result refutes the Copernican hypothesis only if the concept of motion that occurs in the observation statement is the same as the concept of motion that occurs in the Copernican prediction. The observation statement "the stone is falling straight down" must therefore, likewise refer to a movement in (absolute) space. It must refer to a real motion.

Now, the force of an "argument from observation" derives from the fact that the observation statements it involves are firmly connected with appearances. There is no use appealing to observation if one does not know how to describe what one sees, or if one can offer one's description with hesitation only, as if one had just learned the language in which it is formulated. An observation statement, then, consists of two very different psychological events: (1) a clear and unambiguous sensation and (2) a clear and unambiguous connection between this sensation and parts of a language. This is the way in which the sensation is made to speak. Do the sensations in the argument above speak the language of real motion?

They speak the language of real motion in the context of seventeenth-century everyday thought. At least this is what Galileo tells us. He tells us that the everyday thinking of the time assumes the "operative" character of *all* motion,[137] or, to use well-known philosophical terms, it assumes a *naive realism with respect to motion*: except for occasional and unavoidable illusions, apparent motion is identical with real (absolute) motion. Of course, this distinction is not explicitly drawn. One does not first distinguish the apparent motion from the real motion and then connect the two by a correspondence rule. Quite the contrary, one describes, perceives, acts toward the apparent motion as if it were already the real thing. Nor does one proceed in this manner under all circumstances. It is admitted that objects may move which are not seen to move; and it is also admitted that certain motions are illusory (see the example in section 7 above). Apparent motion and real motion are not always identified. However, there are paradigmatic cases in which it is psychologically very difficult, if not plainly impossible, to admit deception. It is from these paradigmatic cases, and not from exceptions, that naive realism with respect to motions derives its strength. These are also the situations in which we first learn our kinematic vocabulary. From our very childhood we learn to react to them with concepts which have naive realism built right into them, and which inextricably connect movement and the appearance of movement. The motion of the stone in the tower argument, or the alleged motion of the earth, is such a paradigmatic case. How could one possibly be unaware of the swift motion of a large bulk of matter such as the earth is supposed to be! How could one possibly be unaware of the fact that the falling stone traces a vastly extended trajectory through space! From the point of view of seventeenth-century thought and language, the argument is, therefore, impeccable and quite forceful. However, notice how *theories* ("operative character" of all motion: essential correctness of sense reports), which are not formulated explicitly, enter the debate in the guise of observational terms. We realize again that observational terms are Trojan horses which must be watched very carefully. How is one supposed to proceed in such a sticky situation?

The argument from falling stones seems to refute the Copernican view. This may be due to an inherent disadvantage of Copernicanism; but it may also be due to the presence of natural interpretations which are in need of improvement. The first task, then, is to discover and to isolate these unexamined obstacles to progress.

It was Bacon's belief that natural interpretations could be discovered by a method of analysis that peels them off, one after another, until the sensory core of every observation is laid bare. This method has serious drawbacks. First, natural interpretations of the kind considered by Bacon are not just added to a previously existing field of sensations. They are instrumental in constituting the field, as Bacon says himself. Eliminate all natural interpretations, and you also eliminate the ability to think and to perceive. Second, disregarding this fundamental function of natural interpretations, it should be clear that a person who faces a perceptual field

without a single natural interpretation at his disposal would be *completely disoriented*; he could not even *start* the business of science. Third, the fact that we do start, even after some Baconian analysis, shows that the analysis has stopped prematurely. It has stopped at precisely those natural interpretations of which we are not aware and without which we cannot proceed. It follows that the intention to start from scratch, after a complete removal of all natural interpretations, is self-defeating.

Furthermore, it is not possible to even *partly* unravel the cluster of natural interpretations. At first sight the task would seem to be simple enough. One takes observation statements, one after the other, and analyzes their content. However, concepts that are hidden in observation statements are not likely to reveal themselves in the more abstract parts of language. If they do, it will still be difficult to nail them down; concepts, just as percepts, are ambiguous and dependent on background. Moreover, the content of a concept is determined also by the way in which it is related to perception. Yet how can this way be discovered without circularity? Perceptions must be identified, and the identifying mechanism will contain some of the very same elements which govern the use of the concept to be investigated. We never penetrate this concept completely, for we always use part of it in the attempt to find its constituents.[138] There is only one way to get out of this circle, and it consists in using an *external* measure of comparison, including new ways of relating concepts and percepts. Removed from the domain of natural discourse and from all those principles, habits, and attitudes which constitute its form of life, such an external measure will look strange indeed. This, however, is not an argument against its use. Quite the contrary, such an impression of strangeness reveals that natural interpretations are at work, and it is a first step toward their discovery. Let us explain this situation with the help of the tower example.

The example is intended to show that the Copernican view is not in accordance with 'the facts.' Seen from the point of view of these 'facts,' the idea of the motion of the earth appears to be outlandish, absurd, and obviously false, to mention only some of the expressions which were frequently used at the time, and which are still heard wherever professional squares confront a new and counterfactual theory. This makes us suspect that the Copernican view is an external measuring rod of precisely the kind described above.

We now can turn the argument around and use it as a detecting device