

LRI Report No. 4C

# **Explanation, Falsifiability, and Rule-following**

Peter G. Ossorio  
1968

LRI Report No. 4C

## Explanation, Falsifiability, and Rule-following

Peter G. Ossorio  
1968



Much has been written and said about the empirical character of psychology and other sciences. Little has been said about the non-empirical aspects. The present paper deals with the question of how what is empirical goes together with what is non-empirical in a science such as physics or psychology. Since a topic of this sort cannot be dealt with in depth in a brief presentation, I shall give a brief review and illustration of a recently developed formulation in which rule-following replaces truth-seeking as the central concept for scientific thinking and procedure.

1. Let us begin directly with a reminder. Psychology consists of two major parts, and they are very distinct parts. One of these is what we normally think of as scientific psychology, that is, psychological theorizing and experimentation. The other part is the discipline of science, what is more commonly called the "Philosophy" of science.

Ordinarily, we do not say that the discipline of science is part of psychology. But for all that we do not say so, we do recognize that it provides us with prescriptions that we accept in regard to the fundamentals of psychological investigation. Among the fundamentals are these five:

- (a) The difference between science and non-science.
- (b) The difference between doing an experiment and doing anything else.
- (c) The difference between empirical and non-empirical.

(d) The difference between what is a scientific explanation and what is not.

(e) The place of definition and experimentation in scientific explanation.

These five are merely illustrative. If our theorizing and experimentation could not be seen as embodying certain decisions in regard to fundamentals such as these five, they would not, and could not, have the significance that they do for us as scientific endeavors. Thus, decisions of this sort are not outsiders' views of what goes on inside psychology. They are not merely philosophers' theories about what psychologists do. They are part of our technical apparatus for doing psychology.

Let us move to another reminder. This division of psychology into two parts goes with the distinction of empirical vs. non-empirical. In the sense in which psychological science is regarded as empirical, the discipline of science is as thoroughly non-empirical as other disciplines such as literature, art criticism, and theology.

Specifically, the discipline of science is non-empirical because the decisions we have accepted in regard to those five fundamentals (for example) do not represent experimental findings or generalizations of experimental findings, or even explanations of experimental results. Nor could they. Neither do they represent stipulative definitions. The psychological scientist is free to define his explanatory concepts and his measurement operations. He is not free to define what an experiment is, or what an explanation is, etc. It is because the fundamentals of psychological procedures are constituted by the application of certain social standards (rather than investigation) and are maintained by the social institutions of psychology and other sciences (rather than by investigation or inquiry) that the designation "the discipline of science" was adopted as being descriptive of this part of psychology.



If, say, one percent of all human behavior is scientific behavior, then the discipline of science may be characterized as a single, non-empirical, non-scientific theory of that one percent of human behavior. The other essential part of psychology consists (for our present purposes) of a variety of empirically oriented behavior theories each of which applies to all behavior. None of these theories overlaps at all with the non-scientific theory of scientific behavior, nor could any of them replace the latter. Thus, we have more than one hundred percent coverage with respect to behavior, and that is curious. The non-scientific type of formulation appears to have the virtue that with it we can say what needs to be said about scientific behavior, i.e., to characterize it effectively, to distinguish it from other forms of human behavior, and to explain it. But our special formulation of scientific behavior leaves unexamined the other ninety-nine percent of human behavior. In contrast, with our scientific theories there is no behavior we cannot talk about, but evidently there is much about behavior that we cannot say. We cannot say as much about behavior as must be the case in order for there to be anything that would qualify as a scientific theory of behavior. So that what falls outside the scope of our scientific explanations of behavior is not irrelevant for the science of behavior or for a scientific understanding of behavior, and so cannot be dismissed by pointing out that science is abstract and always leaves out of consideration some aspects of the concrete phenomena. Curious indeed, and perhaps disquieting, for if our scientific theories of behavior are defective in this way with respect to scientific behavior, we may well wonder whether they are not correspondingly and systematically defective with respect to the other ninety-nine percent of human behavior.

On the face of it, the division of psychology into an incomplete empirical part and an incomplete non-empirical part is undesirable because it guarantees that no psychological theory developed within this framework will ever meet some



of our ordinary standards of adequacy. As it happens, one of the major implications of the rule-following formulation is that this particular division into empirical and non-empirical is in no sense necessary. I shall suggest that one hundred percent coverage is enough, after all, that it is possible within a single psychological behavior theory to say what there is to be said about all behavior, including rat behavior, scientific behavior, and philosophical behavior. There is such a theory (Ossorio, 1966), and it is paradigmatically a rule-following theory of human action. Since the focus of the present paper is methodological rather than theoretical, the theory as such will not be presented, but it will be referred to from time to time as the "general rule-following model."

II. Let us turn to the question of what is empirical about particular sciences, particularly psychology. It appears that our most familiar and generally accepted views of the matter can be summarized in this way: "Scientific laws are empirical because they summarize what we observe, and scientific theories are empirical because they are tested against what we observe."

In order to approach the problem with due caution and remain on neutral ground as long as possible, let us, rather than going directly to psychological problems, begin by examining a classic formula from a classic science. This is the formula that says "A physical body will accelerate in the direction of an applied force." This is an incomplete and colloquial rendering, but nothing hinges on that. The formula is of particular interest because it may be thought of as either an empirical law or a theoretical, hence explanatory, statement.

Using this familiar example as a vehicle, I shall argue for two conclusions. First, that as summaries of what we observe, our empirical laws are not merely falsifiable, but demonstrably false. Second, that our major theories, and particularly our major theories of behavior, cannot be falsified at all. Both conclusions raise some question about our modes of experimentation in psychology



and our traditional accounts of what is empirical about science. It is not so much that these conclusions have been directly denied, but rather that their force has not been adequately dealt with in accounts which represent scientific endeavor as being essentially a search for truth even though we have to settle for confirmation. A rule-following characterization of scientific behavior is presented briefly as an alternative to the traditional soothsaying model, and this is followed by a psychological example which illustrates the new outlook on how empirical and non-empirical fit together in psychological science.

Let us, then return to the formula, "A physical body will accelerate in the direction of an applied force," and let us return with some more reminders. Do we always observe it to be the case that a physical body accelerates in the direction of an applied force? Of course not. If there were any doubt, a negative instance could be produced on the spot. For example, we could apply a horizontal force to a pencil on the table in such a way that it suddenly accelerated in a non-horizontal direction, i.e., when it rolls off the edge. Conversely, if we apply a moderate horizontal force to the table itself, it will not be observed to move at all. Thus, if the primary virtue of such a formula were that it was a general statement of what we observe and testable against future observations, we could discard this particular one because, simply, it is false.

Of course, we do not discard the formula. That it can be falsified in those ways does not bother us at all. That this is so is the important fact. We are not bothered because we have explanations for the failures. In the first case, we neglected the force of gravity, and in the second case we neglected the force of friction between the table and the floor. To be sure. (Of course, these explanations are already a move away from observables.) But suppose we did take account of the force of gravity and the force of friction. Probably we could then do a better job of predicting movements generally, though there would be occasions



when we would not. There would still be some times when our predictions were in error--some times, and no matter how much care we took. This is not a conclusion that is traditionally denied. Rather, it has merely been referred to the "open texture" of scientific concepts and procedures and been declared not to be a problem in many important cases. But then our previous conclusion holds--as an empirical summary, the formula is not merely falsifiable, but demonstrably false.

Still, that conclusion does not bother us, and there is more to it than that it is not always a problem. If we did an experiment and conscientiously established that forces  $x$ ,  $y$ , and  $z$  were operating on our experimental body, and if the body failed to move in the expected direction, we might then claim to have falsified that law of motion empirically. However, such a claim would get us nowhere. The proper physicist would merely scowl at us and say, "Don't be silly. Obviously there was some other force operating." And then we might say, "Oh--so it isn't simply 'a physical body will accelerate in the direction of an applied force.' There has to be a qualifying clause that says, 'unless there is another force operating.'" And the physicist might frown and say, "All right. But the neat way to write the formula, the one that shows its objective and universal character most clearly, is to say, 'A physical body will accelerate in the direction of the resultant of all the forces acting on it.'"

Of course, the physicist is quite right. That does show its objective and universal character. It also shows the formula to be non-falsifiable. Since we do not have a general criterion for that resultant force that is independent of the movement of that body, no set of empirical results is incompatible with the latest versions of the formula. With these latter two we have moved to the second alternative, which is that the formula is no longer an observation summary, but a theoretical statement, hence explanatory.



That the theoretical statement is not falsifiable still does not bother us, nor does it bother the physicist, and that says something about our sophistications, since we also want to say that the theory in question is empirically tested. By now, however, the answers to why it does not bother us are not so easy to come by, nor is it crystal clear any longer that we ought not to be bothered. Certainly, it is beginning to be clear that there is a very substantial logical gap between our scientific accounts of the world and what we observe to be the case. We often hear that such accounts represent what scientists have discovered to be the case, or what they have shown to be the case, and we often talk that way among ourselves. But we might question whether "discovery" or "invention" is the more appropriate description. We might question whether "discovery" and "invention" do not amount to the same thing in this case. And there may well be something disturbing about the conclusion that if we want a formula which explains observations by reference to an objective, universal principle, the price of having it is that the formula be non-empirical and removed from observation. In this connection, we may note that it is not until the formula is protected by the "unless" clause that we say that the body moved because of the forces operating on it--so long as the formula is regarded as an empirical law any such statement would be dismissed as circular and non-explanatory.

Not very long ago the conclusion that our theories are not falsifiable was the occasion for some concern. It was in this connection that the notion of verification was replaced by a more elaborate account in terms of confirmation. In psychology, the story of confirmation has come to be accepted largely in the even more elaborate form of "construct validation." The essential feature of the confirmation story is that we test a theory by making a prediction on the basis of the theory. The theory is confirmed or disconfirmed accordingly as the prediction is true or false. Extensive confirmation of a theory gives us confidence



in its truth, although such evidence is never at all conclusive, and so the theory cannot be shown to be true. Conversely, disconfirmation of a theory provides us with evidence of its falsity. For some theorists, this kind of evidence can be decisive; for most it is not. In any case, disconfirmation is grounds for giving up the theory, or at least, for changing it or changing our minds about it.

No doubt this may be regarded as a barbarously truncated version of the "confirmation" rationale for scientific endeavors. Nevertheless it is enough to point up two kinds of difficulty which it appears may be found in connection with more complete accounts. The first stems from the notion that the point of engaging in experimentation is to have an (empirical) evidential basis for a truth appraisal of the theory and that the ideal would be to have a true theory. The difficulty, briefly, is that with non-falsifiable theories there is no sensible question of their being true or false, and so there would be no sense in any truth appraisal and there could hardly be any sensible question of evidence for or against the truth of such theories.

The second difficulty depends on an alternative reading of the confirmation account. Although there seems little doubt that a concern for the truth of scientific theories is in fact the keystone of the confirmation account it might be claimed (particularly in the light of the first difficulty noted above) that a literal and conservative reading of such accounts will show that they only go as far as saying that extensive confirmation per se, and not any further implication about truth value, is the standard for positive appraisal and acceptance of a scientific theory. On this reading, however, the confirmation account drops out altogether as a rationale for scientific behavior, for it merely repeats in unqualified form something that was already taken for granted subject to qualifications, i.e., that, in fact, under the circumstances that prevail, all other things being equal, a theory which permits extensive correct predictions



is more valued than one which does not. What is called for, and what the confirmation view fails to provide, is a systematic account of what sense it makes to value a highly confirmed theory over one that is not.

Thus, in either case the confirmation story does not appear to provide an adequate account of why as scientists we would want a well confirmed theory as against, for example, one which was merely known to be true, or why it would make sense to change our minds about a non-falsifiable theory in the face of negative experimental results. In point of fact, it appears that in the face of negative findings the theorist gives up his theory or modifies it--unless he doesn't. It is instructive in this connection to reflect on the kind of criticism which we in fact level at ourselves or our colleagues. Aside from gross technical ineptitude, which is perhaps relatively rare, the criticism which counts most with us deals with such things as (a) triviality, or uninformativeness of an experiment, (b) problems of generalizability, or (c) playing it too safe in the selection of occasions or predictions which constitute the empirical "test" of a theory. These grounds of criticism are, of course, interrelated. None of these grounds is touched on in the confirmation story. Yet it would seem that if we had an adequate rationale for scientific behavior, our criticism of particular scientific endeavors would consist primarily in pointing out the ways in which those endeavors fail to conform to the scientific rationale.

III. There is an alternative to the confirmation story as a rationale for scientific behavior. I take it to be simpler and more to the point. Very briefly:

An explanatory formula such as the one about moving bodies does not function primarily as a simple description (hence true or false) of what one observes or expects to observe. Neither does it function as a premiss for such a description. Instead, it is a prescription followed by the scientist in describing what he observes. It is a conditional prescription to the effect that the observed results



must be described in accordance with the format provided by the formula if is is to be a description of a certain kind of phenomenon. Thus, in the previous example, the prescription would be "The acceleration of a body must be expressible as the effect of the resultant of the forces operating on it, unless it isn't a physical body." (Recall that the initial formula was, "A physical body will accelerate....")

If we return to the formula that "A physical body accelerates in the direction of an applied force--unless there is another force operating," we see that the "unless" clause guarantees that the prescription is one which can actually be followed. We can do that because if there is a force which we didn't know about in advance, we can and often do calculate after the fact what it must have been. Sometimes we even repeat the experiment and try to nail it down in advance. If establishing the truth of the theory were the primary point of experimental investigation, then it would seem paradoxical that non-falsifiable theories should be the rule and not the exception. In contrast, if experimentation is primarily a matter of following the prescriptions codified in the theory, then having prescriptions which unquestionably can be followed is a technical and methodological virtue. But also, we must then look further for an account of what the empirical point of experimentation is, since that cannot consist simply in having followed a theoretical prescription which we knew in advance could always be followed.

In the traditional view, what accrues to a theory by virtue of confirmation is truth, or our confidence in its truth. And if we have that confidence in the theory, we need not worry about its range of application, because that will be limited only by logical considerations. If it is true that a physical body moves in the direction of the resultant of all of the forces operating on it, why then it does, and there are no ifs, ands, or buts about it. This is why our standard



of adequacy for a behavior theory is that it should apply to all behavior. It is also why we have that other two-way division in psychology--the division between those who discover the truth and those who apply it. This arrangement has at least two predictable drawbacks. The first is that it would be easy to do (and difficult to avoid doing) a great deal of research that had little or no payoff because it did not add appreciably to our conviction about the truth of the theory in question. Experiments of this kind are among those which we criticize as trivial or uninformative. The second is that it would often be difficult or impossible to generalize the theory in a non-trivial way beyond the laboratory setting or beyond the clinic setting, or whatever the original domain of application of the theory was. That this has been the case historically hardly needs documentation.

The rule-following account of the matter is that nothing accrues to a theory by virtue of confirmation--nothing in particular, and nothing necessarily. (This follows directly from the recognition that experimentation can be trivial and uninformative.) Instead, the open question is where and how our prescription can be followed in a non-trivial way and with effective results. This is a matter of genuine concern, and since we do not in general know the answers to questions of this sort in advance, it is these questions which are the empirical ones. Thus, the point of experimental investigation is to get information of this sort. Within the rule-following framework, it is apparent that the most informative sort of experiment is the one in which we try out our prescription in just those circumstances where there is a real question as to whether it can be done effectively. (In a general way, this corresponds to the "determined efforts to falsify" theories which Popper recommends.) This kind of experiment will be exceptionally informative even if it adds nothing in particular to our confidence in the truth of what we say. In the face of negative findings we change our minds about the range of



effective application of our prescription, since that is what we discover empirically. But of course, we may change our minds in this way in the face of positive findings, too. In the rule-following approach, hypothesis testing has no particular virtue, though it is in no way ruled out. The emphasis is on what we discover, not what we said in advance, (but see below).

It should be clear from the foregoing that the problem of generalization is central, not incidental to the rule-following approach. The range of effective application of a non-empirical formula has to be discovered empirically, and so generalization is something to be done, and it is done by doing an experiment. This contrasts with the traditional approach, where generalization is something that is said, and said after doing an experiment (the experimenter announces that he "generalizes" his results to populations of which his experiment is representative).

Still, there is something missing. Up to this point we have seen that the use of non-empirical formulas is consistent with empirical research, since the empirical questions have to do not with the formulas per se, but with their use. But it may well appear that, under this account, science consists simply of fact-gathering relative to the use of a variety of explanatory formulas. This conclusion would violate the traditional ideology in which scientific endeavor is seen as an open-ended search for more and more general and fundamental explanations. However, something of this sort is a feature of the rule-following approach also. For, given some empirical information (or even some expectations) about the range of effective application of a given formula,  $f_1$ , we might hope to invent a new prescription,  $F_1$ , dealing with the use of  $f_1$  and possibly others as well. One criterion for the effective application of  $F_1$  would be that it should enable us to judge ahead of time on what occasions and in which ways  $f_1$  and others of the same sort could be used effectively. (It is this use of  $F_1$  which would lead us to have expectations or hypotheses in connection with the use of  $f_1$  on



each new occasion.) As will be illustrated by the psychological example below, different prescriptions or explanatory formulas, may be related in ways which are more complex than the simple hierarchy suggested by the foregoing. In any case, in virtue of the relationships among explanatory formulas, the rule-following formulation does provide for depth in explanation as well as scope, but it is the evolution of behaviors that is primary, not the evaluation of theories or the simple gathering of facts. The scientific investigator is an innovator not because he discovers truths which others may subsequently "apply," but because he sets a new example of a form of behavior which others may subsequently follow. And if we can foresee some possible difficulties in trying to follow his example when our circumstances are not entirely the same as his, that will be no more than the difficulty we would face in "generalizing" or "applying" the truth he has "discovered" or "confirmed" when our circumstances are not entirely the same as his.

In sum, the rule-following account of scientific behavior provides a general alternative to the truth-seeking account. It has been presented as providing a more adequate account of the gross facts of scientific behavior, e.g., the common use of non-falsifiable explanatory formulas and the standards by reference to which scientific achievements are criticized and appraised. The psychological example and discussion below are designed to show that the rule-following approach makes unnecessary some of the questionable dichotomies assumed by the truth-seeking approach, e.g., the dichotomy between a non-empirical "philosophy" of scientific behavior and an empirical science of behavior, or between a purely nominal "observation language" and an explanatory "theoretical language." In providing the means for discarding these a priori dichotomies, the rule-following account also discards the reductive and atomistic, hence anti-psychological, bias inherent in the avowedly "neutral" positivistic account.

IV. Let us turn now to a psychological example to illustrate the prescriptive, or rule-following, approach. For this purpose, let us examine another classic formula, the one that says, "Frustration leads to aggression." Originally, this was propounded as an empirical law, and it gave rise to much research and discussion for a number of years. It disappeared from the psychological scene because there was too much negative evidence, and the length to which one had to go to explain away negative findings made it unacceptable as an empirical law. Research on "aggression" has continued, however, and a recent review of such research ends with the suggestion that "aggression" be treated as a hypothetical construct rather than as an observational term. It took roughly thirty years for that suggestion to be made. The following is a schematic presentation of a rule-following approach to the phenomenon. That it did not take thirty years to evolve reflects another dimension of the comparison between the rule-following account and the truth-seeking, confirmation account.

We begin with the formula

- (1) Provocation by O elicits a correspondingly hostile response by P."

This is the analogue of the simple formula concerning physical bodies, and I take it to be a more linguistically sensitive formulation of the frustration-aggression hypothesis. Like both of these latter, our hostility formula is false. What is required is an "unless" clause. In fact there are several such clauses, and so we have:

- (1) Provocation by O elicits a correspondingly hostile response by P.
  - (a) Unless P has another reason for showing anger toward O or for not showing anger toward O. (This is the direct analogue of "unless there is another force operating.") Or

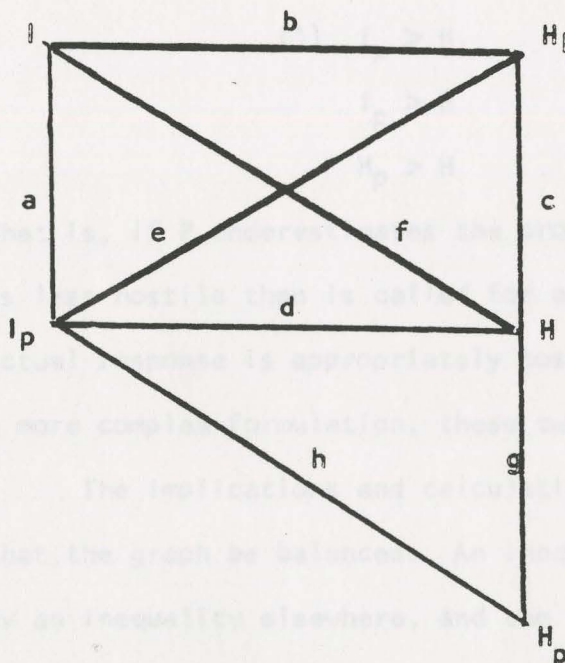


- (b) Unless P doesn't perceive O's behavior as the provocation that it is, or
- (c) Unless P is unable to express his anger in that situation, or
- (d) Unless P believes that what he did was a correspondingly hostile response.

These "unless" clauses may be specified in greater detail. For example, we might talk quantitatively in terms of the relative strength of the provocation by O as against P's other reasons for not showing his anger. Such an analysis has been given an effective computer simulation by Mitchell (1966). Or we might specify discretely what difference it would make if P's reason for not showing anger was fear rather than, say, avoidance of guilt.

Moreover, we have something like a computational scheme which provides us with an analogue (but only an analogue) to calculating what that unknown force must have been. Consider a graph as in Figure 1, with five nodes representing elements of the expanded hostility formula.

Figure 1. The Hostility Graph



- I degree of provocation by O
- $H_O$  degree of hostility appropriate to I
- $I_P$  degree of provocation perceived by P
- H degree of hostility shown by P
- $H_P$  degree of hostility perceived by P to have been shown by P

The "paradigm case" of this graph is one in which all the nodes have the same value, so that all the connecting lines (a through h) have a value of zero. This corresponds to the condition that P correctly perceives the degree of provocation, shows the appropriate degree of hostility, and knows how much hostility he has shown. This paradigm case is like the physical body in uniform linear motion, i.e., it requires no explanation, and moreover, no non-trivial explanation is possible within the conceptual system in question. Thus in the situation represented by the paradigm case of the hostility graph, if a question arose as to why P acted as he did, to point out that he acted in response to that provocation would be an explanation which did not require any further explanation along the same lines.

The paradigm case, however, is only one of the possible configurations of the hostility graph. There are other possibilities. If we begin with any one of the connecting lines (a through h) and suppose it to have a non-zero value (except that line b is exempt) other inequalities are implied, either absolutely or conditionally, and can be calculated analytically. For example, we may begin with the case where P underestimates the provocation by 0 ( $I_p < I; a < 0$ ). Then one of the two following sets of additional inequalities holds:

$$\begin{array}{ll}
 (1) & I_p > H_1 \\
 & I_p > H \\
 & H_p > H \\
 (2) & I_p > H_1 \\
 & H > I \\
 & H > H_1
 \end{array}$$

That is, if P underestimates the provocation by 0, then either his actual response is less hostile than is called for and he perceives it accurately, or else his actual response is appropriately hostile, but he underestimates that also. (In a more complex formulation, these two alternatives are not mutually exclusive.)

The implications and calculations reflect the prescriptive requirement that the graph be balanced. An inequality in one place must be compensated for by an inequality elsewhere, and the second inequality may require a further one



to balance it, and so on. Because of these connections, the total number of balanced graphs is quite small--four basic configurations and their mirror images. Aside from the paradigm case, the balanced graphs either directly represent the use of one of the "unless" clauses in the hostility formula or permit further explanation by reference to one of the "unless" clauses. For example, the case of P underestimating the provocation by 0 falls under clause 2.

V. Let us turn from the hostility example per se to an examination of its methodological interest.

The prescriptive formula, "Provocation by 0 elicits a correspondingly hostile response by P, unless..." represents neither an empirical discovery nor a stipulative definition. Instead, it is a partial formulation of our familiar four-thousand-year-old concept of anger. It is partial because there is more to be said about anger than just that formula. And it is partial because it requires the general rule-following model to say as much as there is to be said about anger. It is because of this part-whole relationship that the "unless" clauses of the hostility formula can make use of other concepts in the larger system. Just as in the physics example the theory of mechanics made it possible to identify other physical bodies as sources of potentially complicating other forces, the general rule-following model permits us to identify fear, guilt, prudence, and so on as reflecting particular reasons, other than anger, for acting. Other motivations of these kinds may therefore reinforce or counter-balance the reason provided by 0's provocation. And it goes without saying that we take P's behavior to be a function of all his reasons jointly, and not just an expression of the reason provided by 0's provocation. ("The acceleration of a physical body is in the direction of the resultant of all the forces acting on it.") The computer implementation mentioned above (Mitchell, 1966) provides an empirical evaluation of two functions for computing a "resultant" reason and predicting behavior therefrom.



One important point to be made in connection with the hostility formula and others like it is that until the set of "unless" clauses is complete, the formula is falsifiable, but only in a peculiar way. An incomplete formula is falsifiable in the sense that some experimental result may be encountered which was not specified in advance as a possible outcome. But this kind of falsifiability is not a case of testability in the truth-seeking sense. It is a spurious testability because positive results do not affect the adequacy of the incomplete formulation--it remains incomplete and can be shown to be so without experimentation. And negative results of this sort are not grounds for giving up the content of the incomplete formulation--only for extending it.

This conclusion is particularly important because it appears that during the past three decades, as a result of equating positivism and truth-seeking with science, psychologists have given top priority and official sanction to implementing prescriptions of "empirical testability," "operationalization," and "parsimony." In giving automatic precedence to implementing these three prescriptions simultaneously in the positivistic style, it appears that psychologists have almost inevitably been led to give oversimplified, incomplete formulations which are empirically testable only in the spurious sense mentioned above, with the result that the most common, and widespread, criticism made by psychologists in regard to psychological research is that it is "trivial" or "uninformative." Let us use the hostility formula again to illustrate how such research might come about.

Suppose we began with the simple hostility formula and set up a classic sort of experiment with an operational definition and a predicted outcome. All three prescriptions are embodied here. The simplicity of the formula gives it the virtue of parsimony, and the fact that one could possibly get negative results for the prediction gives it the virtue of "empirical testability."



If we chose our experimental settings judiciously, we might do many experiments confirming the simple formula, perhaps many years of research, before we encountered a negative result. Let us suppose that we accounted for the negative result by reference to the first "unless" clause on our list, i.e., "unless P has another reason for not showing anger toward O." We could do further confirming experiments until negative results were encountered again, whereupon the next "unless" clause would see the light of day. And so on to the third and fourth "unless" clauses and to the more specific details of each.

Let us remember, however, that if we were not already prepared to allow the additional "unless" clauses, no amount of experimental evidence would require us to do so, and more importantly, no amount of experimental evidence would enable us to do so. For example, if we were not already prepared to accept the second "unless" clause, i.e., "P does not recognize O's provocation for what it is," we would deal with negative experimental results by saying, "P must have had other reasons for not showing his anger toward O, even though it's not clear what they were." And in this case, if someone were to suggest, "Perhaps P doesn't perceive the provocation for what it is," we would not say, "You're wrong," but rather, "What are you talking about? Don't change the subject." Of course, we don't say that.

But even if we recognized the suggestion as apropos, we might well reject it on the basis that the experimental results didn't require that interpretation. With a strong concern for parsimony at work, we would never have to go beyond the first "unless" clause no matter what our experimental findings were. Just as P may have reasons for not showing the anger he has toward O, he may have reasons for not saying what he knows about O. If P is a psychologist with a heavy obligation to be surprised at his experimental results (because they are empirical) and an equal obligation to give reductive "explanations" of O's



behavior on the basis of those results, he will have a variety of reasons for not saying what he knows about O.

To continue: When we were well along with this long, hypothetical series of experiments resulting in the detailed articulation of "unless" clauses, we might, in explaining what we did for the benefit of a colleague or in justifying what we did for the benefit of an editor or a funding agency, describe ourselves as having evolved a "nomological network" and as having empirically established the construct validity of our constructs of provocation and hostility. Of course, the concepts of provocation and hostility need not be expressed literally by the words "provocation" and "hostility." They might be formulated in a different linguistic system (e.g., in French, Spanish, German, etc.) or in idiomatic English (e.g., as "instigative stimulus" and "aggressive response" or as "aversive stimulus" and "aversive response").

But now, what was empirical about it all? Certainly, for example, the hostility graph, the formula, and the detailed "unless" clauses did not pop up out of nowhere. In fact, they emerged in the course of discussions with a colleague, Milton Lipetz. The discussions involved false starts, correction of factual errors, premature closures, checking against clinical examples and clinical explanations, doing thought experiments, and reflecting on our previous research on hostility. It is these historical, biographical features of the situation, what has sometimes been called the "context of discovery," to which the construct validity account of evolving connections and cross-checking them seems most apropos.

In the sense in which it took some doing and did not emerge as a stipulative definition, the hostility formula is "empirical." In this sense, however, art criticism, accounting, and theology are also empirical. What is not empirical is the content of the formula and the logical interconnections among the



elements of the formula, as well as the logical interconnections between the formula and the remainder of the general rule-following model. The fact that we sometimes have to be reminded of some of the connections among our concepts does not make those connections empirical ones, nor does it give our conclusions about those connections the status of empirical discoveries. To use a common analogy, we have to learn arithmetic too, and if we didn't have eyes in our heads and look around us, we might well not learn it, but that does not make the rules of arithmetic empirical--on the contrary, what we have to learn is precisely that they are not empirical.

There is more than a passing analogy between the use of the rules of arithmetic and the use of behavior-descriptive formulas such as the hostility formula. In both cases, the primary value of having something of the sort is not that they are part of a true story about the world, but rather that by virtue of being non-empirical, they can be used effectively in some form of human behavior. Moreover, it appears that their effective use cannot be duplicated without them.

We may here amplify the earlier statement to the effect that what was empirical about explanatory formulas was not their content, but their range of effective application. It is true, for example, that the concepts of hostility and provocation have numerous logical relationships with other behavioral concepts, and these relationships are not empirical. But every such distinction and every such relationship is the logical precondition for a great many empirical possibilities. For example, we may focus on a particular sort of case, e.g., the case where P has a stronger reason not to show anger toward O. We may then ask such empirical questions as

- (a) For whom will which reasons be stronger reasons, and when?
- (b) Are facts of the first sort systematically related to group



membership, historical antecedents, current physiological states, or anything else whatever?

- (c) What human actions could change facts of the first sort? How could facts of the second sort be used to facilitate or inhibit such changes? What human actions could accomplish changes in facts of the second sort?

The point here can be stated simply: The facts we are able to discover empirically are completely and directly dependent on the prescriptive formulas we are able to use, whereas the evolution of new formulas is not a matter of discovery at all and is only very loosely dependent on the facts we are able to discover. But the evolution of new formulas can be expressed directly as the extension of the range of effective application of the concept of human behavior.

Finally, let us return to some of the issues that were raised previously.

It was suggested earlier that there might be something disturbing about the conclusion that the price of having a formula which explains our observations by reference to an objective, universal principle is that the formula be non-empirical and removed from observation. We have seen that the non-empirical feature is obscured in the traditional formulation by being phrased in the truth-testing language of "confirmation," "evidence," and the "empirical" derivation of nomological networks. The non-observability feature is handled in a more positive, but also more radical fashion. The non-observable elements of the explanatory formulas for behavior are identified hypothetically with observables after all, but now of a non-behavioral sort, and the prescriptive basis for this identification is the "hypothesis" of "the unity of science." Following this prescription requires treating psychology not as a science, but as the peripheral fragment of a single, all-embracing science in which the fundamentals lie elsewhere. In this genre, psychological explanations are not real explanations, but merely I. O. U.'s



deposited against the future findings of physiologists, genetecists, and biochemists.

In contrast, the hostility formula and others like it in the general rule-following model, are not I. O. U.'s. For example, if under provocation by O, P is observed not to show hostility, the statement that he had another reason, i.e., fear, for not showing his anger does serve as an explanation, and not a second best one, either, of why he did what he did. The very same system of logical relationships which determines that not showing anger in the face of provocation is a piece of behavior which requires an explanation also provides the standard for what would qualify as an explanation, and it provides the resources for giving one. What qualifies as an explanation is the specification of the applicable "unless" clause from among those "unless" clauses which the system itself provides. Similarly, to return briefly to our heuristic analogy, if a physical object is observed to move otherwise than in accordance with the forces which are known to be operative, then not merely is it clear that an explanation is required, but also the nature of the required explanation is quite clear. It consists of specifying what other force was operating, and no other explanation will do in place of this. This, very briefly, is why the question of what explains what is not an empirical one and why non-behavioral facts cannot explain behavioral facts. No less than the facts we are able to discover, the explanations that we are able to give depend wholly and directly on the prescriptive formulas that we are able to use.

On the whole, therefore, the rule-following account of scientific human behavior can be said to lack the problematic features of the truth-seeking, "confirmation" account. We have seen that in the rule-following formulation there is no dichotomy and no category boundary between a non-empirical discipline of science and an empirical science of behavior, since the prescriptive aspect extends all the way down to the collection of data, the use of data, the



explanation of data, and the generalization of such explanations. No doubt it has not been made explicitly clear that it extends to all data, including rat data, and all explanations, including physiological and physical ones. In any case, because of this continuity, one hundred percent coverage is sufficient for an adequate account of all behavior.

Likewise, with the provocation-hostility example, we have seen that in the prescriptive approach explanation of what is observed is given by the systematic relationships among concepts in a comprehensive system (in contrast to the simple reductive hierarchy embodied in the unity of science "hypothesis"). Because of this, no substantive distinction is required between descriptive and explanatory concepts and no language other than descriptive language is required at all. In the formulation of hostility, there is no distinction to be made between "observation language" and "theoretical language," hence there is no particular encouragement, and certainly no requirement for the curious a priori principle that what looks like behavior is really something else.

VI. Neither the confirmation account nor the rule-following account provides a prescription which, if dutifully followed, guarantees a worthwhile scientific accomplishment. Neither, of course, do we have prescriptions that guarantee good literature, art, music, philosophy, mathematics, education, or psychotherapy. All of our experience in these disciplines reminds us of the crucial gap between substantial achievement and the mere impeccable exercise of technique. At this most important juncture, ironically, it is the truth-seeking, positivistic account which provides us with a prescription that can surely be followed. We cannot always accomplish anything of scientific value, but we can always use simple formulas couched in an explanatory idiom, operationalize, predict, and collect empirical data. In contrast, the rule-following model has no special place for a hypothetical mere science as against good science, any



more than it has a special place for mere arithmetic as against correct arithmetic, or for anything that would be mere music as against musical music or mere psychotherapy as against therapeutic psychotherapy. Neither the existence nor the potential value of verbal, procedural, and computational technologies is denied in the rule-following account. But instead of fixating on these as being what science is, it reminds us of the scientific task of innovating. It discourages the ritual use of such principles as parsimony, operationalization, and testability, but not their effective use. Thus, it puts neither obstacles nor temptations in the way of the working scientist while nevertheless giving him a rational basis on which to work and to relate his work to that of others. It is neither a daring proposal nor an empirical one to suggest that that is a good position for psychologists to be in.

## References

- Mitchell, T. O. The observational judgment of the computer simulation of persons.  
University of Colorado, Dept. of Psychology, 1967 (Mimeo).
- Ossorio, P. G. Persons. Los Angeles: Linguistic Research Institute, 1966.



