that helps us to discover the natural interpretations that exclude the motion of the earth. Turning the argument around, we *first assert* the motion of the earth and *then inquire* what changes will remove the contradiction. Such an inquiry may take considerable time, and there is a good sense in which one can say that it is not yet finished, not even today. The contradiction, therefore, may stay with us for decades or even centuries. Still, *it must be upheld* (Hegel!) until we have finished our examination or else the examination, the attempt to discover the antediluvian components of our knowledge, cannot even start. This, we have seen, is one of the reasons one can give for *retaining*, and, perhaps, even for *inventing*, theories which are inconsistent with the facts: Ideological ingredients of our knowledge and, more especially, of our observations, are discovered with the help of theories which are refuted by them. *They are discovered counter-inductively.*

Let me repeat what has been asserted so far. Theories are tested and possibly refuted by facts. Facts contain ideological components, older views which have vanished from sight or were perhaps never formulated in an explicit manner. These components are highly suspicious, first, because of their age, because of their antediluvian origin; second, because their very nature protects them from a critical examination and always has protected them from such an examination. Considering a contradiction between a new and interesting theory and a collection of firmly established facts, the best procedure is, therefore, not to abandon the theory but to use it for the discovery of the hidden principles that are responsible for the contradiction. Counterinduction is an essential part of such a process of discovery. (Excellent historical example: the arguments against motion and atomicity of Parmenides and Zeno. Diogenes of Sinope, the Cynic, took the simple course that would be taken by many contemporary scientists and all contemporary philosophers: he refuted the arguments by rising and walking up and down. The opposite course, recommended here, led to much more interesting results, as is witnessed by the history of the case. One should not be too hard on Diogenes, however, for it is also reported that he beat a pupil who was content with his refutation, exclaiming that he had given reasons which the pupil should not accept without additional reasons of his own.[139])

Having *discovered* a particular natural interpretation, the next question is how it is to be *examined* and *tested*. Obviously, we cannot proceed in the usual way, i.e., derive predictions and compare them with "results of

observation." These results are no longer available. The idea that the senses, employed under normal circumstances, produce correct reports of real events, for example reports of the real motion of physical bodies, has now been removed from all observational statements. (Remember that this notion was found to be an essential part of the anti-Copernican argument.) But without it our sensory reactions cease to be relevant for tests. This conclusion has been generalized by some rationalists, who decided to build their science on reason only and ascribed to observation a quite insignificant auxiliary function. Galileo does not adopt this procedure.

If *one* natural interpretation causes trouble for an attractive view, and if its *elimination* removes the view from the domain of observation, then the only acceptable procedure is to use *other* interpretations and to see what happens. The interpretation which Galileo uses restores the senses to their position as instruments of exploration, *but only with respect to the reality of relative motion.* Motion "among things which share it in common" is "nonoperative," that is, "it remains insensible, imperceptible, and without any effect whatever." [140] Galileo's first step in the joint examination of the Copernican doctrine, and of a familiar but hidden natural interpretation, consists therefore in *replacing the latter by a different interpretation*, or, considering the function of natural interpretations, *he introduces a new observation language.*

This is, of course, an entirely legitimate move. In general, the observation language which enters an argument has been in use for a long time and is quite familiar. Considering the structure of common idioms on the one hand, and of the Aristotelian philosophy on the other, neither this use nor the familiarity can be regarded as a test of the underlying principles. These principles, these natural interpretations, occur in every description. Extraordinary cases which might create difficulties are defused with the help of "adjuster words," [141] such as "like" or "analogous," which divert them so that the basic ontology remains unchallenged. A test is, however, urgently needed. It is needed especially in those cases where the principles seem to threaten a new theory. It is then quite reasonable to introduce alternative observation languages and to compare them both with the original idiom and with the theory under examination. Proceeding in this way, we must make sure that the comparison is *fair*. That is, we must not criticize an idiom that is supposed to function as an observation language because it is not yet well known and is therefore less strongly connected with our sensory reactions and less plausible than is another and more "common" idiom. Superficial criticisms of this kind, which have been elevated into an entire new "philosophy," abound in discussions of the mind-body problem. Philosophers who want to introduce and to test new views thus find themselves faced not with *arguments*, which they could most likely answer, but with an impenetrable stone wall of well-entrenched *reactions*. This is not at all different from the attitude of people ignorant of foreign languages, who feel that a certain color is much better described by "red" than by "rosso." As opposed to such attempts at conversion by appeal to familiarity ("I know what pains are, and I also know, from introspection, that they have nothing whatever to do with material processes!"), we must emphasize that a comparative judgment of observation languages, e.g., materialistic observation languages, phenomenalistic observation languages, objective-idealistic observation languages, theological observation languages, can start only *when all of them are spoken equally fluently*.

Let me assert at this point that while it is possible to consider and to actively apply various rules of thumb, and while we may in this way arrive at a satisfactory judgment, it is not at all wise to go further and to turn these rules of thumb into necessary conditions of science. For example, one might be inclined to say, following Neurath, that an observation language A is preferable to an observation language B, if it is at least as useful as B in our everyday life, *and* if more theories and more comprehensive theories are compatible with it than are compatible with B. Such a criterion takes into account that *both* our perceptions (natural interpretations included) *and* our theories are fallible, and it also pays attention to our desire for a harmonious and universal point of view. (One always seems to assume that observation languages should be employed not only in laboratories, but also at home, and in the "natural surroundings" of the scientist.) However, we must not forget that we find and improve the assumptions hidden in our observational reports by a method that makes use of inconsistencies. Hence, we might prefer B to A as a starting point of analysis, and we might in this way arrive at a language C which satisfies the criterion even better, but which cannot be reached from A. Conceptual progress like any other kind of progress depends on psychological circumstances, which may prohibit in one case what they encourage in another. Moreover the psychological factors which come into play are never clear in advance. Nor should the demand for practicality and sensory content be regarded as a *conditio sine qua non*. We possess detecting mechanisms whose performance outdistances our senses. Combining such detectors with a computer, we may test a theory directly, without intervention of a human observer. This would eliminate sensations and perceptions from the process of testing. Using hypnosis, one could eliminate them from the transfer of the results into the human brain also, and thus arrive at a science that is completely without experience.[142] Considerations like these, which indicate possible paths of development, should cure us once and for all of the belief that judgments of progress, improvement, etc., are based on rules which can be revealed *now* and will remain in action for all the years to come. My discussion of Galileo has not, therefore, the aim of arriving at the "correct method." It has rather the aim of showing that such a "correct method" does not *and cannot* exist. More especially, it has the limited aim of showing that counterinduction is very often a reasonable move. Let us now proceed a step further in our analysis of Galileo's reasoning!

7. The Tower Argument: Analysis Continued

Galileo replaces one natural interpretation by a very different and as yet (1630!) at least partly unnatural interpretation. How does he proceed? How does he manage to introduce absurd and counterinductive assertions such as the assertion that the earth moves, and how does he manage to get them a just and attentive hearing? One may anticipate that arguments will not suffice—an interesting, and highly important, limitation of rationalism—and Galileo's utterances are indeed arguments in appearance only. For Galileo uses *propaganda*. He uses *psychological tricks* in addition to whatever intellectual reasons he has to offer. These tricks are very successful; they lead him to victory. But they obscure the new attitude toward experience that is in the making, and postpone for centuries the possibility of a reasonable philosophy. They obscure the fact that the experience on which Galileo wants to base the Copernican view is nothing but the result of his own fertile imagination, that it has been *invented*. They obscure this fact by insinuating that the new results which emerge are known and conceded by all, and need only be called to our attention to appear as the most obvious expression of the truth.

Galileo "reminds" us that there are situations in which the nonoperative character of shared motion is just as evident and as firmly believed as the idea of the operative character of all motion is in other circumstances (this latter idea is therefore not the only natural interpretation of motion). The situations are events in a boat, in a smoothly moving carriage, and in any other system that contains an observer and permits him to carry out some simple operations.

Sagredo: There has just occurred to me a certain fantasy which passed through my imagination one day while I was sailing to Aleppo, where I was going as consul for our country . . . If the point of a pen had been on the ship during my whole voyage from Venice to Alexandretta and had had the property of leaving visible marks of its whole trip, what trace—what mark—what line would it have left?

Simplicio: It would have left a line extending from Venice to there; not perfectly straight—or rather, not lying in the perfect arc of a circle—but more or less fluctuating according as the vessel would now and again have rocked. But this bending in some places a yard or two to the right or left, up or down, in length of many hundreds of miles, would have made little alteration in the whole extent of the line. These would scarcely be sensible, and without an error of any moment it could be called part of a perfect arc.

Sagredo: So that if the fluctuation of the waves were taken away and the motion of the vessel were calm and tranquil, the true and precise motion of that pen point would have been an arc of a perfect circle. Now if I had had that same pen continually in my hand, and had moved it only a little sometimes this way or that, what alteration should I have brought into the main extent of this line?

Simplicio: Less than that which would be given to a straight line a thousand yards long which deviated from absolute straightness here and there by a flea's eye.

Sagredo: Then if an artist had begun drawing with that pen on a sheet of paper when he left the port and had continued doing so all the way to Alexandretta, he would have been able to derive from the pen's motion a whole narrative of many figures, completely traced and sketched in thousands of directions, with landscapes, buildings, animals, and other things. Yet the actual, real, essential movement marked by the pen point would have been only a line; long, indeed, but very simple. But as to the artist's own actions, these would have been conducted exactly the same as if the ship had been standing still. The reason that of the pen's long motion no trace would remain except the marks drawn upon the paper is that the gross motion from Venice to Alexandretta was common to the paper, the pen, and everything else in the ship. But the small motions back and forth, to right and left, communicated by the artist's fingers to the pen but not to the paper, and belonging to the former alone, could thereby leave a trace on the paper which remained stationary to those motions.[143]

Or:

Salviati: . . . imagine yourself in a boat with your eyes fixed on a point

of the sail yard. Do you think that because the boat is moving along briskly, you will have to move your eyes in order to keep your vision always on that point of the sail yard and follow its motion?

Simplicio: I am sure that I should not need to make any change at all; not just as to my vision, but if I had aimed a musket I should never have to move it a hairsbreadth to keep it aimed, no matter how the boat moved.

Salviati: And this comes about because the motion which the ship confers upon the sail yard, it confers also upon you and upon your eyes, so that you need not move them a bit in order to gaze at the top of the sail yard, which consequently appears motionless to you. (And the rays of vision go from the eye to the sail yard just as if a cord were tied between the two ends of the boat. Now a hundred cords are tied at different fixed points, each of which keeps its place whether the ship moves or remains still).[144]

It is clear that these situations lead to a nonoperative concept of motion even within common sense.

On the other hand, common sense, and I mean seventeenth-century common sense, also contains the idea of the *operative* character of all motion. This latter idea arises when a limited object that does not contain too many parts moves in vast and stable surroundings, for example, when a camel trots through the desert, or when a stone descends from a tower.

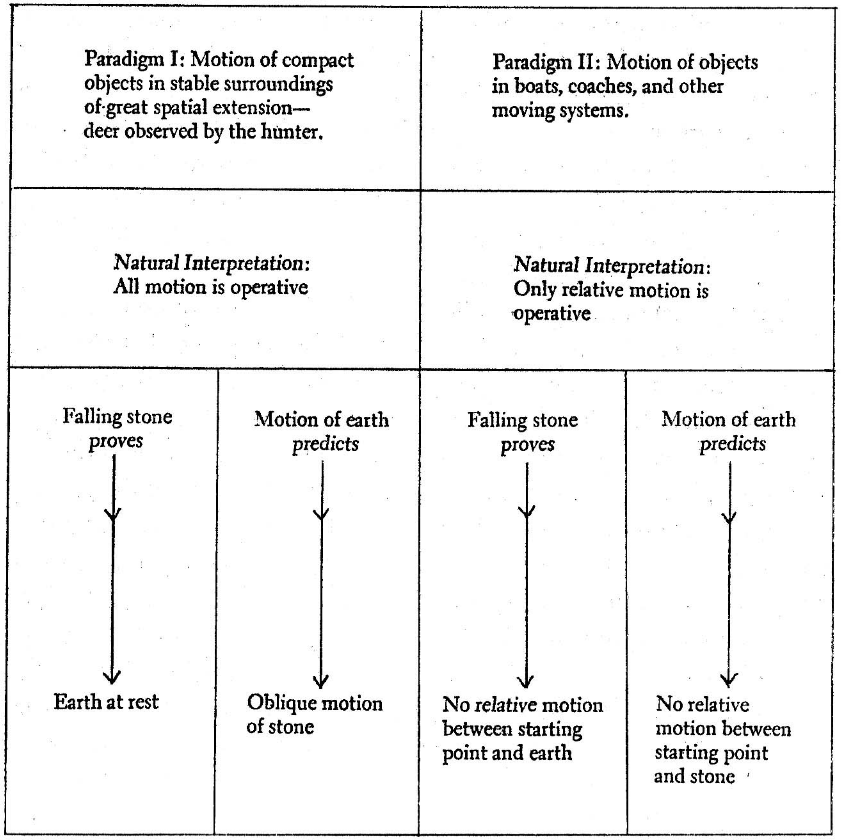
Now, Galileo urges us to "remember" the conditions in which we assert the nonoperative character of shared motion in this case also, and to subsume the second case under the first.

Thus, the first of the two paradigms of nonoperative motion mentioned above is followed by the assertion that "it is likewise true that the earth being moved, the motion of the stone in descending is actually a long stretch of many hundred yards, or even many thousand; and had it been able to mark its course in motionless air or upon some other surface, it would have left a very long slanting line. But that part of all this motion which is common to the rock, the tower, and ourselves remains insensible and as if it did not exist. There remains observable only that part in which neither the tower nor we are participants; in a word, that with which the stone in falling measures the tower." [145]

And the second paradigm precedes the exhortation to "transfer this argument to the whirling of the earth and to the rock placed on top of the tower, whose motion you cannot discern because in common with the rock you possess from the earth that motion which is required for following the tower; you do not need to move your eyes. Next, if you add to the rock a downward motion which is peculiar to it and not shared by you, and which is mixed with this circular motion, the circular portion of the motion which is common to the stone and the eye continues to be imperceptible. The straight motion alone is sensible, for to follow that you must move your eyes downwards." [146]

This is strong persuasion indeed.

Yielding to this persuasion, we now *quite automatically* start confounding the conditions of the two cases and become relativists. This is the essence of Galileo's trickery! As a result the clash between Copernicus and "the conditions affecting ourselves and those in the air above us" [147] dissolves into thin air, and we finally realize "that all terrestrial events from which it is ordinarily held that the earth stands still and the sun and the fixed stars are moving would necessarily appear just the same to us if the earth moved and the others stood still." [148]



Let us now look at the situation from a more abstract point of view. We start with two conceptual subsystems of ordinary thought (see the preceding diagram). One of them regards motion as an absolute process which always has effects, effects on our senses included. The description of this conceptual system which appears in the present paper may be somewhat idealized, but the arguments of the opponents of Copernicus which are quoted by Galileo himself, and which according to him were "very plausible," [149] show that there was a widespread tendency to think in its terms, and that this tendency was a serious obstacle for the discussion of alternative ideas. Occasionally one finds even more primitive ways of thinking, where concepts such as "up" and "down" are used absolutely. Examples are the assertion "that the earth is too heavy to climb up over the sun and then fall headlong back down again," [150] or the assertion that "after a short time the mountains, sinking downward with the rotation of the terrestrial globe, would get into such a position that whereas a little earlier one would have had to climb steeply to their peaks, a few hours later one would have to stoop and descend in order to get there." [151] Galileo, in his marginal notes, calls these "utterly childish reasons [which] suffice[d] to keep imbeciles believing in the fixity of the earth" [152] and he thinks it unnecessary "to bother about such men as these, *whose name is legion*, or to take notice of their fooleries." [153] Yet it is clear that the absolute idea of motion was "well entrenched," and that the attempt to replace it was bound to encounter strong resistance.

The second conceptual system is built around the relativity of motion, and is also well entrenched in its own domain of application. Galileo aims at replacing the first system by the second in *all* cases, terrestrial as well as celestial. Naive realism with respect to motion is to be *completely eliminated.*

Now, we have seen that this naive realism is on occasions an essential part of our observational vocabulary. On these occasions (Paradigm I), the observation language contains the idea of the efficacy of *all* motion. Or, to express it in the material mode of speech, our experience in these situations is the experience of objects which move absolutely. Taking this into consideration, it is apparent that Galileo's proposal amounts to a partial revision of our observation language or of our experience. An experience which partly *contradicts* the idea of the motion of the earth is turned into an experience that *confirms* it, at least as far as "terrestrial things" are concerned.[154] This is what *actually happens.* But Galileo wants to persuade us that no change has taken place, that the second conceptual system is already universally *known*, even though it is not universally *used*. Both Salviati, his representative in the dialogue, and his opponent Simplicio, and also Sagredo, the intelligent layman, connect Galileo's method of argumentation with Plato's theory of *anamnesis*[155]—a clever tactical move, typically Galilean, one is inclined to say. Yet we must not allow ourselves to be deceived about the revolutionary development that is actually taking place.

The resistance against the assumption that shared motion is nonoperative was equated with the resistance which forgotten ideas exhibit toward the attempt to make them known. Let us accept this *interpretation* of the resistance! But let us not forget its *existence*. We must then admit that it restricts the use of the relativistic ideas, confining them to *part* of our everyday experience. *Outside* this part, and that means in interstellar space, they are "forgotten," and therefore not active. But outside this part there is not complete chaos. Other concepts are used, among them those very same absolutistic concepts which derive from the first paradigm. We not only use them, but must admit that they are entirely adequate. No difficulties arise as long as one remains within the limits of the first paradigm. "Experience," that is, the totality of all facts from all domains described with the concepts which are appropriate in these domains, cannot force us to carry out the change which Galileo wants to introduce. The motive for a change must come from a different source.

It comes, first, from the desire to see "the whole [correspond] to its parts with wonderful simplicity" [156] as Copernicus had already expressed himself. It comes from the "typically metaphysical urge" for unity of understanding and conceptual presentation. And the motive for a change is connected, secondly, with the intention to make room for the motion of the earth, which Galileo accepts and is not prepared to give up. The idea of the motion of the earth is closer to the first paradigm than to the second, or at least it was at the time of Galileo. This gave great strength to the Aristotelian arguments, and made them very plausible. To eliminate this plausibility, it was necessary to subsume the first paradigm under the second, and to extend the relative notions to all phenomena. The idea of *anamnesis* functions here as a psychological crutch, as a lever which smoothes the process of subsumption by concealing its existence. As a result we are now *ready* to apply the relative notions not only to boats, coaches, birds, but also to the "solid and well-established earth" as a whole.

And we have the impression that this readiness was in us all the time, although it took some effort to make it conscious. This impression is most certainly erroneous: it is the result of Galileo's propagandistic machinations. We would do better to describe the situation in a different way, as a change of our conceptual system. Or, because we are dealing with concepts which belong to natural interpretations, and which are therefore connected with sensations in a very direct way, we should describe it as a *change of experience* that allows us to accommodate the Copernican doctrine. The change corresponds perfectly to the pattern outlined in an earlier paper: an inadequate view, the Copernican theory, is supported by another inadequate view, the idea of the nonoperative character of shared motion, and both theories gain strength and give support to each other in the process. It is this change which constitutes the transition from the Aristotelian point of view to the epistemology of modern science.

For experience now ceases to be that unchangeable fundament which it is both in common sense and in the Aristotelian philosophy. The attempt to support Copernicus makes experience "fluid" in the very same manner in which it makes the heavens fluid, "so that each star roves around in it by itself." [157] An empiricist who starts from experience, and builds on it without ever looking back, now loses the very ground on which he stands. Neither the earth, "the solid, well-established earth," nor the facts on which he usually relies, can be trusted any longer. It is clear that a philosophy that uses such a fluid and changing experience needs new methodological principles which do not insist on an asymmetric judgment of theories by experience. *Classical physics* intuitively adopts such principles; at least the great and independent thinkers, such as Newton, Faraday, and Boltzmann, proceed in this way. But its *official doctrine* still clings to the idea of a stable and unchanging basis. The clash between this doctrine and the actual procedure is concealed by a tendentious presentation of the *results* of research that hides their revolutionary origin and suggests that they arose from a stable and unchanging source. These methods of concealment start with Galileo's attempt to introduce new ideas under the cover of anamnesis, and they culminate in Newton.[158] They must be exposed if we want to arrive at a better account of the progressive elements in science.

8. The Law of Inertia

Our discussion of the anti-Copernican argument is not yet complete. So far, we have tried to discover what assumption will make a stone *that moves alongside a moving tower* appear to fall "straight down," instead of being seen to move in an arc. The assumption, which I shall call the *relativity principle*, that our senses notice only relative motion, and are completely insensitive to a motion which objects have in common, was seen to do the trick. What remains to be explained is *why the stone stays with the tower*, and why it is not left behind. In order to save the Copernican view, one must explain not only why a motion that preserves the relation among visible objects *remains unnoticed*, but also why a common motion of various objects does not affect their relation. That is, one must explain why such a motion is not a *causal agent*. Turning the question around in the manner explained in section 6, it is now apparent that the anti-Copernican argument of section 5 rests on *two* natural interpretations,[159] viz. the *epistemological assumption* that absolute motion is always *noticed* and the *dynamical principle* that objects (such as the falling stone) which are not interfered with move toward their natural place. The present problem is to supplement the relativity principle with a new law of inertia in such a fashion that the motion of the earth can still be asserted. One sees at once that the following law, the *principle of circular inertia*, as I shall call it, provides the required solution: An object that moves with a given angular velocity on a frictionless sphere around the center of the earth will continue moving with the same angular velocity forever. Combining the appearance of the falling stone with the relativity principle, the principle of circular inertia, and some simple assumptions concerning the composition of velocities, yields an argument which no longer endangers Copernicus's view, but can be used to give it partial support.

The relativity principle was defended in two ways. The first was by showing how it helps Copernicus; this defense is truly ad hoc. The second was by pointing to its function in common sense, and by surreptitiously generalizing that function (see section 7). No independent argument was given for its validity.[160] Galileo's method of support for the principle of circular inertia is of exactly the same kind. He introduces it, again not by reference to experiment or to independent observation, but by reference to what everyone is already supposed to know.

Simplicio: So you have not made a hundred tests, or even one? And yet you so freely declare it to be certain? . . .

Salviati: Without experiment, I am sure that the effect will happen as I tell you, because it must happen that way; and I might add that you yourself also know that it cannot happen otherwise, no matter how you may pretend not to know it . . . But I am so handy at picking people's brains that I shall make you confess this in spite of yourself.[161]

Step by step Simplicio is forced to admit that a body that moves without friction on a sphere concentric with the center of the earth will carry out a "boundless," a "perpetual" motion.[162] We know, of course, especially after the analysis we have just completed of the nonoperative character of shared motion, that what Simplicio accepts is based neither on experiment nor on corroborated theory. It is a daring new suggestion involving a tremendous leap of the imagination. A little more analysis then shows that this suggestion is connected with experiments, such as the "experiments" of the *Discorsi*, by ad hoc hypotheses. (The amount of friction to be eliminated follows not from independent investigations—such investigations commence only much later, in the eighteenth century—but from the very result to be achieved, viz. the circular law of inertia.) Viewing natural phenomena in this way leads, as we have already said, to a complete reevaluation of all experience. We can now add that it leads to the invention of a *new kind of experience* that is not only more sophisticated *but also far more speculative* than is the experience of Aristotle or of common sense. Speaking paradoxically, but not incorrectly, one may say that *Galileo invented an experience that has metaphysical ingredients*.[163] It is by means of such an experience that the transition from a geostatic cosmology to the point of view of Copernicus and Kepler is achieved.

9. The Progressive Role of Ad Hoc Hypotheses

This is the place to briefly mention certain ideas which have been developed by Lakatos, and which throw new light on the problem of the growth of knowledge.

It is customary to assume that good scientists refuse to employ ad hoc hypotheses, and to assert that they are right in their refusal. New ideas, so it is thought, go far beyond the available evidence, and they *must* go beyond it in order to be of value. Ad hoc hypotheses are bound to creep in eventually, but they should be resisted and kept at bay. This is the customary attitude as it is expressed, for example, in the writings of K. R. Popper.

As opposed to this, Lakatos, in lectures, and now also in publications, has pointed out that "adhocness" is neither despicable nor absent from the body of science. New ideas, he emphasizes, are usually almost entirely ad hoc, they cannot be otherwise. And they are reformed only in a piecemeal fashion, by gradually stretching them, so that they apply to situations lying beyond their starting point. Schematically:

Popper: new theories have, and must have, excess content which is, but should not be, gradually infected by ad hoc adaptations.

Lakatos: new theories are, and cannot be anything but, ad hoc. Excess content is, and should be, created in a piecemeal fashion, by gradually extending them to new facts and domains.

The historical material I have just analyzed (and the more extensive material presented in "Problems of Empiricism, Part II") lends unambiguous support to the position of Lakatos. In what follows I shall try to show this in some detail.

First, *kinematic relativity* (cf. section 7, above):

Just like Newtonian physics, Aristotelian physics distinguishes between relative space and absolute space.[164] In addition, it allows one to "operationally" determine absolute places, directions, velocities. One may proceed in the following way: The *center* of the universe is found, for example, by backwardly elongating the direction of two flames, and it is tested by using a third flame. Flames function here as test bodies and not as reference bodies for relative motion. *Distance* from the center is determined by the strength of the upward motion of flames, or of suitable mixtures which may be enclosed in test capsules. Thus, space is traced out, in an entirely physical way, by using known physical laws. *Direction*, finally, is determined by determining the axis of rotation of the stellar sphere. This whole physical background is removed by Galileo. With it, we lose all means of testing for center, distance, and direction. The new relativistic principles (only relative motion is "operative") are therefore metaphysical, and, because adapted to the tower experiment, also ad hoc.

Considering now *dynamical relativity* (section 8), one should remember, first of all, that the natural character of circular motion was not first asserted by Galileo. It was an old assumption, concerning all supralunar entities. The new assumption introduced by Galileo (and by Copernicus, in chapter VIII of *De revolutionibus*) is that circular motion is a natural motion *for terrestrial objects also*. On the one hand, this is an immediate

consequence of having made the earth a star: Stars move in circles. Hence, if the earth is a star, its natural motion will be circular, both its motion around the sun and "its motion with respect to itself," as its rotation was described at the time. Now, does this particular assumption of the rotation of earth assert anything over and above what was known to happen at its surface at Galileo's own time? My attitude, which is in accordance with Lakatos's general theory, is that the answer must be *no*. The only consequence of the assertion is that it connects moving objects rigidly with the framework of the moving, i.e., rotating, earth. This leaves everything as it is, and it especially leaves the results of the tower experiment and the cannon experiment unchanged.[165] No further consequence was implied at the time. (It was different with the motion of the earth around the sun which led one to expect a sizable stellar parallax.) Even the later Newtonian argument that distant objects, moving with the same angular velocity, will hit the earth *ahead* of the tower cannot be used at this stage: it is not at all clear whether Galileo would want distant objects to move with the same angular velocity. (In the case of the planets he notices their *decreasing* angular velocity—the effect of Kepler's third law—and he might have been inclined to treat bodies circulating around the earth in the same way. On the other hand, he calculates the time a stone takes to drop from the moon to the earth by assuming a constant acceleration all the way.[166])

Furthermore, I do not think that bringing in the tradition of the impetus theory will improve matters. For this theory is again ad hoc, this time not with respect to the tower, but with respect to the behavior of objects thrown (which continue to move, contrary to Aristotle's law of inertia). When a circular law is asserted, as seems to be the case with Buridan, the problem is the same as for Galileo.[167] (Besides, the impetus theory is incompatible with Galileo's idea of the nonoperative character of all motion.[168])

Finally, one must not argue against "ad hocness" by pointing to the fact that *experiments* were made in boats, with cannon balls, on towers, and so on.[169] These experiments did not lead to any decisive result. And they did not test any excess *content* of the law of circular inertia, but tried to establish the *fact* which the law then explains in ad hoc fashion. Reference to the experiments with the inclined plane is also beside the point. These experiments test, if that is the right word, the law of free fall. But of course there still remains the task of subdividing *that* motion into an inertial motion and something else. However one looks at the matter, the best conjecture is that at the time in question the circular law of inertia, and to an even greater extent the idea of the relativity of motion, was an ad hoc hypothesis designed to get out of the trouble of the tower.

Now this is such an incredible situation that a little more argument seems to be required. We therefore take a brief look at Galileo's earlier work on mechanics and motion.

In *De motu* motions of spheres in the center of the universe, outside of it, homogeneous, nonhomogeneous, supported at the center of gravity, supported outside of it are discussed, and described as being either natural, or forced, or neither. But about the actual motion of such spheres we hear very little, and what we do hear is by implication only. Thus there appears the question[170] whether a homogeneous sphere made to move in the center of the universe would move forever. We read that "it seems that it should move perpetually," but an unambiguous answer is never given. A marble sphere supported on an axis through the center and set in motion is said to "rotate for a long time" [171] in *De motu* while a perpetual motion is said to be "quite out of keeping with the nature of the earth itself to which rest seems to be more congenial than motion" in the *Dialogue on Motion*.[172] Another argument against perpetual rotations is found in Benedetti's *Diverse Speculations*.[173] Rotations, says Benedetti, are "certainly not perpetual," for the parts of the sphere, wanting to move in a straight line, are constrained against their nature, "and so they come to rest naturally." Again, in *De motu*,[174] we find a criticism of the assertion that adding a star to the celestial sphere might slow it down by changing the relation between the force of the moving intelligences and the resistance of the sphere. This assertion, Galileo says, certainly applies to an excentric sphere. Adding weight to an excentric sphere means that a weight will occasionally be moved away from the center and be raised to a higher level. But "who would ever say that [a concentric sphere] was impeded by the weight, since the weight in its circular path would neither approach, nor recede from, the center." [175] Note that the original rotation is in this case said to be caused by an intelligence; it is not assumed to be taking place all by itself. This is in perfect agreement with Aristotle's *general* theory of motion[176] where a mover is postulated for every motion, and not just for violent motions. Galileo seems to accept this part of

the theory both when letting rotating spheres slow down and when accepting the "force of the intelligences" in the present argument (he also accepts impetus—see below). But in objecting to the idea that a new star will increase resistance he adopts the entirely different view that resistance occurs only when a motion is *forced*, and is absent otherwise. This is neither Aristotelian nor compatible with the version of the impetus theory he holds at the time which attributes *any* prolonged motion to an internal moving force similar to the force of sound that resides in a bell long after it has been struck,[177] and which is again supposed to "gradually diminish." [178]

Looking at these few examples we see that Galileo ascribes a special position to motions which are neither violent nor forced. Such motions may last for a considerable time though they are not supported by the surrounding medium. But *they do not last forever*, and *they need an internal driving force* in order to even persist for a finite time.

Now if one wants to overcome the dynamical arguments against the motion of the earth (and we are here always thinking about its *rotation* rather than about its motion around the sun), then the two italicized principles must both be revised. It must be assumed that the "neutral" motions which Galileo discusses in his early dynamical writings may last forever, or at least for periods comparable to the age of historical records. And these motions must be regarded as "natural" in the entirely new and revolutionary sense that neither an outer nor an inner motor is needed to keep them going. The first assumption is necessary to allow the earth to rotate. The second assumption is necessary if we want to regard motion as a *relative* phenomenon, depending on the choice of a suitable coordinate system.[179] Copernicus, in his brief remarks on the problem,[180] makes both assumptions. Galileo never clearly resolves the problem. He formulates permanence along a horizontal line as a hypothesis in his *Discorsi*[181] and he seems to make both assumptions in the *Dialogue*.[182] Now my guess is that *a clear statement of permanent motion with(out) impetus developed in Galileo only together with his gradual acceptance of the Copernican view*. Galileo changed his view about the "neutral" motions—he made them permanent and "natural"—in order to make them compatible with the rotation of the earth and in order to evade the difficulties of the tower argument.[183] His new ideas concerning such motions are therefore at least partly ad hoc. Impetus in the old sense disappeared partly for methodological reasons (interest in the *how*, not in the *why*—this development itself deserves careful study), partly because of the vaguely perceived inconsistency with the idea of the relativity of all motion. The wish to save Copernicus plays a role in either case. This hypothesis must of course be tested by an examination of Galileo's published writings and his correspondence between 1590 and, say, 1630. Considering what we know already we must admit that it has much plausibility.

Now, if we are right in *assuming* that Galileo framed an ad hoc hypothesis at this point, then we can also *praise him* for his methodological acumen. It is obvious that the moving earth demands a new dynamics. *One* test of the old dynamics consists in the attempt to establish the motion of the earth. Trying to establish the motion of the earth is the same as trying to find a refuting instance for the old dynamics. The motion of the earth, however, is inconsistent with the tower experiment *interpreted in accordance with the old dynamics*. Interpreting the tower experiment in accordance with the old dynamics therefore means trying to save the old dynamics in an ad hoc fashion. If one does not want to do this one must find a different interpretation for the phenomena of free fall. What interpretation should be chosen? One wants an interpretation that turns the motion of the earth into a refuting instance of the old dynamics, without lending ad hoc support to the motion of the earth itself. The first step toward such an interpretation is to establish contact, however vague, with the "phenomena," i.e., with the falling stone, and to establish it in such a manner that the motion of the earth is not *obviously* contradicted. The most primitive element of this first step is to frame an ad hoc hypothesis with respect to the rotation of the earth. The next step would then be to elaborate the hypothesis, so that additional predictions become possible. Copernicus and Galileo take the first and most primitive step. Their procedure looks contemptible only if one forgets that the aim is *to test older views* rather than *to prove new ones*, and if one also forgets that developing a good theory is a complex process that has to start modestly and that takes time. But why, an impatient methodologist might ask, did it take so long before additional phenomena were added? It took so long *because the domain of possible phenomena had first to be circumscribed by the further development of the Copernican hypothesis*. It is much better to remain ad hoc for a while, and in the meantime to develop heliocentrism in all its astronomical ramifications which can then be used as guidelines for a further elaboration of dynamics.

Therefore: Galileo *did* use ad hoc hypotheses. It *was good* that he used them. Had he not been ad hoc, he would have been ad hoc anyway, but this time with respect to an older theory. Hence, as one cannot help being ad hoc, it is better to be ad hoc with respect to a new theory, for a new theory, like all new things, will give a feeling of freedom, excitement, and progress. Galileo is to be applauded because he preferred protecting an interesting hypothesis to protecting a dull one.

10. Summary of Analysis of Tower Argument

I repeat and summarize: An argument is proposed that refutes Copernicus by observation. The argument is inverted in order to discover those natural interpretations which are responsible for the contradiction. The offensive interpretations are replaced by others. Propaganda and appeal to distant and highly theoretical parts of common sense are used to defuse old habits and to enthrone new ones. The new natural interpretations which are also formulated explicitly as auxiliary hypotheses are established partly by the support they give to Copernicus and partly by plausibility considerations and ad hoc hypotheses. An entirely new "experience" arises in this way. Independent evidence is as yet entirely lacking, but this is no drawback as it is to be expected that independent support will take a long time appearing. For what is needed is a theory of solid objects, aerodynamics, hydrodynamics, and all these sciences are still hidden in the future. *But their task is now well defined*, for Galileo's assumptions, his ad hoc hypotheses included, are sufficiently clear and simple to prescribe the direction of future research. Let it be noted, incidentally, that Galileo's procedure drastically reduces the content of dynamics. Aristotelian dynamics was a general theory of change comprising locomotion, qualitative change, generation, and corruption, and it provided a theoretical basis for witchcraft also. Galileo's dynamics and its successors deal with *locomotion* only, and here again only with the locomotion of *matter*. The other kinds of motion are pushed aside with the promissory note, due to Democritos, that locomotion will eventually be capable of explaining *all* motion. Thus, a comprehensive empirical theory of motion is replaced by a much narrower theory[184] plus a metaphysics of motion, just as an "empirical" experience is replaced by an experience that contains strange and speculative elements. *Counterinduction*, however, is now justified both for theories and for facts. It clearly plays an important role in the advancement of science. This concludes the considerations which started in section 2. For details and further examples the reader is again referred to my "Problems of Empiricism, Part II."

11. Discovery and Justification; Observation and Theory

Let us now use the material of the preceding sections to throw light on the following features of contemporary empiricism: first, the distinction between a context of discovery and a context of justification; second, the distinction between observational terms and theoretical terms; third, the problem of incommensurability.

One of the objections which may be raised against the preceding discussion is that it has confounded two contexts which are essentially separate, viz. a context of discovery and a context of justification. *Discovery* may be irrational and need not follow any recognized method. *Justification*, on the other hand, or, to use the Holy Word of a different school, *criticism*, starts only after the discoveries have been made and proceeds in an orderly way. Now, if the example given here and the examples I have used in earlier papers show anything, then they show that the distinction refers to a situation that does not arise in practice at all. And, if it does arise, it reflects a temporary stasis of the process of research. Therefore, it should be eliminated as quickly as possible.

Research at its best is an *interaction* between new theories which are stated in an explicit manner and older views which have crept into the observation language. It is not a one-sided *action* of the one upon the other. Reasoning within the context of justification, however, presupposes that one side of this pair, viz. observation, has frozen, and that the principles which constitute the observation concepts are preferred to the principles of a newly invented point of view. The former feature indicates that the discussion of principles is not carried out as vigorously as is desirable; the latter feature reveals that this lack of vigor may be due to some unreasonable and perhaps not even explicit preference. But is it wise to be dominated by an inarticulate preference of this kind? Is it wise to make it the raison d'être of a distinction that separates two entirely different modes of research? Or should we not rather demand that our methodology treat explicit and implicit assertions, doubtful and intuitively evident theories, known and unconsciously held principles, in exactly the same way, and that it provide means for the discovery and the criticism of the latter? Abandoning the distinction between a context of discovery

and a context of justification is the first step toward satisfying this demand.

Another distinction which is clearly related to the distinction between discovery and justification is the distinction between *observational terms and theoretical terms*. It is now generally admitted that the distinction is not as sharp as it was thought to be only a few decades ago. It is also admitted, in complete agreement with Neurath's original views, that *both* theories *and* observation statements are open to criticism. Yet the distinction is still held to be a useful one and is defended by almost all philosophers of science. But what is its point? Nobody will deny that the sentences of science can be classified into long sentences and short sentences, or that its statements can be classified into those which are intuitively obvious and others which are not. But nobody will put particular weight on these distinctions, or will even mention them, *for they do not now play any role in the business of science*. (This was not always so. Intuitive plausibility, for example, was once thought to be a most important guide to the truth; but it disappeared from methodology the very moment intuition was replaced by experience.) Does experience play such a role in the business of science? Is it as essential to refer to experience as it was once thought essential to refer to intuition? Considering what has been said in section 4, I think that these questions must be answered in the negative. True—much of our thinking *arises* from experience, but there are large portions which do not arise from experience at all but are firmly grounded on intuition, or on even deeper lying reactions. True—we often test our theories by experience, but we equally often *invert* the process; we *analyze* experience with the help of more recent views and we *change* it in accordance with these views (see the preceding discussion of Galileo's procedure). Again, it is true that we often rely on experience in a way that suggests that we have here a solid foundation of knowledge, but such reliance turns out to be just a psychological quirk, as is shown whenever the testimony of an eyewitness or of an expert crumbles under cross-examination. Moreover, we equally firmly rely on general principles so that even our most solid *perceptions* (and not only our *assumptions*) become indistinct and ambiguous when they clash with these principles. The symmetry between observation and theory that emerges from such remarks is perfectly reasonable. Experience, just as our theories, contains natural interpretations which are abstract and even metaphysical ideas. For example, it contains the idea of an observer-independent existence. It is incontestable that these abstractions, these speculative ideas, are connected with sensations and perceptions. But, first of all, this does not give them a privileged position, unless we want to assert that perception is an infallible authority. And, secondly, it is quite possible to altogether *eliminate* perception from all the essential activities of science (see above, section 6 as well as the appendix). All that remains is that some of our ideas are *accompanied* by strong and vivid psychological processes, "sensations," while others are not. This, however, is just a peculiarity of human existence which is as much in need of examination as is anything else.

Now, if we want to be "truly scientific" (dreaded words!), should we then not regard the theses "experience is the foundation of our knowledge" and "experience helps us to discover the properties of the external world" as (very general) hypotheses? And must these hypotheses not be examined just like any other hypothesis, *and perhaps even more vigorously*, as so much depends on their truth? Furthermore, will not such an examination be rendered impossible by a method that either justifies or criticizes "on the basis of experience"? These are some of the questions which arise in connection with the customary distinctions between observation and theory, discovery and justification. None of them is really new. They are known to philosophers of science, and are discussed by them at length. But the inference that the distinction between theory and observation has now ceased to be relevant either is not drawn or is explicitly rejected.[185] Let us take a step forward, and let us abandon this last remainder of dogmatism in science!

12. Rationality Again

Incommensurability, which I shall discuss next, is closely connected with the question of the rationality of science. Indeed, one of the most general objections, either against the use of incommensurable theories or even against the idea that *there are* such theories to be found in the history of science, is the fear that they would severely restrict the efficacy of traditional, nondialectical *argument*. Let us, therefore, look a little more closely at the critical *standards* which, according to some people, constitute the content of a "rational" argument. More especially, let us look at the standards of the Popperian school with whose ratiomania we are here mainly concerned.

Critical rationalism is either a meaningful idea or a collection of slo-

gans (such as "truth"; "professional integrity"; "intellectual honesty") designed to intimidate yellow-bellied opponents (who has the fortitude, or even the insight, to declare that Truth might be unimportant, and perhaps even undesirable?).

In the former case it must be possible to produce rules, standards, restrictions which permit us to separate critical behavior (thinking, singing, writing of plays) from other types of behavior so that we can *discover* irrational actions and correct them with the help of concrete suggestions. It is not difficult to produce the standards of rationality defended by the Popperian school.

These standards are standards of *criticism*: rational discussion consists in the attempt to criticize, and not in the attempt to prove, or to make probable. Every step that protects a view from criticism, that makes it safe, or "well founded," is a step away from rationality. Every step that makes it more vulnerable is welcome. In addition it is recommended that ideas which have been found wanting be abandoned, and it is forbidden to retain them in the face of strong and successful criticism unless one can present a suitable counterargument. Develop your ideas so that they can be criticized; attack them relentlessly; do not try to protect them, but exhibit their weak spots; and eliminate them as soon as such weak spots have become manifest—these are some of the rules put forth by our critical rationalists.

These rules become more definite and more detailed when we turn to the philosophy of science, and especially to the philosophy of the natural sciences.

Within the natural sciences criticism is connected with experiment and observation. The content of a theory consists in the sum total of those basic statements which contradict it; it is the class of its potential falsifiers. Increased content means increased vulnerability; hence theories of large content are to be preferred to theories of small content. Increase of content is welcome; decrease of content is to be avoided. A theory that contradicts an accepted basic statement must be given up. Ad hoc hypotheses are forbidden—and so on and so forth. A science, however, that accepts the rules of a critical empiricism of this kind will develop in the following manner.

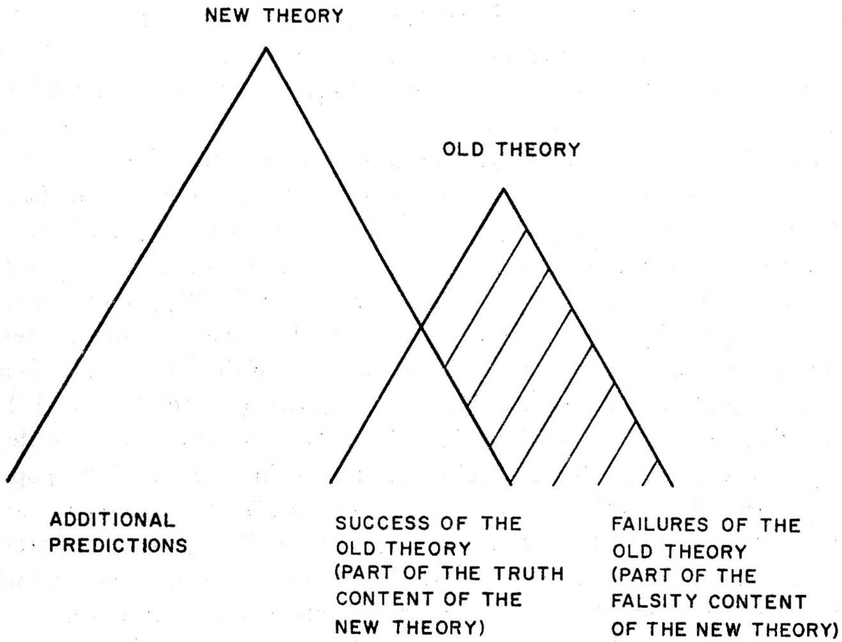
We start with a *problem* such as the problem of the planets at the time of Plato. This problem is not merely the result of *curiosity*, it is a *theoretical result*, it is due to the fact that certain *expectations* have been disappointed: On the one hand it seemed to be clear that the stars must be divine; hence one expects them to behave in an orderly and lawful manner. On the other hand one cannot find any easily discernible regularity. The planets, to all intents and purposes, move in a quite chaotic fashion. How can this fact be reconciled with the expectation and with the principles that underlie the expectation? Does it show that the expectation is mistaken? Or have we failed in our analysis of the facts? This is the problem.

It is important to see that the elements of the problem are not simply *given*. The "fact" of irregularity, for example, is not accessible without further ado. It cannot be discovered by just anyone who has healthy eyes and a good mind. It is only through a certain expectation that it becomes an object of our attention. Or, to be more accurate: this fact of irregularity *exists* because there is an expectation of regularity. After all, the term "irregularity" makes sense only if we have a rule. In our case the rule (which is a more specific part of the expectation that has not yet been mentioned) asserts circular motion with constant angular velocity. The fixed stars agree with this rule and so does the sun if we trace its path relative to the fixed stars. The planets do not obey the rule, neither directly, with respect to the earth, nor indirectly, with respect to the fixed stars.

(In the case just discussed the rule is formulated explicitly, and it can be discussed. This need not be the case. Recognizing a color as red is made possible by deep-lying assumptions concerning the structure of our surroundings and recognition does not occur when these assumptions cease to be available.)

To sum up this part of the Popperian doctrine: Research starts with a problem. The problem is the result of a conflict between an expectation and an observation which in turn is constituted by the expectation. It is clear that this doctrine differs from the doctrine of inductivism where objective facts mysteriously enter a passive mind and leave their traces there. It was prepared by Kant, by Dingler, and, in a very different manner, by Hume.

Having formulated a problem one tries to *solve* it. Solving a problem means inventing a theory that is relevant, falsifiable (to a larger degree than any alternative solution), but not yet falsified. In the case mentioned above (planets at the time of Plato) the problem was to find circular



motions of constant angular velocity for the purpose of saving the planetary phenomena. It was solved by Eudoxos.

Next comes the *criticism* of the theory that has been put forth in the attempt to solve the problem. Successful criticism removes the theory *once and for all* and creates a new problem, viz. to explain (a) why the theory has been successful so far; (b) why it failed. Trying to solve *this* problem we need a new theory that produces the successful consequences of the older theory, denies its mistakes, and makes additional predictions not made before. These are some of the *formal conditions* which *a suitable successor of a refuted theory* must satisfy. Adopting the conditions one proceeds, by conjectures and refutations, from less general theories to more general theories and expands the content of human knowledge. More and more facts are *discovered* (or constructed with the help of expectations) and are then *connected* in a reasonable manner. There is no guarantee that man will solve every problem and replace every theory that has been refuted with a successor satisfying the formal conditions. The invention of theories depends on our talents and other fortuitous circumstances, such as a satisfactory sex life. But as long as these talents hold out the accompanying scheme is a correct account of the growth of a knowledge that satisfies the rules of critical rationalism.

Now, at this point we may raise two questions:

1. Is it *desirable* to live in accordance with the rules of a critical rationalism?

2. Is it *possible* to have both a science as we know it and these rules?

As far as I am concerned the first question is far more important than the second. True—science and other depressing and narrow-minded institutions play an important part in our culture and they occupy the center of interest of most philosophers. Thus the ideas of the Popperian school were obtained by generalizing solutions for methodological and epistemological problems. Critical rationalism arose from the attempt to solve Hume's problem and to understand the Einsteinian revolution, and it was then extended to politics, and even to the conduct of one's private life (Habermas and others therefore seem to be justified in calling Popper a positivist). Such a procedure may satisfy a *school philosopher* who looks at life through the spectacles of his own specific problems and recognizes hatred, love, happiness only to the extent to which they occur in these problems. But if we consider the interests of *man* and, above all, the question of his freedom (freedom from hunger, despair, from the tyranny of constipated systems of thought, *not* the academic "freedom of the will"), then we are proceeding in the worst possible fashion.

For is it not possible that science as we know it today (the science of critical rationalism that has been freed from all inductive elements) or a "search for the truth" in the style of traditional philosophy will create a monster? Is it not possible that it will harm man, turn him into a miserable, unfriendly, self-righteous mechanism without charm and without humor? "Is it not possible," asks Kierkegaard, "that my activity as an objective [or a critico-rational] observer of nature will weaken my strength as a human being?" [186] I suspect the answer to all these questions must be affirmative and I believe that a reform of the sciences that makes it more anarchistic and more subjective (in Kierkegaard's sense) is therefore urgently needed. But this is not what I want to discuss in the present essay. Here I shall restrict myself to the second question and I shall ask: is it possible to have both a science as we know it and the rules of a critical rationalism as just described? And to *this* question the answer seems to be a resounding *no*.

To start with we have seen, though rather briefly,[187] that the actual development of institutions, ideas, practices, and so on often does not start from a problem but rather from some irrelevant activity, such as play-

ing, which, as a side effect, leads to developments which later on can be interpreted as solutions to unrealized problems. Are such developments to be excluded? And if we *do* exclude them, will this not considerably reduce the number of our adaptive reactions and the quality of our learning process?

Secondly, we have seen, in sections 4ff, that a strict principle of falsification, or a "naive falsificationism" as Imre Lakatos calls it, combined with the demand for maximum testability and non-adhocness would wipe out science as we know it, and would never have permitted it to start. This has been realized by Lakatos who has set out to remedy the situation.[188] His remedy is not mine, it is not anarchism. His remedy consists in slight modification of the "critical standards" he adores. (He also tries to show, with the help of amusing numerological considerations, that it is already foreshadowed in Popper.)

According to naive falsificationism, a theory is judged, i.e., either accepted or condemned, as soon as it is introduced into the discussion. Lakatos gives a theory time, he permits it to develop, to show its hidden strength, and he judges it only "in the long run." The "critical standards" *he* employs provide for an interval of hesitation. They are applied "with hindsight." If the theory gives rise to interesting new developments, if it engenders "progressive problem shifts," then it may be retained despite its initial vices. If on the other hand the theory leads nowhere, if the ad hoc hypotheses it employs are not the starting point but the end of all research, if the theory seems to kill the imagination and to dry up every resource of speculation, if it creates "degenerating problem shifts," i.e., changes which terminate in a dead end, then it is time to give it up and to look for something better.

Now it is easily seen that standards of this kind have practical force only if they are combined with a *time limit*. What looks like a degenerating problem shift may be the beginning of a much longer period of advance, so—how long are we supposed to wait? But if a time limit *is* introduced, then the argument against the more conservative point of view, against "naive falsificationism," reappears with only a minor modification. For if you can wait, then why not wait a little longer? Besides there are theories which for centuries were accompanied by degenerating problem shifts until they found the right defenders and returned to the stage in full bloom. The heliocentric theory is one example. The atomic theory is another. We see that the new standards which Lakatos wants to defend either are *vacuous*—one does not know when and how to apply them—or else can be *criticized* on grounds very similar to those which led to them in the first place.

In these circumstances one can do one of the following two things. One can stop appealing to permanent standards which remain in force throughout history, and govern every single period of scientific development and every transition from one period to another. Or one can retain such standards as a *verbal ornament*, as a memorial to happier times when it was still thought possible to run a complex and catastrophic business like science by a few simple and "rational" rules. It seems that Lakatos wants to choose the second alternative.

Choosing the second alternative means abandoning permanent standards *in fact*, though retaining them *in words*. *In fact* Lakatos's position now is identical with the position of Popper as summarized in the marvelous (because self-destructive) Appendix i/15 of the fifth edition of the *Open Society*.[189] According to Popper, we do not "need any . . . definite frame of reference for our criticism," we may revise even the most fundamental rules and drop the most fundamental demands if the need for a different measure of excellence should arise.[190] Is such a position irrational? Yes and no. Yes, because there no longer exists a single set of rules that will guide us through all the twists and turns of the history of thought (science), either as participants or as historians who want to reconstruct its course. One can of course *force* history into a pattern, but the results will always be poorer and less interesting than were the actual events. No, because each particular episode is rational in the sense that some of its features can be explained in terms of reasons which were either accepted at the time of its occurrence or invented in the course of its development. Yes, because even these local reasons which change from age to age are never sufficient to explain *all* the important features of a particular episode. One must add accidents, prejudices, material conditions, e.g., the existence of a particular type of glass in one country and not in another for the explanation of the history of optics, the vicissitudes of married life (Ohm!), superficiality, pride, oversight, and many other things, in order to get a complete picture. No, because, transported into the climate of the period under consideration and endowed with a lively and curious intelligence, we might have had still more to say; we might have tried to overcome accidents, and to "rationalize" even the most

whimsical sequence of events. But, and now I come to a decisive point for the discussion of incommensurability, how is the transition from certain standards to other standards to be achieved? More especially, what happens to our standards, as opposed to our theories, during a period of revolution? Are they changed in the manner suggested by Mill, by a critical discussion of alternatives, or are there processes which defy a rational analysis? Well, let us see!

That standards are not always adopted on the basis of argument has been emphasized by Popper himself. Children, he says, "learn to imitate others . . . and so learn to look upon standards of behavior as if they consisted of fixed, 'given' rules . . . and such things as sympathy and imagination may play an important role in this development." [191] Similar considerations apply to those grownups who want to continue learning, and who are intent on expanding both their knowledge and their sensibility. This we have already discussed in section 1. Popper also admits that new standards may be discovered, invented, accepted, imparted to others in a very irrational manner. But, he points out, one can criticize them *after* they have been adopted, and it is *this* possibility which keeps our knowledge rational. "What, then, are we to trust?" he asks after a survey of possible sources for standards.[192] "What are we to accept? The answer is: whatever we accept we should trust only tentatively, always remembering that we are in possession, at best, of partial truth (or rightness), and that we are bound to make at least some mistake or misjudgement somewhere—not only with respect to facts but also with respect to the adopted standards; secondly, we should trust (even tentatively) our intuition only if it has been arrived at as the result of many attempts to use our imagination; of many mistakes, of many tests, of many doubts, *and of searching criticism*."

Now this reference to tests and to criticism, which is supposed to guarantee the rationality of science, and, perhaps, of our entire life, may be either to *well-defined procedures* without which a criticism or test cannot be said to have taken place, or to a purely *abstract* notion, so that it is left to us to fill it now with this, and now with that concrete content. The first case has just been discussed. In the second case we have again but a verbal ornament. The questions asked in the last paragraph but one remain unanswered in either case.

In a way even this situation has been described by Popper, who says that "rationalism is necessarily far from comprehensive or self-contained." [193]

But our present inquiry is not whether *there are* limits to our reason; the question is *where* these limits are *situated*. Are they outside the sciences so that science itself remains entirely rational; or are irrational changes an essential part even of the most rational enterprise that has been invented by man? Does the historical phenomenon "science" contain ingredients which defy a rational analysis, although they may be described with complete clarity in psychological or sociological terms? Can the abstract aim to come closer to the truth be reached in an entirely rational manner, or is it perhaps inaccessible to those who decide to rely on argument only? These are the problems which were raised, first by Hegel and then, in quite different terms, by Kuhn. They are the problems I wish to discuss.

In discussing these further problems, Popper and Lakatos reject considerations of sociology and psychology, or as Lakatos expresses himself, "mob psychology," and assert the rational character of *all* science. According to Popper, it is possible to arrive at a judgment as to which of two theories is closer to the truth, even if the theories should be separated by a catastrophic upheaval such as a scientific or other revolution. (A theory is closer to the truth than another theory if the class of its true consequences, its truth content, exceeds the truth content of the latter without an increase of falsity content.) According to Lakatos, the apparently unreasonable features of science occur only in the material world and in the world of (psychological) thought; they are absent from the "world of ideas, from Plato's and Popper's 'third world.'" It is in this third world that the growth of knowledge takes place, and that a rational judgment of all aspects of science becomes possible.

Now in regard to this convenient flight into higher regions, it must be pointed out that the scientist is, unfortunately, dealing with the world of matter and of psychological (i.e., subjective) thought also. It is *mainly* this material world he wants to change and to influence. And the rules which create order in the third world will most likely be entirely inappropriate for creating order in the brains of living human beings (unless these brains and their structural features are put in the third world also, a point that does not become clear from Popper's account[194]). The numerous deviations from the straight and rather boring path of rationality which one can observe in actual science may well be *necessary* if we want to achieve progress with the brittle and unreliable material (instruments; brains; assistants; etc.) at our disposal.

However, there is no need to pursue this objection further. There is no need to argue that science as we know it may differ from its third-world shadow *in precisely those respects* which make progress possible.[195] For the Popperian model of an approach to the truth breaks down even if we confine ourselves to ideas entirely. It breaks down because there are incommensurable theories.

13. Incommensurability

Scientific investigation, says Popper, *starts* with a problem, and it proceeds by *solving* it.

This characterization does not take into account that problems may be wrongly formulated, that one may inquire about properties of things or processes which later research declares to be nonexistent. Problems of this kind are not *solved*, they are *dissolved* and removed from the domain of legitimate inquiry. Examples are the problem of the absolute velocity of the earth, the problem of the trajectory of an electron in an interference pattern, or the important problem whether incubi are capable of producing offspring or whether they are forced to use the seeds of men for that purpose.[196]

The first problem was dissolved by the theory of relativity which denies the existence of absolute velocities. The second problem was dissolved by the quantum theory which denies the existence of trajectories in interference patterns. The third problem was dissolved, though much less decisively so, by modern (i.e., post-sixteenth century) psychology and physiology as well as by the mechanistic cosmology of Descartes.

Now changes of ontology such as those just described are often accompanied by conceptual changes.

The discovery that certain entities do not exist may force the scientist to redescribe the events, processes, observations which were thought to be manifestations of them and were therefore described in terms assuming their existence. Or, rather, it may force him to use new *concepts* as the older *words* will remain in use for a considerable time. Thus the term "possessed" which was once used for giving a causal description of the behavioral peculiarities connected with epilepsy was retained, but it was voided of its devilish connotations.

An interesting development occurs when the faulty ontology is *comprehensive*, that is, when its elements are thought to be present in every process in a certain domain. In *this* case every description inside the domain must be changed and must be replaced by a different statement (or by no statement at all). Classical physics is a case in point. It has developed a comprehensive terminology for describing the most fundamental mechanical properties of our universe, such as shapes, speeds, and masses. The conceptual system connected with this terminology assumes that the properties *inhere* in objects and that they change only if one interferes with the objects, not otherwise. The theory of relativity teaches us, at least in one of its interpretations, that there are no such inherent properties in the world, neither observable, nor unobservable, and it produces an entirely new conceptual system for description inside the domain of mechanics. This new conceptual system does not just *deny* the existence of classical states of affairs, it does not even permit us to *formulate statements* expressing such states of affairs (there is no arrangement in the Minkowski diagram that corresponds to a classical situation). It does not, and cannot, share a single statement with its predecessor. As a result the formal conditions for a suitable successor of a refuted theory (it has to repeat the successful consequences of the older theory, deny its false consequences, and make additional predictions) cannot be satisfied in the case of relativity versus classical physics and the Popperian scheme of progress breaks down. It is not even possible to connect classical statements and relativistic statements by an *empirical hypothesis*.[197] Formulating such a connection would mean formulating statements of the type "whenever there is possession by a demon there is discharge in the brain" which perpetuate rather than eliminate the older ontology. Comprehensive theories of the kind just mentioned are therefore completely disjointed, or *incommensurable*. The existence of incommensurable theories provides another difficulty for critical rationalism (and, a fortiori, for its more positivistic predecessors). We shall discuss this difficulty by discussing and refuting objections against it.

It was pointed out that progress may lead to a complete replacement of statements (and perhaps even of descriptions) in a certain domain. More especially, it may replace certain natural interpretations by others. This case has already been discussed (see above, section 6). Galileo replaces the idea of the operative character of all motion by his relativity principle in order to accommodate the new views of Copernicus. It is entirely natural to proceed in this way. A cosmological theory such as the heliocentric theory, or the theory of relativity, or the quantum theory (though

the last one only with certain restrictions) makes assertions about the world as a whole. It applies to observed and to unobserved (unobservable, 'theoretical') processes. It can therefore demand to be used always, and not only on the theoretical level. Now such an adaptation of observation to theory, and this is the gist of the *first objection*, removes conflicting observation reports and saves the theory in an ad hoc manner. Moreover, there arises the *suspicion* that observations which are interpreted in terms of a new theory can no longer be used to refute that theory. It is not difficult to reply to these points.

As regards the objection we point out, in agreement with what has been said before (toward the end of section 4), that an inconsistency between theory and observation may reveal a fault of our *observational terminology* (and even of our sensations) so that it is quite natural to change this terminology, to adapt it to the new theory, and to see what happens. Such a change gives rise, and should give rise, to new auxiliary subjects (hydrodynamics, theory of solid objects, optics in the case of Galileo) which may more than compensate for the empirical content lost by the adaptation. And as regards the suspicion we must remember that the predictions of a theory depend on its postulates, the associated grammatical rules, *as well* as on initial conditions while the meaning of the "primitive" notions depends on the postulates (and the associated grammatical rules) only.[198] In those rare cases, however, where a theory *entails* assertions about possible initial conditions[199] we can refute it with the help of *self-inconsistent observation reports* such as "object A does not move on a geodesic" which, if analyzed in accordance with the Einstein-Infeld-Hoffmann account reads "singularity α which moves on a geodesic does not move on a geodesic."

The *second objection* criticizes the interpretation of science that brings about incommensurability. To deal with it we must realize that the question "are two particular comprehensive theories, such as classical celestial mechanics (CM) and the special theory of relativity (SR) incommensurable?" is not a complete question. Theories can be interpreted in different ways. They will be commensurable in some interpretations, incommensurable in others. Instrumentalism, for example, makes commensurable all those theories which are related to the same observation language and are interpreted on its basis. A realist, on the other hand, wants to give a unified account, both of observable and of unobservable matters, and he will use the most abstract terms of whatever theory he is contemplating for that purpose.[200] This is an entirely natural procedure. SR, so one would be inclined to say, does not just invite us to rethink *unobserved* length, mass, duration; it would seem to entail the relational character of all lengths, masses, durations, whether observed or unobserved, observable or unobservable.

Now, and here we only repeat what was said not so long ago, extending the concepts of a new theory, T, to all its consequences, observational reports included, may change the interpretation of these consequences to such an extent that they disappear from the consequence classes either of earlier theories or of the available alternatives. These earlier theories and alternatives will then all become incommensurable with T. The relation between SR and CM is a case in point. The concept of length as used in SR and the concept of length as presupposed in CM are different concepts. Both are *relational* concepts, and very complex relational concepts at that (just consider determination of length in terms of the wave length of a specified spectral line). But relativistic length, or relativistic shape, involves an element that is absent from the classical concept and is in principle excluded from it.[201] It involves the relative velocity of the object concerned in some reference system. It is of course true that the relativistic scheme very often yields numbers which are practically identical with the *numbers* obtained from CM, but this does not make the *concepts* more similar. Even the case $c \rightarrow \infty$ (or $v \rightarrow O$) which yields identical predictions cannot be used as an argument for showing that the concepts must coincide, at least in this special case. Different magnitudes based on different concepts may give identical values on their respective scales without ceasing to be different magnitudes. The same remark applies to the attempt to identify classical mass with relativistic *rest* mass.[202] This conceptual disparity, if taken seriously, infects even the most "ordinary" situations. The relativistic concept of a certain *shape*, such as the shape of a table, or of a certain temporal sequence, such as my saying "Yes," will differ from the corresponding classical concept also. It is therefore futile to expect that sufficiently long derivations may eventually return us to the older ideas.[203] The consequence classes of SR and CM are not related in any way. A comparison of content and a judgment of verisimilitude cannot be made.[204]

The situation becomes even clearer when we use the Marzke-Wheeler interpretation of SR. For it can be easily shown that the methods of measurement provided by these authors, while perfectly adequate in a relativistic universe, either collapse or give nonsensical results in a classical world (length, for example, is no longer transitive, and in some coordinate systems it may be impossible to assign a definite length to *any* object[205]).

We are now ready to discuss the *second* and most popular objection against incommensurability. This objection proceeds from the version of realism described above. "A realist," we said, "will want to give a unified account, both of observable and of unobservable matters, and he will use the most abstract terms of whatever theory he is contemplating for his purpose." He will use such terms in order either to *give* meaning to observation sentences or else to *replace* their customary interpretation. (For example, he will use the ideas of SR in order to replace the customary CM-interpretation of everyday statements about shapes, temporal sequences, and so on.) Against this, it is pointed out that theoretical terms receive their interpretation by being connected with a preexisting observation language, or with another theory that has already been connected with such an observation language, and that they are devoid of content without such connection. Thus Carnap asserts[206] that "[t]here is no independent interpretation for L_T [the language in terms of which a certain theory, or a certain world view, is formulated]. The system T [the axioms of the theory and the rules of derivation] is in itself an uninterpreted postulate system. [Its] terms . . . obtain only an indirect and incomplete interpretation by the fact that some of them are connected by the [correspondence] rules C with observation terms . . ." Now, if theoretical terms have no "independent interpretation," then surely they cannot be used for correcting the interpretation of the observation statements, which is the one and only source of their meaning. It follows that realism as described here is an impossible doctrine.

The guiding idea behind this very popular objection is that new and abstract languages cannot be introduced in a direct way, but must be first connected with an already existing, and presumably stable, observational idiom.[207]

This guiding idea is refuted at once by noting the way in which children learn to speak and in which anthropologists and linguists learn the unknown language of a newly discovered tribe.

The first example is instructive for other reasons also, for incommensurability plays an important role in the early months of human development. As has been suggested by Piaget and his school[208] the child's perception develops through various stages before it reaches its relatively stable adult form. In one stage objects seem to behave very much like afterimages,[209] and they are treated as such. In this stage the child follows the object with his eyes until it disappears, and he does not make the slightest attempt to recover it, even if this would require but a minimal physical (or intellectual) effort, an effort, moreover, that is already within the child's reach. There is not even a tendency to search; and this is quite appropriate, "conceptually" speaking. For it would indeed be nonsensical to "look for" an afterimage. Its "concept" does not provide for such an operation.

The arrival of the concept and of the perceptual image of material objects changes the situation quite dramatically. There occurs a drastic reorientation of behavioral patterns, and, so one may conjecture, of thought. Afterimages, or things somewhat like them, still exist, but they are now difficult to find and must be discovered by special methods. (The earlier visual world therefore *literally disappears.*) Such special methods proceed from a new conceptual scheme (afterimages occur in *humans*, not in the outer physical world, and are tied to them) and cannot lead back to the exact phenomena of the previous stage (these phenomena should therefore be called by a different name, such as "pseudo-afterimages"). Neither afterimages nor pseudo-afterimages are given a special position in the new world. For example, they are not treated as "evidence" on which the new notion of a material object is supposed to rest. Nor can they be used to *explain* this notion: afterimages arise *together with it*, and are absent from the minds of those who do not yet recognize material objects. And pseudo-afterimages *disappear* as soon as such recognition takes place. It is to be admitted that every stage possesses a kind of observational "basis" to which one pays special attention and from which one receives a multitude of suggestions. However, this basis (i) *changes* from stage to stage; and (ii) is *part* of the conceptual apparatus of a given stage; it is *not* its one and only source of interpretation.

Considering developments such as these, one may suspect that the family of concepts centering upon "material object" and the family of concepts centering upon "pseudo-afterimage" are incommensurable in precisely the sense that is at issue here. Is it reasonable to expect that conceptual and perceptual changes of this kind occur in childhood only? Should we welcome the fact, if it is a fact, that an adult is stuck with a stable perceptual world and an accompanying stable conceptual system which he can modify in many ways, but whose general outlines have forever become immobilized? Or is it not more realistic to assume that fun-

damental changes, entailing incommensurability, are still possible, and that they should be encouraged lest we remain forever excluded from what might be a higher stage of knowledge and of consciousness? (Cf. on this point again section 1, especially on the role of scientific and other revolutions in bringing about such a higher stage.) Besides, the question of the mobility of the adult stage is at any rate an empirical question, which must be attacked by *research* and which cannot be settled by methodological *fiat*. The attempt to break through the boundaries of a given conceptual system and to escape the reach of "Popperian spectacles" (Lakatos) is an essential part of such research (and should be an essential part of any interesting life).[210]

Looking now at the second element of the refutation, anthropological field work, we see that what is anathema here (and for very good reasons) is still a fundamental principle for the contemporary representatives of the philosophy of the Vienna Circle. According to Carnap, Feigl, Nagel, and others, the terms of a theory receive their interpretation in an indirect fashion, by being related to a different conceptual system which is either an older theory or an observation language.[211] This older theory, this observation language, is not adopted because of its theoretical excellence. It cannot possibly be: the older theories are usually refuted. It is adopted because it is "used by a certain language community as a means of communication." [212] According to this method, the phrase "having much larger relativistic mass than . . ." is partially *interpreted* by first connecting it with some *prerelativistic* terms (classical terms, common-sense terms), which are "commonly understood" (presumably, as the result of previous teaching in connection with crude weighing methods), and it is *used* only after such connection has given it a well-defined meaning.

This is even worse than the once quite popular demand to clarify doubtful points by translating them into Latin. For while Latin was chosen because of its precision and clarity, and also because it was conceptually richer than the slowly evolving vulgar idioms,[213] the choice of an observation language or of an older theory as a basis for interpretation is justified by saying that they are "antecedently understood": the choice is based on sheer *popularity*. Besides, if *prerelativistic* terms which are pretty far removed from reality (especially in view of the fact that they come from an incorrect theory implying a nonexistent ontology) can be taught ostensively, for example, with the help of crude weighing methods (and one must assume that they can be so taught, or the whole scheme collapses), then why should one not introduce the *relativistic* terms *directly*, and *without* assistance from the terms of some other idiom? Finally, it is but plain common sense that the teaching or the learning of new and unknown languages must not be contaminated by external material. Linguists remind us that a perfect translation is never possible, even if one is prepared to use complex contextual definitions. This is one of the reasons for the importance of *field work* where new languages are learned *from scratch*, and for the rejection, as inadequate, of any account that relies on 'complete' or 'partial' translation. *Yet just what is anathema in linguistics is taken for granted by logical empiricism*, a mythical "observation language" replacing the English of the translators. Let us commence field work in this domain also, and let us study the language of new theories not in the definition factories of the double language model, but in the company of those metaphysicians, theoreticians, playwrights, courtesans who have constructed new world views! This finishes my discussion of the guiding principle behind the second objection against realism and the possibility of incommensurable theories.

Another point that is often made is that there exist *crucial experiments* which refute one of two allegedly incommensurable theories and confirm the other (example: the Michelson-Morley experiment, the variation of the mass of elementary particles, the transverse Doppler effect, are said to refute CM and confirm SR). The answer to this problem is not difficult either: adopting the point of view of relativity, we find that the experiments, *which of course will now be described in relativistic terms*, using the relativistic notions of length, duration, speed, and so on,[214] are relevant to the theory. And we also find that they support the theory. Adopting CM (with, or without an ether), we again find that the experiments, which are now described in the very different terms of classical physics, i.e., roughly in the manner in which Lorentz described them, are relevant. But we also find that they *undermine* CM, i.e., the conjunction of classical electrodynamics and of CM. Why should it be necessary to possess terminology that allows one to say that it is the *same* experiment which confirms one theory and refutes the other? But did we not ourselves use such terminology? Well, for one thing it should be easy though somewhat laborious to express what was just said *without* asserting identity. Secondly, the identification is of course not contrary to our thesis, for we are now not *using* the terms of either relativity or classical physics, as is done in a test, but are *referring* to them and their relation to the physical world. The language in which *this* discourse is carried out can be classical, or relativistic, or ordinary. It is no good insisting that scientists act as if the situation were much less complicated. If they act that way, then they are either instrumentalists (see above) or mistaken (many scientists are nowadays interested in *formulas*, while the subject here is *interpretations*). It is also possible that being well acquainted with both CM and SR, they change back and forth between these theories with such speed that they seem to remain within a single domain of discourse.

It is also said that by admitting incommensurability into science we can no longer decide whether a new view explains what it is supposed to explain, or whether it does not wander off into different fields.[215] For example, we would not know whether a newly invented physical theory is still dealing with problems of space and time or whether its author has not by mistake made a biological assertion. But there is no need to possess such knowledge. For once the fact of incommensurability has been admitted, the question which underlies the objection does not arise. Conceptual progress often makes it impossible to ask certain questions and to explain certain things; thus we can no longer ask for the absolute velocity of an object, at least as long as we take relativity seriously. Is this a serious loss for science? Not at all! Progress was made by the very same "wandering off into different fields" whose undecidability now so greatly exercises the critic: Aristotle saw the world as a super *organism*, as a *biological* entity, while one essential element of the new science of Descartes, Galileo, and their followers in medicine and in biology is its exclusively *mechanistic* outlook. Are such developments to be forbidden? And if they are not, what, then, is left of the complaint?

A closely connected objection starts from the notion of *explanation* or *reduction* and emphasizes that this notion presupposes continuity of concepts; other notions could be used for starting exactly the same kind of argument. (Relativity is supposed to explain the valid parts of classical physics; hence it cannot be incommensurable with it!) The reply is again obvious. As a matter of fact it is a triviality for anyone who has only the slightest acquaintance with the Hegelian philosophy: why should the relativist be concerned with the fate of classical mechanics except as part of a historical exercise? There is only *one* task we can legitimately demand of a theory, and it is that it should give us a correct account of the world, i.e., of the totality of facts *as seen through its own concepts*. What have the principles of explanation got to do with this demand? Is it not rea-