From the Descriptive to the Normative in Psychology and Logic

Paul Thagard

Philosophy of Science, Vol. 49, No. 1 (Mar., 1982), 24-42.

Stable URL:
http://links.jstor.org/sici?sici=0031-8248%28198203%2949%3A1%3C24%3AFTDTTN%3E2.0.CO%3B2-Y

Philosophy of Science is currently published by The University of Chicago Press.

FROM THE DESCRIPTIVE TO THE NORMATIVE IN PSYCHOLOGY AND LOGIC*

PAUL THAGARD†

Department of Humanities
University of Michigan-Dearborn

The aim of this paper is to describe a methodology for revising logical principles in the light of empirical psychological findings. Historical philosophy of science and wide reflective equilibrium in ethics are considered as providing possible models for arguing from the descriptive to the normative. Neither is adequate for the psychology/logic case, and a new model is constructed, employing criteria for evaluating inferential systems. Once we have such criteria, the notion of reflective equilibrium becomes redundant.

1. Introduction. Recently Stephen Stich and Richard Nisbett (1980) have made a strong case for the philosophical relevance of recent empirical work on reasoning. They argue that psychological studies concerning the pervasiveness of inferential errors in human subjects reveal serious flaws in Nelson Goodman's (1965) dissolution of the problem of induction. Goodman proposed to abandon Hume's problem of finding a foundational justification of inductive principles, and instead to pursue the task of mutually adjusting our inferential rules and our inferential practices. Stich and Nisbett borrow from Rawls (1971) the term "reflective equilibrium" to describe the state in which rules and principles have been satisfactorily adjusted to each other. Goodman's idea is that when we have reached a reflective equilibrium, the inferential rules arrived at are normatively justified. Stich and Nisbett argue that it is all too easy to get people into reflective equilibrium despite poor practices such as the gambler's fallacy or the making of regression errors. Hence reflective equilibrium alone does not justify inferential rules. They go on to consider a *social* component in justification.

This paper takes a different approach to the problem of justifying logical principles on the basis of inferential practice. I shall try to construct

*Received January 1981; revised August 1981.

†For helpful discussions, I am grateful to Richard Nisbett, Daniel Hausman, the Philosophy Department at the University of Waterloo, and Fellows of the 1981 Summer Institute on Philosophy and Psychology. I have benefited from unpublished writings on ethics by Daniel Hausman and Peter Railton. I am also grateful to the University of Michigan Program in Cognitive Science for generous support.

Philosophy of Science, 49 (1982) pp. 24–42.

a methodology for revising normative (prescriptive) logical principles in the light of descriptive psychological findings. A methodological model is sought by first considering the structure of two more familiar models for using descriptive matters in the development of normative principles: historical philosophy of science and wide reflective equilibrium in ethics. Neither model is adequate for the psychology/logic case, but components can be adapted to yield a new and richer account of how inferential practice can be a guide to what is normatively correct. It follows from this account that reflective *equilibrium* is at best incidental to the process of developing normative principles. Moreover, when the dispensability of equilibrium considerations becomes evident, Stich and Nisbett's emphasis on the social component of justification becomes avoidable, along with its potentially relativistic implications.

2. Historical Philosophy of Science. In recent years, philosophers of science have increasingly shifted from the methods of logical reconstruction pioneered by the Vienna Circle to methods involving detailed study of historical or contemporary examples of scientific practice (Laudan 1979). Instead of approaching science with the view that it needs to be cleaned up in order to reach standards commensurate with empiricist epistemology and the rigor of symbolic logic, the historical approach treats the philosophy of science as a kind of empirical discipline (Hausman 1980). Close attention must be paid to what scientists actually do, and prescriptions about what scientific method ought to amount to should be founded in actual practice. The implicit manifesto of this sort of approach was Kuhn's *Structure of Scientific Revolutions,* originally published in 1962. The point drawn by many from the historical studies of Kuhn and others was that the elegant work of logical positivists on axiomatics, reduction, deductive explanation, and formal confirmation theory was often irrelevant to science, and accordingly did not provide adequate normative standards. A few philosophers of science have followed Feyerabend (1975) in anarchistically concluding that *no* normative standards of scientific method are appropriate: anything goes. But this is not a direct consequence of the de-emphasis of logical reconstruction and the embrace of historical methods, for we can hope to generate from historical studies methodological principles whose sway can be extended normatively to cover the general practice of science. But how does this work?

Consider first the nature of theory evaluation which can provide the first rough approximation to the selection of logical principles. Recent studies of theory evaluation in science suggest that theory choice is based on methods which are *comparative, coarse-grained,* and *dynamic.* (For illustrations, see Thagard 1978, Schaffner 1970, Lakatos 1970, and Laudan 1977.) Theory choice is comparative in that it involves the assessment

of a number of competing theories with respect to the empirical evidence, not just the adequacy of a particular theory. In the terminology of Thagard (1978), theory choice is inference to the best of competing explanations. Theory choice is coarse-grained in that the units for comparing the relative explanatory strength of competing theories are not individual statements deducible from the theories but large classes of facts explained by theories, or, in the terminology of Laudan (1977), problems solved by the theories. Finally, theory choice is *dynamic* in that assessment of theories is not just a matter of the relation of theories to the evidence at a specific time, but must also take into account the way in which the research program incorporating the theory develops over time in response to the empirical challenges (*cf.* Lakatos 1970). We can expect that the assessment of normative standards in philosophy of science, ethics, and logic will also be comparative, coarse-grained, and dynamic.

But doing historical philosophy of science is in many respects different from doing empirical science. In the first place, the selection of case studies is very important. We do not typically study what Henry Snerd was doing at the Miscellaneous Technology Corporation research lab in 1934; rather, most studies concentrate on cases recognized as exemplary accomplishments of science. Galileo's physics, Newton's mechanics, Lavoisier's chemistry, and Darwin's theory of evolution have for obvious reasons been favored objects of study: they have been licensed by the subsequent history of science as genuine achievements. The progressive development of science enables us to pick out typical examples of scientific method which can be assumed, at least provisionally, to be typical not only of how science is done but of how it *should* be done. In doing empirical science, we might study a particular phenomenon because we think it to be typical of a wide range of phenomena, but there is no normative association. In historical philosophy of science, on the other hand, case studies acquire normative significance because of the background belief that such scientists as Galileo, Newton, and Darwin generally knew what they were doing.

Admittedly there are lots of pitfalls. Although we intensively examine Newton's work on mechanics and optics, we are leery of his studies of alchemy and astrology, since that work has not been vindicated by subsequent scientific developments. Studies abound on Darwin's theory of natural selection, but historians as well as philosophers have spent little time on his discredited theory of pangenesis. Hence we do not assume that the actual practice of even an esteemed scientist is always to be taken as normatively significant. Moreover, it must be granted that the philosophical historiography of science is a highly theory-laden activity. What methodology a philosopher finds in the papers and diaries of the scientists studied is undoubtedly influenced by what the philosopher's antecedent

philosophical expectations lead him or her to look for. Historical philosophy of science is unavoidably based on philosophical history of science. Philosophical history of science is not however an arbitrary enterprise, since even the philosophical historian must feel constrained by the actual statements of the subjects of investigation.

Historical philosophy of science is comparative in that we ought to consider which of different methodologies best describes and explains what is going on in a concrete case. It is coarse-grained in that we are not trying to account deductively for a general array of scientific practices, but aim to explain a restricted number of practices deemed historically significant. And it is dynamic, in that one of the criteria we should use in evaluating a philosophical account of methodology concerns how well the account leads to subsequent illuminating historical work. We may of course decide that a particular instance of scientific work is not of philosophical significance because the scientist in question was *not* employing the correct methodology. But this does not invalidate our general historical procedure, since a judgment would have to be based on reflection on other cases.

I have only scratched the surface in discussing how descriptive case studies in the history of science can be relevant to normative issues in the philosophy of science. But that is not the main concern of this paper. I want now to summarize the elements of the descriptive/normative relation in historical philosophy of science which may prove relevant for the question of the relevance of psychology to logic. The result is a schema for going from the descriptive to the normative which I shall label "HPS":

(HPS)

1. We select cases of actual scientific practice. Selection is made on the basis of subsequent events in the history of science which have marked the cases as significant contributions to the growth of scientific knowledge.
2. We develop case studies which describe scientific practice.
3. We assume—or this can be argued for—that scientists have generally been successful in achieving the epistemic goals of science.
4. Then the actual methods of the scientists in our case studies are at least an approximation to what the methods *ought* to be. We describe the methods found.
5. We reflect philosophically on the methods found in the case studies, developing more complex normative models, which can then be applied to other case studies.

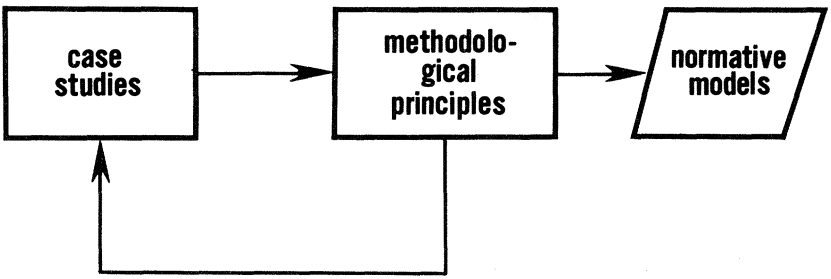

Figure One. HPS: Historical philosophy of science.

This description of HPS is schematic, and its linearity is highly misleading. It seems to suggest that we proceed by first doing history and then by deriving methodological principles. But of course our historiography is unavoidably influenced by expected methodological conclusions. History and philosophy of science should be viewed as part of a dynamic *system* of recurring influences, best represented not by a sequence of steps, but as a cybernetic process as in Figure One (*cf.* Rescher 1977). Methodological norms are arrived at only after the historical/methodological loop has been repeatedly run. Let us now compare this process with what happens in ethical theory.

3. Wide Reflective Equilibrium. John Rawls (1971) recommended an ethical method aimed at achieving "reflective equilibrium" of particular moral judgments and general moral principles. This method has been developed further by Norman Daniels (1979) on which most of the following discussion draws. The method of *wide* reflective equilibrium is described by Daniels as follows (1979, pp. 258f.):

> The method of wide reflective equilibrium is an attempt to produce coherence in an ordered triple of sets of beliefs held by a particular person, namely, (a) a set of considered moral judgments, (b) a set of moral principles, and (c) a set of relevant background theories. We begin by collecting the person's initial moral judgments and filter them to include only those of which he is relatively confident and which have been made under conditions conducive to avoiding errors of judgment. . . . We then propose alternative sets of moral principles that have varying degrees of "fit" with the moral judgments. We do *not* simply settle for the best fit of principles with judgments, however, which would give us only a *narrow* equilibrium. Instead we advance philosophical arguments intended to bring out the relative strengths and weaknesses of the alternative sets of principles (or competing moral conceptions). These arguments can be constructed as

> inferences from some set of relevant background theories (I use the term loosely). Assume that some particular set of arguments wins and that the moral agent is persuaded that some set of principles is more acceptable than others. . . . We can imagine the agent working back and forth, making adjustments to his considered judgments, his moral principles, and his background theories. In this way he arrives at an equilibrium point that consists of the ordered triple (a), (b), (c).

As with scientific theory choice and historical philosophy of science, we have here a method which is comparative, coarse-grained, and dynamic. The particular moral judgments are not at all like incorrigible intuitions to which moral principles must conform: the judgments can be revised in the light of the principles just as the principles can be revised in the light of the judgments. The revisability of particular moral judgments does not show a complete difference in kind from the empirical evidence used to evaluate scientific theories; for in the first place even observation statements in science are not incorrigible, and in the second place assessment of theories is not usually in terms of how they explain particular observations, but in terms of how they explain general classes of facts expressed in empirical generalizations which may well turn out to be false (see Thagard 1978). But it must be granted that there is at least a difference in degree between the status of the moral judgments used in establishing moral principles and the status of empirical evidence used in assessing scientific theories: moral judgments are, or ought to be, *more* corrigible. Historical philosophy of science provides an intermediate case of corrigibility. We must be prepared to admit that our interpretation of a particular case study is wrong, or even that for general reasons the case study was not an appropriate one, perhaps because the scientist's description of his procedure had been corrupted by unfortunate attention to some contemporary misguided philosopher. Nevertheless, corrigibility is limited by our antecedent conviction that work by the distinguished scientist we have chosen is representative of the best scientific work.

One of the most important aspects of Daniels' characterization of wide reflective equilibrium is the role of (c) a set of background theories. Coherence is to be achieved not only between moral judgments and principles, but also with theories concerning such matters as the nature of human beings and of society. Psychological theories about the actual or possible moral behavior of individuals will play a role in the interplay of particular moral judgments and general principles. But this is only one of the respects in which the method of wide reflective equilibrium takes us from the descriptive to the normative. The general moral principles which are developed in the move to equilibrium are clearly normative in

content, but the status of the particular judgments is ambiguous. True, a moral judgment that a particular sort of action is wrong has normative content, but it functions in the method of wide reflective equilibrium partly as a descriptive report of the attitude of a person toward the sort of action. Through a process of reflection, descriptions of how a person feels about a certain sort of action are supplanted by a judgment of how a person ought to feel about an action.

Empirical background theories can play a particularly important role in the assessment of moral principles when they concern the physical or psychological limitations of human agents. Most ethical theorists accept the principle that "ought" implies "can", so that conversely if it is empirically impossible for an agent to perform a certain action, then the moral principle that says that the agent ought to perform the action must be held very suspect. For example, the (false) psychological theory that people are capable of acting only in their own self-interest would require the rejection of an ethical theory which maintained that people have a general obligation to be altruistic. In less extreme cases, we might only have the psychological result that there are certain things which are *very hard,* but not impossible, for people to do. In this case, the "ought" implies "can" principle does not have a direct bearing, but the psychological difficulty and resultant costs of carrying out certain actions should figure in our estimation of the coherence of our judgments, principles, and background theories. This is especially clear in the context of explicitly consequentialist ethical theory, where part of the set of consequences to be taken into account in assessing the rightness of an action concerns the psychological effects on the agent.

Let us now attempt to schematize the model for reaching normative conclusions found in Daniels' description of wide reflective equilibrium. The order of steps does not represent any fixed temporal order, but is only meant to suggest a possible way of proceeding. Call the model "WRE":

(WRE)

1. We have a set of particular moral judgments about what is right or wrong, selected for expected freedom from error.
2. We postulate a number of general moral principles which explain and justify the particular judgments.
3. We attempt to come up with a maximally coherent set of beliefs, consisting not only of the moral judgments and principles, but also taking into account our background theories, especially concerning psychological limitations.
4. We arrive at a state of reflective equilibrium, and con-

clude that the acceptance of the moral principles in the final set of beliefs is justified.

As with HPS, this description is misleadingly linear: the process of reaching reflective equilibrium is better represented by the flow chart in Figure Two. Normative principles are outputs from the system only after repeated adjustments of moral judgments and principles in the light of background theories have been made. Does the procedure of assessing logical principles in the light of psychological findings at all resemble WRE?

4. HPS and WRE as Models for the Relevance of Psychology to Logic. HPS and WRE are important models for relating the descriptive and the normative in their own spheres, but neither is adequate for answering questions about the relation of the descriptive and the normative in psychology and logic. I shall attempt to show this in the present section, with a view to developing in the next a model which incorporates the aspects of HPS and WRE which are most relevant to the psychology/logic issue.

First more must be said about the nature of the problem of the relevance of psychology to logic. We are concerned with deductive, inductive, and practical inference. (Practical inference concerns what to *do*—how to make rational decisions.) These three branches of logic are clearly related to philosophy of science and ethics in some respects: one of the main problems of philosophy of science is the inductive logic of assessing scientific theories, and another concerns the foundations of statistical inference, so my discussion of HPS is not distinct from discussions of inductive inference; and ethics is often concerned with what ethical decisions to make, so it is closely related to practical inference. Hence HPS and

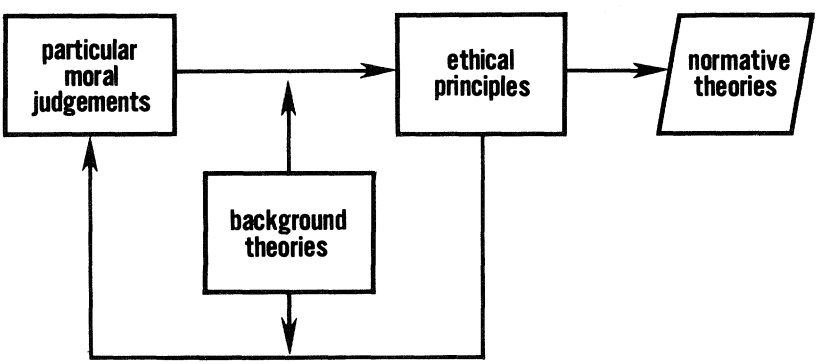

Figure Two. WRE: Wide reflective equilibrium.

WRE should not be thought of as models taken from some field alien to logic, but merely as examples of relatively explicit characterizations of how descriptive matters can have logical relevance.

Recent empirical work in psychology shows numerous systematic discrepancies between popular inferential practice and accepted logical norms (Tversky and Kahneman 1974, Kahneman and Tversky 1979b, Nisbett and Ross 1980). Such discrepancies elicit three possible responses, which can be crudely characterized as follows:

(1) People are dumb. They simply fail to follow the normatively appropriate inferential rules.
(2) Psychologists are dumb. They have failed to take into account all the variables affecting human inferences, and once all the factors are taken into account it should be possible to show that people are in fact following the appropriate rules.
(3) Logicians are dumb. They are assessing the inferential behavior of human thinkers with respect to the wrong set of normative standards.

In this discussion, I assume the egalitarian position that everybody gets to be dumb some of the time. The problem is to establish a methodology for mediating among people's inductive behavior, psychologists' descriptions of that behavior, and the logicians' normative principles used to judge the appropriateness of inferential behavior.

HPS will not furnish this methodology. The crucial difference between doing historical philosophy of science and using psychological findings to revise logical principles is the special status of the case studies in HPS. When we select Darwin or Newton for special scrutiny in HPS, we do so with the understanding that the scientific thinking of the subject is exemplary. Subsequent developments in the history of science have warranted our belief that our stellar scientists knew what they were doing, at least tacitly. When the psychologists tell us that people frequently do not take into account regression phenomena (Tversky and Kahneman 1974), we do not assume that the subjects know what they are doing: instruction in statistics can be expected to change their inferential behavior in desirable ways.

But neither can we find experts in general inferential behavior who are analogous to our exemplary scientists. We need to distinguish between two sorts of experts: those who are expert at performing a task, and those who are expert about explicitly saying how a task should be done. The two kinds of expertise need not coincide, as we see in athletics in the contrast between the inarticulate star performer and the pedagogically helpful but athletically inept coach. In HPS, we ought to pay more attention to what the subjects of our case studies do than to what they *say*

they do, since scientists' explicit methodological pronouncements are as likely to be reflections of what they think they are expected to say as they are to be reflections of their actual methodological practice. In logic, we do not have as clear a set of people certified as experts with respect to their *practice,* independent of their pronouncements.

We might suppose that for deductive inference we could take mathematicians as our practicing experts, for inductive inference we could take statisticians, and for practical inference we could take high-level managers. But in all these cases practice is severely infected by philosophical views. Intuitionistic mathematicians eschew the use of some classical logical principles. Statisticians have grave disputes about the foundations of their work, while practicing statisticians generally use a grab-bag of whatever methods—Bayesian, Neyman-Pearson, or Fisher—seem useful in a particular context. And the situation is even more indeterminate with high level managers, since to the extent they use formal principles, they probably use ones taught to them in business school by people who are more certified as experts in logic (especially practical logic) than as practicing expert managers. The situation in logic is rather like the situation in doing historical philosophy of *social* science. Whereas in natural science there is sufficient consensus in the scientific community that selection of cases for study is quite uncontentious, the schisms in the various social sciences preclude doing case studies whose validity would be universally accepted. For example, a study of explanation in economics would get very different results if the investigator concentrated on Marx's capital theory rather than neo-classical theories. In psychology, one would end up drawing a very different methodological picture as the result of attention to Gestalt psychologists' practice rather than behaviorists'. Similarly, logical practice among the alleged experts in deductive, inductive, and practical inference is not sufficiently uniform and historically validated to allow the use of the case study methodology of HPS. It can be argued that there has been progress in natural science, but the existence of progress is much more problematic in the social sciences and especially in everyday inferential practice. Without a background argument for the sort of progress that underlies our confidence in scientific case studies, we cannot use an HPS methodology in going from psychology to logic.

That is not to say that psychological studies might not be very useful in determining how people do or perhaps should reason, but unlike the situation in HPS, we will not be able to move to normative judgments primarily on the basis of the studies. Of course in historical philosophy of science we do not immediately leap from "is" to "ought" either. But there the leap is at least indirectly possible through assumptions about the nature of the growth of scientific knowledge.

Unlike HPS, WRE is similar to the psychology/logic problem in that

there are no case studies or particular moral judgments with assumed prior validity. This naturally makes the justification of ethical principles much shakier than the justification of methodological principles in the philosophy of science. We would need some argument that there is or can be progress in the development of ethical theories analogous to the more visible progress in natural scientific theories, if we are to have confidence that WRE really does serve to justify ethical principles. I cannot address this difficult issue here. But there are at least several respects in which WRE offers an attractive model for establishment of logical principles. (It must be noted in passing that Rawls (1971) cites Goodman (1965) as the inspiration for Rawls' views on reflective equilibrium, so the confluence is not accidental.)

Like particular moral judgments, common inferential practice is revisible in the light of overriding normative principles. Just as in WRE we have a dynamic of particular judgments and general moral principles, so in the psychology/logic case we can look for a process of development of inferential practice and normative logical principles as the result of critical assessment of both. In the logical case as in the ethical, we are seeking a wide rather than a narrow reflective equilibrium. This means that we want coherence of not only inferential practice and normative principles, but of both of these with background theories and beliefs.

The relevant background information is of two kinds. First, we need an account of the inferential capacities of human beings. As Goldman (1978) has suggested, the principle of "ought" implies "can" is relevant to epistemology and logic as well as ethics. We should not demand of a reasoner inferential performance which exceeds the general psychological abilities of human beings. For example, we can not prescribe that a cognizer believe all the logical consequences of his or her beliefs, since none of us has infinite storage capacity. In the same spirit, we do not want to prescribe normative logical principles which are too horribly difficult for humans to follow. What does "too difficult" mean here? To answer that, we need a second kind of background information, concerning the *goals* of the inferential behavior. With deductive and inductive inference, we have the minimal goals of achieving true beliefs and avoiding false ones, but that is a much too simple view of the matter. Other epistemic goals include achieving explanations and holistically coherent belief systems. Much of scientific knowledge, as well as most of everyday knowledge, has instrumental import, so that what deductive and inductive strategies we adopt will depend in part on our practical aims. Deductive or inductive principles which are inordinately costly in psychological terms may be supplanted by principles which *prima facie* are intellectually inferior. Especially in the logic of decision making we see the relevance of the psychological and social limitations on human cog-

nition to the question of what standards are normatively correct. As March (1978) points out, the application of apparently optimal decision strategies may not be optimal given restraints on human abilities.

The application of WRE to the psychology/logic case thus suggests that we should strive to reach reflective equilibrium among the following four factors:

(a) common inferential practice,
(b) normative logical principles,
(c) background theories about the cognitive capacities and limitations of human beings, and
(d) background views about the goals of inferential behavior.

However, this is still a too simple account of the matter, for in disputes about logical principles we often find other, more philosophical concerns brought to bear. For example, debates about the foundations of statistical inference often concern in part what philosophical interpretation of the probability calculus should be adopted: proponents and critics of Bayesianism debate the merits of subjective probabilities versus objective logical or frequentist views. Dummett (1978) argues for intuitionistic logic and mathematics largely on the basis of verificationist theories of meaning. Thus we have to add to the matrix of elements taken into account in achieving reflective equilibrium:

(e) background philosophical theories.

Cohen (1981a) asserts without argument that the psychology/logic case involves a *narrow* reflective equilibrium. He compares the case of fitting logical principles with logical practice to devising a grammar which fits a population's linguistic practice. The latter is indeed a case of narrow reflective equilibrium (Daniels 1980): we are only concerned with a fit between principles and practice, and background theories do not play a role. But in constructing a set of logical principles we are doing much more than simply matching up with actual practice. Practice can be *improved*. Logical practice has improved enormously with the developments in deductive, inductive, and practical logic of the past several hundred years. There is no analogous sense in which linguists aim to improve the overall grammar of a linguistic population; their task is descriptive. The logician, on the other hand, is concerned to develop a set of principles which is inferentially optimal given the cognitive limitations of reasoners. This requires reference to background psychological and philosophical theories and to the goals of inferential behavior. Hence logical principles could only be arrived at by a process of *wide* reflective equilibrium.

In his (1981b), Cohen gives two reasons for considering narrow rather than wide reflective equilibrium. He says that the background issues

about philosophical problems are too controversial and too finegrained to be brought to bear on experimental studies of human rationality. All that follows from this is that establishing logical principles using wide reflective equilibrium will be very difficult, since it requires at least provisional answers to hard philosophical questions which lurk in the background of disputes about logical principles. But we already know from the longevity and complexity of disputes about inductive and deductive principles that the establishment of such principles is never easy. We can not evade philosophical issues about meaning, ontology, and inferential goals. Consideration of these issues as well as background psychological theories requires wide reflective equilibrium, rather than narrow.

In a situation where narrow reflective equilibrium is sought, the descriptive inputs are relatively incorrigible. The linguist is concerned to report grammatical practice, not to reform it. Similarly, in historical philosophy of science, our concern is not usually to say what the exemplary scientists should have done. The extreme case of relative incorrigibility is scientific theory choice, where, despite the theory ladenness of observations, it is unusual to completely override an observation. In contrast, in ethics, where we seek wide reflective equilibrium, we expect background knowledge and the improvement of principles to lead to revision and improvement in intuitive ethical judgments. Similarly in logic and psychology, the descriptive input is subject to change as education brings about improvement in inferential practice. Cohen's assumption that basic inferential practice *must* be rational is as insupportable as intuitionism in ethics.

5. From Psychology to Logic. We can now give a rough characterization of what the above discussion suggests is the appropriate methodology for assessing the relevance of psychological studies to normative principles. The following model, called "FPL" for "from psychology to logic" suggests a possible procedure for attempting to resolve disputes about normative principles. As with HPS and WRE, the steps have no rigid temporal significance.

(FPL)

1. We do empirical studies to describe inferential behavior.
2. We generate sets of logical principles which explain and justify that inferential behavior.
3. When inferential behavior deviates from logical norms, we consider whether new norms are needed or whether we can just revise inferential behavior to bring it in line with existing norms.
4. This consideration is based on an attempt to develop

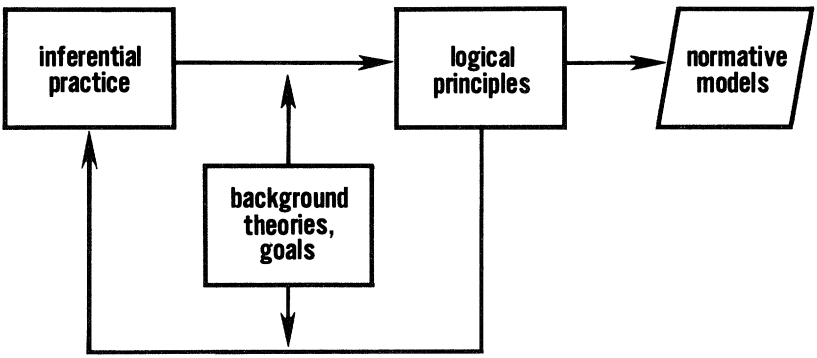

Figure Three. FPL: From psychology to logic.

a maximally coherent set of beliefs about peoples' actual behavior, their optimal behavior given their cognitive limitations and the goals of inferential behavior, and background philosophical issues.

5. The logical principles among the maximally coherent set of beliefs are then deemed to be justified.

Once again we have a process whose dynamic features are best represented in a flow chart, as in Figure Three.

By analogy with WRE, we repeatedly adjust logical practice and principles in the light of background theories until reflective equilibrium is reached; then the principles arrived at in the equilibrium state are "outputted" as normatively correct. But we cannot escape asking: what is it which determines that we have reached a state of reflective equilibrium? In terms of the linear description of FPL, we must ask what it is to reach step 4 and achieve a "maximally coherent" set of beliefs. Without an account of how to evaluate coherence among practice, principles, goals, and background theories, FPL has little content. I shall now describe a list of criteria for assessing the coherence of such a set of beliefs. Then in Section 6 I shall argue that the existence of such a set renders any discussion of reflective equilibrium redundant.

An *inferential system* is a matrix of four elements: normative principles, descriptions of inferential practice, inferential goals, background psychological and philosophical theories. Schematically a system $S = \langle NP, IP, G, T \rangle$. How do we assess the coherence of such a system? Most importantly, how can we say that one system is more coherent than another? I propose three main criteria: robustness, accomodation, and efficacy. A system is robust if its normative principles account for inferential practice in a wide range of situations. Robustness in inferential

systems is analogous to consilience in theory choice, where a theory is consilient if it explains a wide range of facts (Thagard 1978). In a robust system, our normative principles justify and explain a variety of inferential behavior. However, we do not expect the principles to account for *all* inferential behavior, since we need to leave open the possibility that even after considerable reflection people's behavior still deviates from logical norms. We can *accommodate* such behavior by using background psychological theory to explain why in some cases people deviate from the system's logical norms. Thus accommodation provides a criterion supplemental to robustness: to the extent that principles in S do not robustly account for inferential practice, we expect to be able to explain that practice by reference to psychological factors which interfere with the application of logical principles; the deviant practice is thereby accommodated.

By *efficacy* of a system I mean the extent to which the principles and practices of a system lead to satisfaction of the relevant inferential goals. This is in part an empirical matter. We must observe how well principles enable us to satisfy such goals as preserving truth in deduction, achieving explanatory theories in induction, and meeting human needs in practical inference. Efficacy also should take into account what background psychological theories tell us about how easily principles will be applicable given human information processing mechanisms.

Schematically, we can summarize these three criteria by the following questions:

(1) Robustness: to what extent do the normative principles account for inductive practice?
(2) Accommodation: to what extent do background theories account for deviations of inductive practice from the normative principles?
(3) Efficacy: given background theories, to what extent does following the normative principles promote the satisfaction of the inferential goals?

These criteria can be used to assess the comparative coherence of competing systems. In a given domain, we can assume that T and G will be common to competing systems, and this gives us some hope of reaching an objective conclusion that one system is more coherent than the other. In particular, comparison of the *efficacy* of the two systems may enable us to make choice of systems more than a matter of purely internal coherence. Choice will obviously be highly complex and non-algorithmic, but nevertheless may be determinate and objective. The core, then, of the method FPL is development of an inferential system which is highly coherent according to the above criteria.

6. Beyond Reflective Equilibrium. But now we can abandon the presumption that the achievement of reflective equilibrium is somehow essential to the justification of logical norms. The notion of reflective equilibrium has been very useful in getting us to our current model FPL, since it has enabled us to elucidate the Goodmanian process of mutually adjusting principles and practice, while also taking into account background theories. But we can quickly see that the notion of equilibrium is of no help in actually justifying principles.

The problem, as Stich and Nisbett (1980) pointed out, is just that equilibrium may be too easily reached. An individual may reach reflective equilibrium while possessing an inferential system which is resoundingly non-efficacious: people can rest contentedly with the gambler's and other fallacies. Stich and Nisbett's move to discussion of *expert* reflective equilibrium is of no help, since they recognize that the experts can achieve specious equilibria too. At the end of their paper, they seem reduced to the relativist conclusion that rational resolution of debates between conservative experts and "cognitive rebels" may be impossible.

This conclusion is avoided by seeing that the justification of a set of normative principles is based, not on the reflective equilibrium of any individual or group, but on the place of the principles in a defensible inferential system. Defense is based on arguments that the system is coherent according to the criteria discussed above. General and expert inferential behaviors are relevant, since their description is part of *IP* which is one component of the system *S*. But other factors besides *IP* play a role in determining which *S* and set of normative principles *NP* will be optimal. The criteria of accommodation brings out the fact that we may well expect to have general or expert inferential behavior which does *not* conform to *NP*.

If one takes the notion of reflective equilibrium too seriously, one is pushed to unsatisfactory answers to the question: *whose* reflective equilibrium? There are two possible answers, one populist and one elitist. The populist strategy, favored by Cohen (1981), is to emphasize the reflective equilibrium of the average person. This founders, because education in sophisticated inferential techniques can be expected to provide the individual with a much more efficacious system. The elitist strategy, favored by Stich and Nisbett (1980), is to emphasize the reflective equilibrium of experts. This too is inadequate, for it leaves us no way of saying why the experts *should* be in equilibrium, or of mediating disputes among experts. On my account, the experts—or for that matter ordinary persons—*ought* to have their principles and practices in equilibrium *if* they have a highly coherent inferential system. Coherence is to be evaluated according to criteria to which the achievement of reflective equilibrium is *irrelevant*. What we are really after is not equilibrium, but

progress: the development of better and better inferential systems. Improvement of inductive systems may well come about through an oscillating process of richer and more efficacious principles, practices, theories, and goals even if we never during the process achieve equilibrium.

Compare the situation in scientific theory choice. One can imagine the claim that a theory becomes acceptable when one is able to reach a "reflective equilibrium" between the theory and the observations it is supposed to explain. But that would clearly obscure the nature of the justificatory process. What makes the theory acceptable is that it provides a better explanation of the observations than competing theories, according to criteria such as explanatory breadth, simplicity, and analogy (Thagard 1978). Reflective equilibrium would be in this case a mere epiphenomenon of the justified acceptance of the theory. Similarly, once we can provide criteria for assessing logical norms *vis-à-vis* competing inferential systems, then reflective equilibrium is seen not to be an essential feature of the justification of those norms. At best, the notion of reflective equilibrium provides a metaphor for describing the complex process, better represented in Figure 3, of how the development of a justified system of logical norms involves the interaction of numerous components including descriptions of inductive practice.

I have argued that reflective equilibrium—Goodman's fit between inferential practice and normative principles—is not in itself a source of justification of the principles. It may however provide an indirect way of telling what principles are justified at a particular time. The point here is social, not epistemological: the fact that the experts or others are in reflective equilibrium does not justify anything; for this we need arguments that the inferential system in question is optimal with respect to the criteria discussed. However, when we face the practical problem of deciding what are most likely to be the best logical principles to use, we would be wise to follow the advice of Stich and Nisbett and consult the inferential experts. Since they are more familiar with alternative inferential practices, background theories, and inferential goals than are ordinary people, experts are more likely to have highly coherent inference systems.

The metamethodological question naturally arises: what legitimates the three proposed criteria for inferential coherence? Here we have a development from the descriptive to the normative at a higher level: I have advocated these criteria because they seem to be the ones actually used when we set out to evaluate inferential practices, and because they seem to promote the establishment of the sorts of inferential principles we want. In short, the criteria are, in an extended sense, robust and efficacious. There is no circularity here of the sort that is found in inductive justifications of induction. We have long since abandoned the search for a full

foundationalist justification of inferential practice. (*cf.* Rescher 1977.)

It would have been helpful to illustrate the application of FPL and the three criteria for coherence through detailed discussion of actual cases of debate over inferential principles, but that requires separate articles. I have elsewhere suggested ways in which psychological considerations may require revision of standard canons of the logic of decision (Thagard 1981). In addition, applications of empirical studies to the logic of confirmation of generalizations can be found in Thagard and Nisbett (forthcoming). Other disputes to which the methodology might be applied include critiques of Bayesian foundations of statistics (Shafer 1976).

These disputes can not be rationally settled if we are solely concerned with the reflective equilibria of experts or people in general. Settlement requires assessment of competing systems, dynamically developed through a dialectic of principles and practices against a backdrop of goals and theories. The model FPL and the criteria of robustness, accommodation, and efficacy provide a way of understanding how we can objectively apply the descriptive to the normative in psychology and logic.

REFERENCES

Braine, M. (1978), "On the Relation Between the Natural Logic of Reasoning and Standard Logic", *Psychological Review 85*: 395–416.

Cohen, L. J. (1977), *The Probable and the Provable*. Oxford: Clarendon Press.

Cohen, L. J. (1980), "Where is the Fallacy? A Rejoinder to Daniel Kahneman and Amos Tversky", *Cognition 8*: 89–92.

Cohen, L. J. (1981a), "Can Human Irrationality be Experimentally Demonstrated?" *Behavioral and Brain Sciences 4*: 317–331.

Cohen, L. J. (1981b), "Reply to Commentators: Are There Any A Priori Constraints on the Study of Rationality?", *Behavioral and Brain Sciences 4*: 359–367.

Daniels, N. (1979), "Wide Reflective Equilibrium and Theory Acceptance in Ethics", *Journal of Philosophy 76*: 256–282.

Daniels, N. (1980), "On Some Methods of Ethics and Linguistics", *Philosophical Studies 37*: 21–36.

Dummett, M. (1978), *Truth and Other Enigmas*. Cambridge: Harvard University Press.

Feyerabend, P. (1975), *Against Method*. London: New Left Books.

Goldman, A. I. (1978), "Epistemics: The Regulative Theory of Cognition", *Journal of Philosophy 75*: 509–523.

Goodman, N. (1965), *Fact, Fiction, and Forecast*. Second edition. Indianapolis: Bobbs-Merrill.

Hausman, D. (1980), "How to Do Philosophy of Economics", in P. Asquith and R. Giere, eds., *PSA 1980*, vol. 1. East Lansing: Philosophy of Science Association: 353–362.

Henle, M. (1962), "On the Relation Between Logic and Thinking", *Psychological Review 69*: 366–378.

Kahneman, D. and Tversky, A. (1979a), "On the Interpretation of Intuitive Probability: A Reply to Jonathan Cohen", *Cognition 7*: 409–411.

Kahneman, D. and Tversky, A. (1979b), "Prospect Theory: An Analysis of Decision Under Risk", *Econometrica 47*: 263–291.

Kuhn, T. (1970), *The Structure of Scientific Revolutions*. Second edition. Chicago: University of Chicago Press.

Lakatos, I. (1970), "Falsification and the Methodology of Scientific Research Programs",

in I. Lakatos and A. Musgrave, eds., *Criticism and the Growth of Knowledge*. Cambridge: Cambridge University Press: 91–195.
Laudan, L. (1977), *Progress and Its Problems*. Berkeley: University of California Press.
Laudan, L. (1979), "Historical Methodologies: An Overview and Manifesto", in P. Asquith and H. Kyburg, eds., *Current Research in Philosophy of Science*. East Lansing: Philosophy of Science Association: 40–54.
March, J. G. (1978), "Bounded Rationality, Ambiguity, and the Engineering of Choice", *Bell Journal of Economics 9*: 587–608.
Nisbett, R. and Ross, L. (1980), *Human Inference: Strategies and Shortcomings of Social Judgment*. Englewood Cliffs: Prentice-Hall.
Rawls, J. (1971), *A Theory of Justice*. Cambridge: Harvard University Press.
Rescher, N. (1977), *Methodological Pragmatism*. New York: New York University Press.
Schaffner, K. (1970), "Outline of a Logic of Comparative Theory Evaluation with Special Attention to Pre- and Post-Relativistic Electrodynamics", in *Minnesota Studies in the Philosophy of Science V*: 311–354. Minneapolis: University of Minnesota Press.
Shafer, G. (1976), *A Mathematical Theory of Evidence*. Princeton: Princeton University Press.
Slovik, P., Fischhoff, B. and Lichtenstein, S. (1977), "Behavorial Decision Theory", *Annual Review of Psychology 28*: 1–39.
Sober, E. (1978), "Psychologism", *Journal for the Theory of Social Behavior 8*: 165–191.
Stich, S. and Nisbett, R. (1980), "Justification and the Psychology of Human Reasoning", *Philosophy of Science 47*: 188–202.
Thagard, P. (1978), "The Best Explanation: Criteria for Theory Choice", *Journal of Philosophy 75*: 76–92.
Thagard, P. (1981), "Beyond Utility Theory", in M. Bradie *et al.*, eds., *Reason and Decision*. Bowling Green, Ohio: Bowling Green State University Department of Philosophy.
Thagard, P. and Nisbett, R. E. (forthcoming), "Variability and Confirmation", *Philosophical Studies*.
Tversky, A. and Kahneman, D. (1974), "Judgment Under Uncertainty: Heuristics and Biases", *Science 185*: 1124–1131.
Wason, P. and Johnson-Laird, P. (1972), *Psychology of Reasoning*. Cambridge: Harvard University Press.