SCIENCE AS PRACTICE: A METHODOLOGICAL CRITIQUE AND CASE STUDY

by

Ranjan Chaudhuri

Thesis submitted to the faculty of the
Virginia Polytechnic Institute and State University
in partial fulfilment of the requirements for the degree of

MASTER OF SCIENCE

in

Science and Technology Studies

APPROVED:

Steve Fuller

Steve Fuller, Chair

Peter Barker

Peter Barker

Ellsworth Fuhrman

Ellsworth Fuhrman

May 1992

Blacksburg, Virginia

SCIENCE AS PRACTICE: A METHODOLOGICAL CRITIQUE AND CASE STUDY

by
Ranjan Chaudhuri

Chair: Steve W. Fuller
Science and Technology Studies

(ABSTRACT)

In this thesis, I attempt to develop the rudiments of a 'practice' conception of scientific knowledge and activities as the basis for a suitable methodology for Science and Technology Studies. In order to do this, I examine the methodologies proposed by two sociologists who can be very broadly construed to be working within the tradition of the 'Sociology of Scientific Knowledge', Harry Collins and Steve Woolgar in the context of their application to a specific case, and attempt to develop an alternative conception by contrast.

The thesis is structured as follows. I begin by describing Collins' and Woolgar's methodologies for the analysis of scientific knowledge development in some detail in Chapter One. In Chapter Two, I examine the application of these methodologies to the 'computer models of scientific discovery' case (the case is of some interest because it is held to 'refute' the possibility of the sort of analysis of scientific activities that Collins and Woolgar propose). I then use the material of the first two Chapters in Chapter Three to diagnose the shortcomings of Collins' and Woolgar's methodologies as illustrated by their application to the computer models of discovery case. This sets the stage for an alternative analysis of the computer models case which does not suffer from these shortcomings in Chapter Four. Finally, I develop a practice-based conception of scientific knowledge development in Chapter Five (which I derive from the material of Chapters One through Four), contrast it with the methodologies of Collins and Woolgar, and use it to illustrate and evaluate my alternative analysis.

Acknowledgements

The C.S.S.S. and STS program at Virginia Tech has proven to be a wonderful intellectual environment, and I have greatly benefitted from my teachers and friends of the past two years, all of whom I obviously cannot thank in laundry-list fashion here. In connection with this thesis, I would however like to thank the following people in particular for their help and encouragement. The members of my committee, Peter Barker, Skip Fuhrman and Steve Fuller patiently read and corrected earlier drafts. I would especially like to thank Steve for initially suggesting the topic, and for advice along the way which I have incorporated but cannot acknowledge in each instance due to its sheer volume; Peter for the opportunity afforded by his course 'Science as Social Practice' in the Fall of 1992 which helped get me started and more; and Skip for insisting that I be reflexive both in letter and in spirit. My fellow students helped provide a happy working environment, and support in forms ranging from a sympathetic ear to constructive criticism to occasional nutritional supplements. I would especially like to thank Doug, Sujatha and Teresa here. Carolyn helped me with word-processing problems and allowed me the use of her equipment at not inconsiderable inconvenience to herself, for which I am truly grateful. My housemates Curt, Garrit, Gonzo and Numbers accommodated my erratic schedule and put up with some of my darker moods with humor and understanding. My brother Anjan and sister Debrani were a cheering presence from afar. Without the love and support of my parents, none of this would have been possible, and it is to them that this thesis is dedicated.

Table of Contents

Introduction

The primary motivation behind this thesis is to engage the issue of what Science Studies or STS is about. With the demise of foundationalist attempts to demarcate (or justify) 'Science' in terms of either aims or methods or subject-matter, the trend in Science Studies work of late has been toward a heterogenous collection of investigations of the relationships between scientific activities and other activities within the society. I wish to address the issue of what generally (if anything) STSers ought to be doing in the course of their studies of current developments in science and technology. My overall aim is a critical and reflective one, and I ultimately wish to come up with normative and exhortive proposals for STSers to follow as a result; this thesis is intended to be a beginning to that end.

Specifically, the strategy adopted in the thesis is the following. I eventually attempt to conceptualize both scientific activity and that of the STSer studying it as 'practices', which I minimally define as collective rule governed activities. I believe that doing so allows for the creation of a framework within which to answer questions such as 'What, if anything, is the role of the STS practitioner in examining any current scientific activities?', 'Are there any goals that STSers share?' and 'What methods might be appropriate to the attainment of these ends?'. An adequate conception of STS practice would then (in the course of drawing attention to the nature and point of the activity which STSers ought to engage in) provide some minimal prescriptive and evaluative standards to guide and judge Science Studies research.

I proceed to identify (some of) the relevant 'meta-issues' along the above lines surrounding work in Science Studies by examining the positions of Harry Collins and Steve Woolgar in the context of their analysis of a particular case in the first three Chapters of the thesis. The purpose behind laying all this groundwork is twofold. Examining the methodologies of Collins and Woolgar allows me to analyze their implicit conceptions of STS practice, so as to formulate criticisms of their conceptions and develop an alternative by contrast. Examining the applications of these methodologies to a particular case then allows me to reflexively illustrate the shortcomings of their analyses and explicitly show how an adequate alternative analysis would differ from theirs. I have chosen Collins and Woolgar not arbitrarily, but because their methodologies seem to have some prima facie drawbacks that are especially antithetical to the practice conception I wish to develop. These become apparent in the detailed analysis of their methodologies and their application to the case study which I give in the first three Chapters. Without preempting what follows, I believe that both Collins and Woolgar take as the starting point for their methodological constructions the tenability of epistemological relativism as grounds for adopting an agnostic stance to 'internal' questions in their case studies of scientific activities. I find this approach wanting on various grounds on the conception of STS practice that I finally set out in the last Chapter.

The case that I have selected is itself of some interest in this regard. It involves the claims of Peter Slezak, a cognitive scientist who believes that recent Cognitive Science (and particularly Artificial Intelligence) research serves to 'empirically refute' all STS approaches to 'Science' (which I take to be all approaches to understanding scientific activity which insist on their ineliminably 'social' character). Slezak argues that the case of 'computer modelling of scientific discoveries' taken from within AI research, in conjunction with other developments in cognitive science serve to show that scientific discovery essentially involves no 'social' factors.

Cognitive Science and AI, (particularly in Slezak's rendition of them) is a good case for my purposes for the following reasons. In the first place, Cognitive Science as a 'scientific' discipline is worthy of consideration by the STS analyst in its own right. But as a 'field' it also has some other characteristics that make it interesting from an STS perspective. The case of Cognitive Science provides a locus for the discussion of several issues of interest to STSers, methodological as well as substantive, that make investigating it a good way to begin developing an adequate conception of STS practice.

Cognitive Science is a loose conglomeration of researchers drawn from relatively well-established disciplines, but is itself arguably not a discipline with a well established perceived identity. Conceptualizing 'cognitive science' research becomes an interesting issue for the STS analyst as a consequence. Secondly, cognitive scientists are interested in, and believe that they have the resources to address some of the 'same' questions that STSers are interested in- for example that of the relationship between 'cognitive' and 'social' factors in scientific inquiry, for Cognitive Science is often construed at least in part as a reflexive inquiry into the nature of science itself. The STS analyst can therefore consider the import of cognitive science research for such questions (and also consider the significance of the 'empirical' research results for STS practice) in the course of addressing other questions about cognitive science as a 'scientific activity' in its own right (such as the relationship between the 'internal' and 'external' views of cognitive science as a scientific activity, the competence required of STS practitioners in order to assess and critique the scientific activities they examine, and the dimensions along which normative recommendations might be made as a consequence of the STSer's analysis).

Slezak's presentation of Cognitive Science and AI is especially pertinent to answering questions like the ones above because it can be taken as an instance of a 'scientific practice' being used to directly challenge the fundamental tenets of (a version of) STS practice, and hence throws issues like the above ones into sharp relief: since Collins and Woolgar are compelled to address these issues in their responses to Slezak, this allows me to engage in a reflexive examination of sorts in the final two Chapters, where I set out what I take to be the salient 'meta-issues' involved in conceptualizing STS as a practice by studying how Collins and Woolgar approach Cognitive Science practice. A critique of their methodologies and applications in the case of cognitive science therefore allows me to formulate, in the final Chapter, methodological directives for STS practitioners as well as some suggestions for conceptualizing their objects of inquiry, which is the intended payoff of engaging in the exercise.

Chapter One: The Methodologies of Collins and Woolgar

In this Chapter I set out the methodologies that Collins and Woolgar propose for the analysis of scientific knowledge-development. I then go on to examine their application to a specific case in the following Chapter.

(i) Collins' methodology: The 'Empirical Programme of Relativism'

Collins' methodology derives from his studies of the performance and replication of novel scientific experiments[1] within the 'Empirical Program of Relativism' (or EPOR).[2]

Collins (1981a) takes EPOR based analyses of scientific change to proceed in three stages. In the first stage the EPOR researcher (henceforth 'researcher') engages in participant-observation of scientific experiments meant to produce novel results, during which she collects 'data' for the purpose of understanding the process of scientific-knowledge development. Several case studies of ongoing experimental research are carried out to study the processes by which such questions as what the experiment is supposed to demonstrate, what counts as a correctly done experiment and what counts as the correct result are settled.

The second stage of EPOR involves formulating and verifying various 'hypotheses'[3] about

---

[1]. Collins has concentrated on experimental work because of his belief that replicability of experimental results is what grounds the widespread acceptance of corresponding theory change: "From the scientist's point of view, the establishment of a conceptual change amounts to the widespread acceptance that the corresponding empirical results are replicable" (Collins (1985), p. 129). Whether this is so or not, I'll assume that the methodology developed here is to be equally applicable to the study of all scientific change, whether conceptual or empirical.

[2]. Collins (1981a) provides an overview of the original programme and assembles together a number of case studies constructed under its aegis.

[3]. The scare quotes around 'data' and 'hypotheses' are to draw attention to the fact that the sorts of data and hypotheses (as well as the identified processes) that Collins has in mind are not necessarily supposed to be of any 'standard' scientific sort.

the actual processes by which the above issues are settled (i.e. the actual process by which new scientific knowledge is produced). The third and final stage of EPOR requires the researcher to link up (in the specified way) the processes identified in the second stage to the wider social context in which the experiments are performed, thereby providing a complete explanation of the nature of the processes involved.

In subsequent work, Collins has worked out the EPOR methodology and its rationale in greater detail. Collins (1981b) calls this approach to the study of scientific-knowledge production the 'Radical Program' (RP) to distance it from the 'Strong Programme' advocated in Bloor (1976) and the approaches of the 'rationalist philosophers of science' (collectively referred to as the Normal Programme, or NP). Collins (1981b) (a) argues that a stance of epistemological relativism towards natural scientific claims is a methodological imperative of RP (b) declares the SP tenets of 'impartiality' and 'symmetry' to be 'meta-methodological' commitments of RP, and (c) rejects the SP tenets of 'causality' and 'reflexivity'.

(a) Epistemological relativism as a methodological imperative emerges from Collins' critique of NP approaches to explaining the emergence of new scientific knowledge, i.e. those explanations that involve recourse to what is 'TRASP' (true, rational, successful or progressive). Collins divides NP explanations into two types: rational-actor explanations (which take the explicit appeal to what is TRASP by the scientists to suffice in explaining the new developments); and hidden-hand explanations (which take abstract 'principles of TRASPness' to suffice, whether the scientists are themselves aware of these or not).

In the rational-actor case, Collins holds that the correctness of a new scientific claim (or experimental result) cannot be (entirely) the result of currently accepted 'principles of rationality' precisely because of its novelty, and that post-hoc rational-actor explanations cannot (completely) account for the process of knowledge creation as it actually occurs because they would involve an illegitimate appeal to the *outcome* of the process to explain it. Since appeals to what is TRASP do not suffice to explain the process of scientific knowledge creation, the researcher is to avoid *using* (only) these notions in her explanations of the process, and is hence to take an agnostic stance toward all TRASP claims made by the scientists performing the experiment.

This methodological stance is taken to be provided with a measure of empirical support[4] from the case studies carried out in the first stage of EPOR. These case studies are taken to establish certain theses about the generation of new scientific knowledge by experiments which count against rational-actor TRASP explanations: for example, the participation of the researcher in the process reveals the skill-like character of experimental ability, and the impossibility of its complete and formal explication (Collins (1985), p. 129).[5] This in turn leads to the possibility of a regress in the explanation of what an experiment establishes and how it is correctly performed when disagreements arise.[6] In the case of disagreements, therefore, there is a certain 'interpretative flexibility' involved in the answers to the above questions, which allows for an investigation of the processes by which the flexibility is 'managed' and disagreements are settled. These case studies go some way toward establishing that these

[4]. This suggests that the methodological imperative prescribing a relativist stance is empirically vindicated. It seems that the case studies are supposed to vindicate the methodological imperative, and at the same time the methodological imperative is supposed to guide the case studies.

[5]. *Participant*-observation is required of the researcher so that the skill-like nature of the experimenter's ability is experienced first-hand. Also, by participating, the researcher is in a better position to demonstrate that TRASP-based explanations are indeed inadequate accounts of the process, and that a more complete explanation is required.

[6]. In Collins' words, the 'experimenter's regress' arises because "the skill-like nature of experimentation means that the competence of experimenters and the integrity of experiments can only be ascertained by examining *results*, but the appropriate results can only be known from competently performed experiments, and so forth" (Collins (1985), p. 130).

processes involve factors other than the principles of rationality explicitly invoked by the disputants to settle the issue in their own favor.

Hidden-hand TRASP explanations are summarily dismissed by Collins because

> [they] do not seem to be in the temper of the times since they rest on the sort of confidence in the stability of) our contemporary understanding of the natural world that is largely lacking in current philosophy. (Collins (1981b), p. 222).

(b) Symmetry and impartiality are taken to be meta-methodological commitments of RP as they lead to the methodological commitments independently argued for and empirically established as described above (e.g. the commitment to epistemological relativism with respect to natural scientific claims, participant-observation of the process of knowledge development and so forth).[7] Impartiality is compatible with the relativistic stance because during the actual experimental process one does not know whether the putative new results or claims will be accepted as true or rejected as false; symmetry is likewise compatible because the explanation of the process must be of sufficient generality (i.e. each particular explanation must be an instance of the same type) in order to be applicable to the outcome in either case. In the case studies carried out in the first stage of EPOR, therefore, the researcher treats the claims of the various scientists involved impartially. (And the overall explanation produced after the second and third stages would be symmetrical).

(c) The other two SP tenets, reflexivity and causality, are rejected. Reflexivity is rejected because that would require the researcher to take a stance of epistemological relativism toward her own explanations of the scientific-knowledge development process. Collins holds that on

---

7. These are called meta-methodological commitments because they are neither empirical nor a priori epistemological claims. They are presuppositions that put one in the "appropriate frame of mind for doing the sociology of knowledge because it leads to the right kind of methodology" (Collins (1985), p. 174). These presuppositions are perhaps best conceived as the 'hard-core' of EPOR in Lakatosian terms.

the contrary, the researcher should assume that her own explanations of the process provide 'objective' and 'real' knowledge.[8] The rationale is that there is no reason for the researcher to believe that the explanations deriving from EPOR will be compromised by the relativistic stance toward the 'natural world' since the explanations produced by the researcher need not be of the same kind (as they do not apply to the same domain) as those produced by the natural scientist. The rejection of causality is related to this point. If the theories produced in EPOR need not be of the same sort as natural scientific ones, then the forms of explanation offered need not be 'causal' ones in any strict sense either.[9] Thus, the case studies produced in the first stage of EPOR produce 'data' and lead to the development of explanations in later stages that are not necessarily expressible in the form of a causal law.

These considerations are used by Collins in developing the theories and 'patterns of explanation' to be employed in the second and third stages of EPOR (the details are to be found in Collins 1985), especially Chs. 1 and 6).[10] Recall that the case studies done in the first stage established the 'experimenter's regress' and the 'interpretative flexibility' involved in construing experimental results in the case of disagreements. The second stage involves the identification of the mechanisms for constraining this flexibility and the third stage involves explaining how the regress is avoided in general by invoking the 'wider social context' as an explanatory resource. In completing the EPOR framework for explaining how new scientific knowledge is

[8]. Collins continues to endorse this sort of 'social realism', see Collins and Yearley (1992).

[9]. This seems to suggest that EPOR does not produce any naturalistic or 'scientific' theories despite its empirical character. Collins does seem to suggest otherwise elsewhere (see Collins (1985), p. 172); perhaps he wants to retain the honorific label 'scientific' for EPOR explanations, despite their difference from the paradigmatic causal explanations of e.g. physics.

[10]. The 'Wittgensteinian' views developed here are very similar to those in Bloor (1983) except that Collins seems to reject the call for systematic, causal analysis of what Bloor calls 'language games' and he calls 'forms of life'.

actually produced, Collins uses the notions of a 'network model of society' and 'core sets' in these second and third stages.

I now describe Collins' version of the 'network model', and the idea of the 'core set' to complete my exposition of the EPOR framework. On the network model[11], any society supports a 'network of concepts'. The links between these concepts that produce the network are the generalizations that are made using them within the society. (Collins gives the example of the concepts of 'emerald' and 'green' being linked together in their being used to make the generalization that 'Emeralds are green'). Collins holds, following Winch (1958), that what sustains the links between these concepts (i.e. what supports these generalizations) are "rules embodied and institutionalized in forms of life" (Collins (1985), p. 132).

Each generalization that links concepts is used by groups ('institutions') in specific ways (on the basis of the relevant rules) in the course of their activities. The overall network of concepts of a society therefore corresponds to the activities of the various groups of people using them. And conversely, the activities of these groups of people overlap in various ways, corresponding to the links between the various concepts expressed in the generalizations they make. Concepts and the links between them are held on this model to be created and revisable as the activities of their human users and the rules for their use may change (or novel kinds of activities may come about). Given the linked nature of concepts, therefore a change in any concept may have ramifications for all the other existing concepts to which it is linked and the generalizations made using them in the course of the activities of their users, and

> the extent to which a change... is likely to reverberate through the system as a whole affects the ease with

[11]. Collins traces the ancestry of this model to Hesse (1974) and Barnes (1983). The resemblance to Bloor (1983) has been indicated above.

which that change is brought about. (Collins (1985), p. 133).

'Scientific' concepts and the activities that utilize them are no different in these respects. Thus scientific concepts also form a network, and the concepts in this network are linked to various (other, that is 'nonscientific') concepts in use in the wider society as well. The development of new scientific knowledge claims involves the creation of new scientific concepts or changing the use of preexisting ones (and the corresponding changes in the activities of the groups of people utilizing them). The existing scientific concepts making up the 'scientific conceptual network' are taken to be grounded in the institutionalized rules for their use within the groups that make up the 'scientific community'. These institutionalized rules are taken to be social conventions, or procedures for use defined and accepted within the relevant groups of users. The interpretative flexibility that arises when new claims are being established (as demonstrated in the first stage of EPOR) is then unsurprising because of the initial lack of accepted procedures for use in the case of new concepts or concept revisions. The process that constrains this flexibility can then be understood as the process by which acceptable procedures of use are created (or 'negotiated').

Given all this, Collins sees two (complementary) sets of changes that need to be dealt with in the second and third stages of EPOR in order to give a complete explanation of the development of scientific knowledge- the changes in concepts that (may) occur when for example a new experiment is performed, and the corresponding changes in the institutionalized use of those concepts by the various groups using it. Both changes in the 'conceptual order' as well as changes in the 'social order' underlying it would have to be explained.[12] The second

[12]. From what has been said so far, it might seem that the conceptual and social orders were 'isomorphic' for Collins, so that an explanation in conceptual terms would be *equivalent* to the corresponding explanation in social or institutional terms. But the explanation in terms of the social order is clearly supposed to *add* to the explanation in terms

stage of EPOR deals largely with understanding the changes in the conceptual order while the third stage explicitly connects these to the larger social structures or institutions within which it is embedded. The tool that Collins uses to articulate and link together the patterns of explanation developed in these second and third stages is the notion of the 'core-set'.[13]

As noted above, on Collins' view all potentially new scientific claims involve the revision of existing concepts or the creation of new ones. As these concepts are linked to others in generalizations that have various institutionalized patterns of use within the scientific conceptual network, all those whose own concepts are linked to the ones being revised become involved in the concept-revision process (and the greater the perceived potential impact on their own concepts and generalizations, the deeper the involvement). The 'core-set' is thus the set of people that become *directly* involved in the process of negotiating the revision of the use of a particular concept. The members of the core-set involved in negotiating the revision of the use of a particular concept are a heterogenous group that come together in contingent circumstances, each having

> greatest knowledge of those parts of the conceptual web that make up their own discipline... [thus their]... arguments and attitudes... will be affected by... their background and training... their perceptions of their place in the web and their ambitions and strategies. (Collins 1985), p. 142).

The second stage of EPOR involves identifying the mechanisms by which the process of negotiation is actually carried out in core-set negotiations. In the case of new experiments, these mechanisms do include negotiations over such items drawn from the existing and accepted

---

of the conceptual order, for talking about social institutions allows for a structural mechanism for explaining change not available to explanations given solely in conceptual terms since the concepts are said to be grounded in these various institutions or groups.

13. This notion is initially developed in Collins (1979).

'scientific' repertoire of 'tests of tests' available to the various members of the core set, such as "calibration of instruments" and "use of surrogate phenomena" (Collins (1985), Chs. 4 and 5). But the availability of the tests themselves do not suffice to determine how they will be used or which ones will turn out to be convincing and to whom, for given the results of the first stage, it is clear that the 'scientific' considerations alone cannot suffice to block the potential regress (see above). There are therefore always some 'nonscientific' tactics and strategies involved in core-set activity. Some of the 'social' tactics uncovered in the second stage of EPOR include those of "creating contradiction" and "using prior agreements" while using the 'scientific' tests available to settle the issue (see Collins (1985), pp. 134-6). The rationale for participant-observation and the case study approach of EPOR also become apparent at this point. Participant-observation is required because the researcher must find out *which* concepts are being renegotiated and how they are linked to the other activities of each of the core-set members, so that she can explain *how* the process occurs.[14] Examination of many cases is required so that the patterns of explanation of the process of negotiation discerned by the researcher can be generalized.

In the third stage, the researcher completes the explanation of the process of scientific change by linking the activities of the various members of the core-set in the context of the negotiation to their positions in the larger scientific conceptual network (and the other, 'external' networks in the society to which there may be links). These third stage explanations in terms of the 'wider social structures' involves identifying the groups or institutions to which each of the core set members belongs and identifying *their* perceptions of their positions in the

[14]. Actually participating in the process also lends credence to the researcher's claim that the various core-set members' version of events is not the whole story.

network. These could then explain why each member was using their particular strategy in the core-set negotiation, and what sorts of resources (conceptual and material) they were bringing to the negotiation from the 'outside'. The third stage explanation would then document the interplay between these diverse interests, resources and strategies as the correct use of the new concept was 'fixed' in the negotiation process (i.e. as its links with the other concepts in the network were finalized by producing a new set of social conventions or institutionalized rules for the use of the concept).

(ii) Woolgar's methodology: The interrogation of representations

Woolgar's methodology for the analysis of scientific activities is reconstructed here on the basis of his writings from 1981 on. Woolgar (1981) is a critique of the SP methodology[15] for analyzing and explaining scientific knowledge-development, and hints at an alternative which is subsequently worked out in greater detail. Woolgar's proposed alternative calls for a complete abandonment of the SP-type approach when contrasted with the lesser modifications of SP tenets suggested by Collins which have been discussed above. Woolgar (1981) takes issue with SP-type 'forms of explanation' on two basic counts, namely (a) The use of the notions of 'naturalism' and 'causality'; and (b) the identification of 'interests' as a primary explanatory resource.

(a) Woolgar thinks that appeals to naturalism and causality tend to be somewhat mystifying, given the unexplicated sense in which SP accounts utilize these concepts. For example, in its most general form an SP explanation (in its most provocative rendering) is supposed to be a causal one, with 'social interests' being at the foundation of the causal chains that determine the

[15]. The main targets are Bloor (1976) and particularly, Barnes (1978). Although other issues are raised as well, the focus of Woolgar's criticism is the 'interest model' based explanations proposed in these accounts.

contents of particular scientific theories. Woolgar notes that typically, supposedly causal SP explanations fall short of this ideal. Thus SP case studies are held to demonstrate that certain identified interests

> influence rather than determine knowledge production, or that particular scientific episodes can be better understood in the light of the particular interests of the involved parties (Woolgar (1981), p. 369),

but the precise causal linkages are never indicated. It is often insinuated that the reason for this is the complexity of the causal connections involved, but nonetheless it is claimed that given sufficient time and ingenuity, the exact causal influences of the various interests could be identified in principle. Similarly, the sense in which SP accounts are supposed to be naturalistic is equivocated upon (Woolgar (1981), p. 370).[16]

(b) According to Woolgar, taking interests (or any other such factor for that matter) as *the* primary explanatory resource assumes that they

> not only enjoy an unproblematic existence, to be drawn on at will by the investigator, but that their) existence is separate and distinct from the scientific content they are said to explain. (Woolgar (1981), p. 369).

Furthermore, while accepting that all causal explanations must take some concepts to be basic and unexplicated,[17] Woolgar is nonetheless puzzled by the SP choice of *interests* as a special explanatory resource over all other possible 'social' explanans in the absence of explicit justification for such a choice. Thus it is never made clear by SP proponents why interests are themselves not considered to be in need of explanation; why of all things, it is interests that

---

[16]. Woolgar suggests that there are two senses of naturalism appealed to in SP accounts, e.g. Barnes (1978): one which refers to the 'natural scientific' approach (which is experimental, stressing the objective, observable features of phenomena); and the other which is more 'phenomenological' (which eschews generalizations and posits a plurality of methods for studying phenomena). Woolgar thinks that SP accounts appeal to the latter sense of naturalism when examining the phenomena to be explained, yet attempt to cast their explanations in a form compatible with naturalism in the former sense. (Woolgar (1981), p. 370).

[17]. Though for Woolgar, this suggests a shortcoming of causal explanations, see below.

are *not* to be treated as 'actively construed assemblages of conventions or meaningful cultural resources, to be understood and assessed in terms of their role in activity'... (Woolgar (1981), p. 370; Barnes (1978) is being cited)

Woolgar believes, on the contrary, that creation and development of 'interests' must be taken to be *constitutive* of scientific activity, and hence something that must also be accounted for in any satisfactory account of scientific practices:

> Scientists themselves can be seen to be constantly engaged in monitoring, evaluating, attributing... the potential presence or absence of interests in the work and activities of both others and themselves... the construction and use of interests is an aspect of scientific activity that demands treatment in its own right. (Woolgar (1981), p. 371).

Foundational 'social' factors other than interests, if used similarly in causal explanations of the SP variety, would be potentially met with the same objection from Woolgar. It is essentially for this reason that Woolgar rejects all causal forms of explanation of scientific activity and developments: they all make use of resources that are themselves assumed not to be in need of explication, when in fact (and this is crucial) they clearly *are*, given the role they play in scientific activity. Woolgar takes all explanations of the causal variety to be beset by similar 'methodological problems' (I discuss Woolgar's view of the general nature of these methodological problems below).

As an alternative, Woolgar proposes an *ethnomethodological* approach to scientific activity.[18] Such an approach would be geared toward answering the prior questions about the 'practical management' of these methodological difficulties in various contexts (i.e. it would develop an understanding of the nature of the explanatory techniques and strategies used to

---

[18]. The exemplar of this sort of approach, of course, is Latour and Woolgar (1979).

manage these difficulties). He argues that this sort of approach would not suffer the abovementioned drawback in connection with the use of unexplicated resources. His reasoning is that even if the ethnomethodological approach would itself make some unexplicated presuppositions, the goal would not be to provide complete explanations of all features of the scientific activity under consideration. So even if use was made of some unexplicated feature that might be constitutive of the overall scientific activity in question, it would not necessarily be the case that the unexplicated features would be constitutive of *those specific features* of the scientific activity being examined on their basis. This would allow for any set of unexplicated features used by the analyst on one occasion to themselves be the subject of investigation on another occasion. This piecemeal approach to scientific (or, for that matter any other kind of) activity would be possible because the ethnomethodological forms of explanation would not be causal ones (i.e. they would not require an unexplicated or foundational set of factors to explain all aspects of scientific activity; see Woolgar (1981), p. 371). Woolgar thus thinks that SP explanations too can be shown to be utilizing such strategies of managing methodological difficulties to make their explanations 'work'.[19] Some of the motivation for Woolgar's approach should be apparent at this point; I now turn to a more detailed description of his project and its rationale.

Woolgar's primary concern is with the concept of representation, and 'The Problem' associated with it.[20] 'The Problem' concerns the "[grounds which] provide the warrant for the

---

[19]. Some of these strategies have been hinted at above when discussing Woolgar's treatment of 'causality' and 'naturalism' and 'interests', for details see Woolgar (1981), pp. 379-388.

[20]. Woolgar draws attention to the similarity between his approach to the social study of science on the one hand, and the concerns of the French structuralists and post-structuralists with the signifier-signified relationship in the context of the connection between language and world on the other. (See Woolgar (1986), p. 313).

relationship between the objects of study and the statements made about those objects" (Woolgar (1983), p. 240). 'The Problem' thus concerns the nature of the relationship between that which is represented and that by means of which the representation is achieved, or more generally, any relationship between 'surface signs' and 'underlying realities'.[21] It is a problem given his belief that all grounds given to provide such warrant are in principle defeasible on any occasion on which they are invoked. Hence Woolgar regards The Problem as an epistemological one, and takes a skeptical stance when he claims that it is "a general and irresolvable one... which requires artful management whenever it makes its appearance" (Woolgar (1983), p. 241).

The Problem, being an irresolvable one, leads to a series of 'Methodological Horrors' which are never resolved, but only managed or 'momentarily staved off' on each occasion on which they arise. Chief among these horrors are indexicality, inconcludability and reflexivity.[22] Woolgar uses the term 'reflexivity' to denote a feature of the representans-representandum relationship derived from Garfinkel (1967). In this sense reflexivity refers to the

> intimate interdependence between representation and represented object... such that the sense of the former is elaborated by drawing on the knowledge of the latter and the knowledge of the latter is elaborated by what is known about the former. (Woolgar (1988b), p. 33)[23]

Along the same lines, reflexivity is used to denote "the inseparability of a 'theory' of representation from the heterogeneous social contexts in which representations are composed

---

21. Thus a representational relationship exists between such diverse pairs as photograph-scene, action-intention, what is said-what is meant, voltmeter reading-voltage etc. (For a larger list, see Woolgar (1988a), p. 31).

22. 'Indexicality' connotes the 'different' reality referred to by the 'same' representation on the different occasions of use of the representation. 'Inconcludability' highlights the fact that "the task of exhaustively and precisely defining the underlying pattern (meaning) of any one representation is in principle endless" (Woolgar (1988b), p. 32).

23. 'Reflexivity' in Woolgar's sense is thus neither a methodological tenet to be applicable to one's own explanations, nor an injunction for periodic appraisal of one's own theories and actions, but a constitutive (hence unavoidable) *fact* about all representational activity.

and used" (Woolgar and Lynch (1990), p. 12). The 'Horror' arises from the fact that each of these factors contributes potentially to the defeasability of the justification given for any representation on any particular occasion.

Given Woolgar's belief that The Problem is an all pervasive and irresolvable one, the natural question for him to ask is: how is it that The Problem isn't apparent to (i.e. isn't a problem for) those engaged in representational activities? His answer is that it is their 'argumentative strategies' and 'practical managing skills' which circumvent the horrors or minimize their implications. Woolgar's project, therefore, is to investigate the general 'forms of explanation' and the 'management strategies' used by the creators of representations to defuse the threat of the horrors.[24] Since Woolgar takes the notion of representation to be of fundamental importance for the study of scientific activities, the development of his own methodology and his assessment of others is made on this basis.

Woolgar identifies three orientations in the social studies of science in terms of their stance toward The Problem and the attendant Horrors, which he calls the (a) reflective, (b) mediative, and (c) constitutive positions respectively (see Woolgar (1983), pp. 242-3).[25] Woolgar argues that orientations (a) and (b) fail to engage The Problem and satisfactorily account for the management of the Horrors, and that a development of (c) is the most fruitful line of inquiry to pursue vis-a-vis the analysis of scientific (and other representational) activities.

(a) The reflective view is identified with all scientific representational activities in general and

---

[24]. It is with this goal in mind that Woolgar formulates 'key questions' which any adequate sociology of science must answer, see Woolgar (1981), p. 389. (Also see Woolgar (1985c), p. 162, Woolgar (1988b), pp. 7-8 and Woolgar and Lynch (1990), pp. 124-5).

[25]. A similar tripartite division of approaches is suggested in Woolgar (1988b), p. 7; in Woolgar (1988a), pp. 39-41; and in Woolgar (1989b), pp. 204-5.

with the Mertonian tradition in the Sociology of Science in particular. This tradition takes an uncritical attitude toward the idea of representation; in fact it fails to even acknowledge the existence of The Problem. It is grounded in a (naive) realist ontology and an objectivist epistemology: on this view

> real world entities enjoy an existence independent of their description... this view) locates the origin of knowledge in the character of the natural world. (Woolgar (1983), p. 243)

From the analyst's (of scientific activities) point of view, 'science' is held to have a real 'nature' or 'essence' (though it might be claimed to be in fact too complicated to pin down precisely; see Woolgar (1988a), p. 20). The *representations* created by the analyst as well as by the scientist are taken to be unproblematic in that they are assumed to mirror the natures or essences of what is being represented (usually in virtue of the 'method' used to construct the representations). At best, the analyst might be called upon to diagnose the reasons for failure of the proper application of the method when *mis*representations are produced (these are variously taken to be 'social' and 'external' factors as opposed to 'cognitive' and 'internal' ones respectively; see Woolgar (1989b), p. 205). The reflective view is rejected because it fails to acknowledge the existence of The Problem and thus cannot account for the management of the Horrors.

(b) The mediative view (which is taken to include the Strong Programme and the position of Collins as sketched above) holds that

> [scientific] accounts [or generally, representations] can be thought of as products of the social, cultural and historical circumstances which *intervene* between reality and the produced accounts... social circumstances *mediate* in the production of knowledge accounts. (Woolgar (1983), p. 244).

This view is held to be an improvement over (a) because it does acknowledge the existence of the problem in some sense. Typically, proponents of the mediative view argue for the underdetermination of (scientific) accounts by 'reality', and claim that the flexibility of the (multiple) construals that might arise are constrained by social or external factors as well as by purely scientific considerations. On such accounts, from the analyst's point of view an anti-realist stance is taken toward the scientific representations under investigation. The Problem is hinted at with the assertion that *scientific* representations are problematic, and the link here between representation and underlying reality is not as secure as in (a); scientific representations, being underdetermined, do not accurately capture any underlying reality.

However, the analyst does take her own account of scientific representation creation to be a faithful representation of 'what actually happens/happened'.[26] Here the analyst takes a realist stance toward the 'social' world insofar as her representations of events and methodologies for constructing accounts are not taken to suffer the defect of underdetermination or being products of historical, social and cultural circumstances. Woolgar agrees with critics who have pointed out the Strong Programme's lack of consistency on this point in its asymmetrical treatment of the natural and social worlds.

Woolgar has other criticisms of mediative view as well. For one thing, advocates of the reflective and mediative views alike use the strategy of 'ontological gerrymandering' in their analyses of phenomena. Such accounts

> depend upon making problematic the truth of certain states of affairs selected for analysis and explanation, while backgrounding or minimizing the possibility that the same problems apply to the assumptions on which the analysis depends. (Woolgar and Pawluch (1985a), p. 216)

Mediative accounts typically use the strategy of arguing that the scientists' own explanations (for

[26] Though there is often judicious equivocation over the question of what role 'reality', i.e. the natural world is supposed to have played in any specific instance; see Woolgar (1983), p. 262.

the choice of a particular theory, say) are problematic, but that their alternative analyses are not. Historical and cultural relativism (as well as irony and the Horrors) are typically *used* to undermine the claims of scientists and philosophers of science. As a consequence of their desire to provide alternative accounts of 'what actually happens/happened' defenders of the mediative view fail to recognize the full import of The Problem. Hence they end up "[paying] less attention to the fundamentals of argument and persuasion" (Woolgar (1983), p. 262), and to "the fundamentals of knowledge production" (Woolgar (1988b), p. 17) . They also run the risk of being subsumed into reflective type accounts when all is said and done (Woolgar (1989b), p. 206).

(c) On the constitutive view

> accounts [or representations] are *constitutive* of reality... there is no a priori distinction to be drawn between accounts and reality... instead accounts *are* the reality; there is no reality beyond the constructs we imply when we talk of reality. (Woolgar (1983), p. 245)

Proponents of the constitutive view accordingly hold that 'there is nothing outside of the representations', which leads to a relativism which is taken to apply to all representations, including those of the analyst. In other words, a relativist stance is taken toward both the social and the natural worlds, and essentialism/realism in any form is rejected. This is not to suggest that

> there is no distinction between a thing and what is said about it, [but rather] these distinctions [are] actively created achievements rather than pre-given features of our world. (Woolgar (1986), p. 314).[27]

At the same time, the constitutive view acknowledges The Problem squarely, accepting that its

---

[27]. In this regard, a 'splitting-inversion' model is proposed by Woolgar in the case of representations of scientific discoveries (see Woolgar (1983), p. 246; Woolgar (1988a), pp. 68-9). On this model the active constitution or construction of a representation involved in a 'discovery' (and the reflexivity etc. of the situation) is forgotten as the construction of the representation is completed, and the constructed ('discovered') properties are said to have been there all along.

own representations are as much subject to the Horrors as any others.

Woolgar thinks that a thoroughgoing constitutive view is required to avoid the pitfalls of the reflective and mediative ones, and that one especially has to look out for the omnipresent danger of slipping back into these in one's own analyses. Woolgar thinks that these dangers arise whenever the analyst is tempted to use the constitutive view as an *instrument* for other ends (which are often ideological, being motivated by political agendas, see Woolgar (1983), p. 262) rather than as an end in itself. Thus, instead of regarding, e.g. 'reflexivity' as a tool to be used in crafting 'social' explanations of scientific activities, it should be regarded as a project (i.e. as something to be investigated in its own right, Woolgar (1988b), pp. 16-17).[28]

The rationale for this approach is to be found in Woolgar's conception of representation as an ideology and the role of the agent/self in perpetuating it. Recall that representation, the relationship between surface signs and underlying realities, is the source of the Problem and the Horrors, and science is taken to be a highly institutionalized form of representational practice (although by no means the only one). Representation is taken to be *ideological* because the ones doing the representing (the 'agents' of representation) present themselves as passively representing the 'real' nature of that which is represented (thereby presenting themselves as dispassionate and neutral observers) when they are in fact actively constituting it (Woolgar (1988a), pp. 100-1). Representation (or the act of representing) thus has the effect of "[building] rhetorical distance between the observer and the observed object, and [establishing) the

[28]. For a similar distinction between irony as project and irony as instrument in the social study of science, see Woolgar (1983), pp. 257-61.

antecedence of the latter" (Woolgar (1988a), p. 108).[29] Since the representing agent also claims and is generally accorded *authority* in the matter of her representations, the possibility of the represented entity becoming 'disempowered' (whether willingly or not) arises[30], and uncritical acceptance of the ideology of representation may therefore have repressive (even if unintended) effects. It is for this reason that selective relativism of the SP variety does not go far enough for Woolgar's liking, and hence becomes a case of uncritical acceptance of the ideology of representation:

> By insisting on a distinction between science (as object) and our disciplines [i.e. sociology] as resource, we are in danger of mistaking relativist critiques of science for an adequate appreciation of the more general phenomena of representation. (Woolgar (1988a), p. 101)

Woolgar thinks that the first step in overcoming the ideology of representation involves the reestablishment of The Problem as a problem for those who unselfconsciously use representational devices (see Woolgar (1983), p. 263; Woolgar (1985), p. 225; Woolgar and Pawluch (1985a), p. 362). The idea is for the analyst to pursue the constitutive view as a project, to "continually interrogate and find strange the process of representation while we engage in it" (Woolgar (1988b), pp. 28-9) and create practices "in which the interrogation of methods proceeds simultaneously with, and as an integral part of, the investigation of the object" (Woolgar (1988b), p.8). These strategies are intended to break the stranglehold of the ideology of representation, and the "obsession with technical rationality" which is characteristic of contemporary society (Woolgar (1988a), p. 101). Woolgar's development of the projects of

---

[29]. On Woolgar's view neither the agent of representation nor the represented entity need be human beings. For example various scientific instruments or 'inscription devices' are taken to be agents of representation. Also see the more detailed discussion of this in the next Chapter.

[30] See Fuller (1992a) for details.

reflexivity and new literary forms (see Woolgar (1988b)) are meant to be a step in this direction, attempts to undermine one stronghold of the ideology of representation- namely conventional *textual* representational practice.

As noted above, the overall goal of pursuing the constitutive view as a project where Woolgar is concerned is the hope that it will lead to new ways of asking questions about the nature of the (scientific) knowledge-production process; ways that will not be flawed in the sense of taking representations to be anything other than actively constructed and occasioned products. This would, thinks Woolgar, provide a way to transcend the distinction between relativism and objectivism and clear the way for an investigation of how explanations and representations are actually constructed and managed in practice in various settings.

Chapter Two: Collins' and Woolgar's analyses of Slezak

In a recent paper published in the journal Social Studies of Science[1], Peter Slezak has argued that recent developments in the interdisciplinary field of Cognitive Science (or more specifically, the development of certain computer models of the process of scientific discovery within Artificial Intelligence [AI] research) amount to an empirical refutation of the Strong Programme in the Sociology of Scientific Knowledge.[2]

In this Chapter, I (i) give a nutshell-version of Slezak's argument for the empirical refutation of SSK, placing it in the broader context of his conception of Cognitive Science; and (ii) describe Collins' and Woolgar's responses to Slezak with reference to their methodologies for the analysis of scientific practices as explicated in the previous Chapter[3]. This allows me to formulate criticisms of them in the following Chapter.

(i) Slezak's 'empirical refutation of SSK'

Two points should be noted at the outset. Firstly, while Slezak is ostensibly concerned to refute the 'Strong Programme' in SSK, one should not be misled into thinking that it is the Edinburgh School of Barnes and Bloor which is his particular target. His argument is a more general one, meant to be a "demonstration of a case in which scientific discovery is totally

---

[1]. See Slezak (1989a).

[2]. Slezak is Director of the Centre for Cognitive Science at the University of New South Wales, Australia.

[3]. While I use Collins (1989) and Collins (1990) for the details of his response to Slezak, I have more or less 'constructed' a hypothetical response on Woolgar's behalf as he does not himself say very much on the issue.

isolated from all social and cultural factors whatever" (Slezak (1989a), p. 563).[4] The refutation is therefore meant to count against all variants of the Sociology of Knowledge or philosophical accounts that make use of any 'social' considerations whatsoever. Secondly, while it might appear that it is the specific case of the computer modelling of the processes of scientific discovery which *by itself* constitutes the evidence for

> a 'pure' or socially uncontaminated instance of inductive inference, [as these programs] are capable of autonomously deriving classical scientific laws from the raw observational data (Slezak (1989a), p. 563),

Slezak actually wants the evidence for such modelling to be taken in conjunction with the other developments in the larger interdiscipline of 'Cognitive Science' (henceforth 'CS') taken as a whole; this evidence is ultimately a small contribution to the coherent and converging lines of research in the component subdisciplines: linguistics, philosophy, psychology, neuroscience and AI. According to Slezak, these converging developments point to a computational, information processing and symbol manipulating model of the mind which in turn furnishes a materialistic, causal-law based model for the explanation of cognition and action in human beings. These models obviously extend to the case of scientific cognition and activity as well, thereby vindicating the 'traditional view of science' and refuting any kind of attempt at explaining science by appealing to 'social' considerations.[5]

(a) Slezak's overall argument:

First, the interrelationship between CS and computer models of the scientific discovery process:

---

[4]. All subsequent references to SSK or the strong program in this section should be taken to refer to this conception.

[5]. The 'traditional view' being the one for Slezak which holds that "there are principles of rationality and a 'scientific method' that are independent of social factors" (Slezak (1989a), p. 566).

(i) CS is effectively part of the project of naturalizing epistemology.

(ii) CS is unified by a shared conceptual model of cognition- one that abstractly defines a novel way of understanding cognition in individuals as a mechanical process (namely, the Turing Machine concept).

(iii) This abstract model is physically instantiable in many ways (and it is instantiated in human beings in one specific such way).

(iv) One of other the ways in which this abstract model is (explicitly) instantiated is in the modern digital computer.

(v) The computer can therefore be used to model and test/verify hypotheses about cognition deriving from this abstract model.

(vi) There is evidence that certain features of cognition (as defined by the abstract model) are independent of social and other contextual considerations.

(vii) The essential features of scientific cognition, too, can be shown to be uncontaminated by social considerations, given the evidence provided by computer models of the process of scientific discovery.[6]

Claims (i)-(vi) involve the broader considerations drawn from CS, while (vi) and (vii) pertain to the actual empirical refutation of SSK itself. I now describe the argument of (vii).[7]

(b) <u>The argument from the computer models of scientific discovery</u>:

As noted above, Slezak thinks that there is evidence from the computer modelling of cognition to support the claim that certain general features of cognition are insulated from

[6]. See also Slezak (1986), Slezak and Albury (1988) and Slezak (1990).

[7]. While I touch on some aspects of theses (i)-(v) in the discussion of Collins' and Woolgar's responses to his arguments in (vi) and (vii) below, the greater part of my discussion of (i)-(v) is in Chapter 4.

'social' considerations. Computer models of scientific discovery are supposed to demonstrate that these, too, involve only such insulated general features. Slezak describes these models as tests specifically designed "to abstract science from its contingent embodiment in a ubiquitous social context" (Slezak (1989), p. 556).[8] According to Slezak, the refutation of SSK follows because these models provide a 'sufficient demonstration' or 'existence proof' of the possibility of scientific discovery occurring without social considerations entering into the picture (i.e. the very existence of the models together with the fact that they 'work' suffices to refute SSK).[9]

The scientific discovery process is conceived of as a problem-solving activity very much like problem solving in other ordinary, everyday contexts, "sharing [with them] many methods, techniques and strategies [or 'heuristics'] of solution" (Slezak (1989), p. 567). The models themselves[10] involve suitably formalized versions of these heuristics in various scientific domains to demonstrate they can 'solve problems' of scientific discovery:

> The BACON programs employ a small set of data-driven heuristics or rules for detecting constancies and trends in data, thereby defining theoretical terms to describe the data parsimoniously and formulate hypotheses... the heuristic methods employed by the program, by introducing invariants, permit it to postulate certain intrinsic properties of objects (Slezak (1989a), p. 568),

thereby 'inducing' scientific laws from the input data. Similarly,

> ...Lenat's AM program, using theory-driven processes [heuristics]... in the domain of set theory... was able to discover the natural numbers as well as the operations of addition, subtraction, multiplication and

---

[8] By 'science' here I take Slezak to include general principles of rationality and method as well as the content of specific theories.

[9]. Slezak doesn't suggest that these models are supposed to capture the process of discovery as it actually might have happened (though he does not reject the possibility that certain aspects of the actual process itself too can be modelled).

[10] Slezak uses Langley and Simon's BACON programs and Lenat's AM program- see Langley, Bradshaw and Simon (1980) and Lenat (1978)- to illustrate his case.

> division... (Slezak (1989a), p. 570),

hence illustrating heuristic-based 'theory' development.

Slezak thinks that these heuristic-based models provide two (related) arguments against accounts of the scientific discovery process that would resort to using social or contextual factors in their explanations. The first is from the abstract nature of the heuristics used and the second is from their generality. The first argument is that since these heuristics have been derived (following Langley, Simon Bradshaw and Zytkow (1989) in this) through a process of abstraction from detailed study of actual cases of problem-solving, what results are

> purely formal heuristics which are the embodiment of quite *general* problem-solving techniques which would be difficult to connect in any way to specific social phenomena- least of all those that might have prevailed [at the time of the discovery, especially if that discovery had been made quite some time ago] (Slezak (1989a), pp. 573-4).

This argument is meant to shift the burden of proof onto those who would deny that the heuristics were sufficiently abstract, challenging them to point out the 'socially contaminating' features (and presumably, Slezak would follow up by insisting that any 'contaminating' factors detected could also be eliminated). The assumption behind this argument is that essentially the *same* formal heuristics or rules of discovery (or formally equivalent ones) are used when a particular discovery is originally made (say, Boyle's Law as discovered by Boyle from the data), subsequently reproduced by others (say by me in a physics laboratory class), or modelled on a computer (by BACON).

The second argument relies on the generality of the heuristics- not only is the same set of heuristics used to make the same discovery at different times, but the same set of heuristics is used to make different discoveries. This is because the heuristics

> do not use or require prior information about the problem domain... the very possibility of such general

methods of discovery, applicable to a wide diversity of problem domains, brings into sharp relief the problems they pose for social constructivists- namely the futility of trying to correlate the content of theories with the [obviously different] social circumstances that might have prevailed for the original discoverers (Slezak (1989), p. 570).

And while the BACON heuristics cover only a single aspect of the process of scientific discovery (that too in a very restricted sense), namely the process of making inductive generalizations from data, Slezak would no doubt claim that similar heuristics which govern the other aspects of scientific discovery could surely be discovered as well, once these were clearly identified).

Finally, the link between these formal heuristics as used in computer models and the case of discovery by human beings is as follows. Slezak believes that any human being who makes a particular discovery must have (somehow) possessed the heuristics on the basis of which the discovery is made; hence both Boyle and BACON possess the heuristics required to induce Boyle's Law from the data.[11] As to the question of how human beings 'possess' these heuristics (i.e. 'where' they are), Slezak believes that they are 'internally represented' in human beings in a manner analogous to the way the BACON program is represented in particular machines running it.

(ii)(a) Collins' response to Slezak

Collins' response to Slezak hinges on the issue of whether computer models of scientific discovery can be said to instances of scientific discovery strictly speaking.[12] He concedes that

---

[11]. However, the computer model may still have normative import because the 'inducing-laws-from-data' heuristic it possesses is the pure article, uncontaminated by the social and contextual factors that may have been mixed up with the 'same' heuristic as possessed by Boyle.

[12]. See Collins (1989) for his response to Slezak's original article, and the more detailed account in Collins (1990).

if a computer model like BACON can be said to be making scientific discoveries, then social considerations can indeed be deemed to be irrelevant to the process. Since he thinks that social factors are always involved in the process of scientific discovery, he argues that computer models "including all future versions of BACON and its ilk that do not involve unforeseen design principles" (Collins (1989), p. 614) cannot, in principle, be said to be discovering scientific laws. Collins thinks that the claim that the computer modelling of scientific discoveries can indeed be considered instances of discovery is an ambiguous one open to two interpretations, which he calls the (1) 'encapsulated community' and (2) 'scientist-mimic' interpretations respectively[13] (Collins (1989), p. 614). He argues that neither of these can be sustained, and hence the attempted refutation from computer models fails.

(1) The encapsulated community interpretation:

On this interpretation, a model like BACON[14] is conceived of as independently reproducing the discoveries made by whole scientific communities:

> The encapsulated community interpretation is that BACON can do everything that sociologists of scientific knowledge claim is the prerogative of the community (Collins (1989), p. 616).

Collins thinks that if the encapsulated community interpretation is true, BACON should in principle be able to autonomously reproduce *all* (and only) presently accepted scientific discoveries, i.e. reproduce the history of science.[15] If BACON could do this then all

---

[13]. The encapsulated community interpretation takes the model to be simulating a society, whereas the human-mimic model takes it to be equivalent to an individual within that society.

[14]. All uses of 'BACON' in the remainder of this section should be interpreted as 'BACON or models of the BACON type'.

[15]. One might ask whether this is too strong a requirement on Collins' part. After all, surely the 'community' theory would be refuted even if BACON should reproduce even a single scientific discovery. The stronger requirement that BACON be able to reproduce entire the history of science has polemical import given that 'community' activity is only conceivable as extended over time, so that a BACON would have to model (or otherwise account for) this

'community' based theories of discovery would be refuted, as an isolated machine is not a social collectivity.

Collins argues for the in principle impossibility of BACON simulating/reproducing the discoveries of an entire scientific collectivity as follows.[16] In the case of the processes by which the data are obtained, he notes that there is an immediate difficulty for BACON in that it is not able to perform its own experiments and gather its own data for the purposes of making scientific discoveries (Collins (1989), p. 614); such data have to be 'fed' to it. But this is problematic, because the process of gathering such data is itself usually performed by a community of experimentalists, who engage in

> filtering what is to count as proper 'data' from the results of badly designed experiments or observations, the results of 'preliminary runs', the results of unspecified errors...[etc] (Collins (1989), p. 614-5).

Hence, by giving BACON a particular set of data, we have ensured that "the human social collectivity has made its presence felt in filtering the data- and thus in predetermining the results" (Collins (1989), p. 616). This doesn't mean that Collins would want to say that choosing a particular set of data would inevitably predetermine their 'correct' *interpretation* (though there is a connection here, see below)- but rather that the preselection of the data given to BACON eliminates that part of the process of 'constraining interpretive flexibility' of the data which a truly isolated BACON should be required to go through- it is in this sense that it isn't clear "when our imaginary isolated BACON asks for the data on, say, suspended oil drops...

---

dimension.

[16]. Collins' argument here appears confused. Thus Slezak claims in his response (Slezak (1989b), p. 681) that Collins seems to have conflated the cases of (i) processes by which data from experiments are obtained and (ii) the processes by which the data obtained are interpreted. While it is true that what Collins has to say on pp. 614-617 is badly stated, I think that a consistent argument can be reconstructed from Collins' point of view- one which even makes a virtue of the above conflation, see below.

whose numbers should be provided" (Collins (1989), p. 616). A genuinely isolated BACON would gather its own data, but Collins doesn't see how it could possibly do this.

A similar difficulty arises with respect to the process of *interpretations* of data. That is, even if the above problem is ignored, Collins does not see how BACON could choose between the conflicting interpretations of what are regarded as the same set of data. The process of establishment of an interpretation, too, is a matter of negotiation between various parties which results in the 'correct' interpretation being fixed. In the case of an autonomous BACON

> it is hard to know what criteria BACON would use to choose an interpretation and a solution... given the same data, one BACON, whose program had reached the early years of this century would believe in integral electric charges, while another would believe that the unit of charge could be divided indefinitely (Collins (1989), pp. 616-7).

Insofar as any actual BACON did manage to come up with what turned out to be the correct interpretation, Collins would claim that this would be because the heuristics it possessed (again in virtue of human intervention) would have predetermined the correct interpretation. Again, it isn't clear to him how an isolated BACON would come up with these heuristics by itself. In any event, it should be noted that for Collins these two cases are not really distinct ones- after all, on his view the processes of obtaining data are inextricably intertwined with the processes of interpreting them.[17] The problem of gathering data and the problem of interpreting them for Collins are then really two aspects of the same problem.

To summarize, Collins makes essentially the same argument in both cases- it is not conceivable that the processes of scientific discovery, involving dynamic negotiation between

---

[17] Collins would deny a rigid theory-observation (or content-context) distinction, or argue for a 'theory-ladenness' of observations in keeping with his commitment to epistemological relativism. This is not to say that Collins doesn't think that there is a distinction to be made between theoretical concepts and experimental results, but what these happen to be are themselves fixed by the relevant communities.

groups of people that result in sui generis outcomes can be reproduced by an autonomous machine on the basis of abstract heuristics. The guiding intuition behind these arguments is that the notion of an autonomous machine simulating a scientific community is clearly absurd, because one of the tasks performed by the community is the fixing of meanings and interpretations (e.g. by the settlement of controversies through negotiation), and an isolated machine will lack the resources to accomplish this task by itself. The connection between this argument for the untenability of the encapsulated community interpretation and Collins' methodology for the analysis of scientific practices should be evident, as the argument is really a restatement of Collins' methodological commitments in another guise. The 'encapsulated community' that BACON would have to model in order to make any particular scientific discovery would include the 'core set' and the 'network of concepts' of the society whose scientific discoveries were being modelled. With the establishment of each new discovery with the completion of the negotiation process within the core-set, the 'network of concepts' would undergo various changes, the details of which would be unpredictable in advance.[18] Hence, it is inconceivable that 'BACON and its ilk' could replicate the process: when all is said and done, "BACON is not a social collectivity" (Collins (1989), p. 614).

(2) The 'scientist-mimic' interpretation:

Collins thinks that the interpretation on which BACON could be taken to be simulating "individual scientists' contributions to discovery is neither incoherent, nor obviously false" (Collins (1989) p. 617). At the same time, he believes that if it were true, it would refute not only SSK,

[18]. Recall that Collins holds that even if the 'rules and procedures' used in negotiation are fairly well established at any given point in time (though they too are diachronically mutable) and do not necessarily involve any overt 'contextual' considerations, their *deployment* is a purely contingent matter, grounded in tacit conventions.

but any idea of a social science that takes social collectivities to be the primary unit of analysis... [i.e. all] approaches [which] take human abilities [including the ability to participate in making scientific discoveries] to exist by virtue of membership of social collectivities (Collins (1989), p. 617).

The refutation would follow because "machines are not part of social collectivities" and successful replacement of an individual scientist by BACON would show that it wasn't necessary to be part of a social collectivity in order to participate in the process of scientific discovery.

Collins' reasons for thinking that machines are not (and could not possibly be) part of social collectivities involve the 'degree of connectedness' that human beings have with the society to which they belong. This connectedness results from the fact that human action derives from the 'rules embodied and institutionalized in forms of life' that they have immediate access to, or awareness of.[19] Though these rules are (in principle) mutable and subject to unpredictable extensions in future cases, two individuals who share membership of a social collectivity do so in virtue of an overwhelming number of shared static, institutionalized rules that ground their actions (and that of the other members of the society- see Collins (1985, p. 16). It is in this sense that the community or society is the primary unit of analysis, and the ability to act (and to understand action) presupposes the community within which the actions are embedded.

Collins' argument that machines lack this degree of connectedness with social life derives from the fact that e.g. what BACON does is only a small (and relatively insignificant part) of what human beings do when they engage in the activity of making a scientific discovery. What the machine lacks in terms of greater connectedness than this is covered by the catchall phrase

---

19. The connection between Collins' phenomenological leanings and his relativism is detailed on pp. 15-16 of Collins (1985).

"the ability to anticipate the social nexus into which the discovery will be cast..." (Collins (1989), p. 618). Collins goes on to add that "no current or foreseeable machine" could do this, i.e. make the requisite judgements to tell the difference between "what might make a sensible discovery, what might make a risky discovery and what might make a nondiscovery" (Collins (1989), p. 619). Thus machines cannot in principle be members of social collectivities, machines "just do not do what humans do".[20] While it might be granted (and Slezak would readily admit) that no current machine does anticipate the 'social nexus' within which it operates, Collins' insistence that no foreseeable machine could do this is at odds with his general methodology.[21] What Collins should be doing in this instance is what he has purportedly done in his other case studies: he should try to identify the 'core-set' in this dispute, and then the internal linkages of the core set members to each other and their external linkages to the larger society, and then proceed to document the various negotiations and their ultimate outcome. Collins is candid enough to admit this: on his account the analyst of scientific practices is merely supposed to describe; taking a stand on this issue means "moving away from relativistic sociological analysis in order to participate in the debate" (Collins (1990), p. 190).

Why, then, does Collins insist on going on to participate in the debate in this instance? The answer to this question is tied up with what Collins sees as the explanation of BACON's partial success in 'recovering Boyle's Law from the data'. Collins believes that BACON will be able to reproduce that part of the process of discovery "which requires rote reproduction of what has

[20]. This judgment presumably expresses the societal perception, institutionalized in our forms of life in various ways.

[21]. If anything, on Collins' account the ultimate decision on whether BACON could be taken to be discovering scientific laws, including the decision on what criteria would be used for making this judgement would follow from a process of negotiation within the relevant community.

been learned [or what can be done] mechanically" (Collins (1989), p. 622), which includes in this instance the algorithmic procedure for inducing the relationship between given pairs of data points. The same is true of all human activities that can be reproduced by machines: machines can mimic humans (or become substitutes for humans within a society) only to the extent that humans do what they do in a machine-like way. And in turn, humans need act in a machine-like way only to the extent that they are prepared to forget that there are essential differences in human actions generally and that subset of their actions which machines can imitate. These essential differences involve a distinction that Collins draws between 'behaviour-specific' and 'concerted' actions (corresponding to the difference between behavior and action, see Collins (1990), Chs. 3 and 4. I discuss this in the following Chapter).[22]

Given these differences, Collins fears that accepting machines of this sort as members of society or as 'social beings' will lead to two unpalatable general consequences:
(1) we will start to behave more like machines ourselves as our appreciation of these essential differences will become lost, and this will involve the loss of our freedom of action, and in due course (2) our image of ourselves will become more like our image of machines, and "we will see departures from the machine like ideal as a matter of human frailty rather than human creativity" (Collins (1990), pp. 222-4). The net result would be 'dehumanization', hence the importance of Collins' taking a stand on the matter:

> if we use [intelligent machines] with too much uncritical charity, or if we start to think of ourselves as machines, or to model our behaviour on their behaviour, not only will we lose sight of what we are, but the tools we make will not be good ones (Collins (1990), p. 224).

---

[22]. It is interesting to note that Collins does not base the difference in the fact that BACON is a strictly algorithmic rule-following machine. Thus, with regard to connectionist machines which "appear to build their knowledge in sets of connections that are not obviously the equivalent of explicit rules", he claims that this does not mean that "they are, by that fact, potential members of our society... they are still not social beings" (Collins (1990), p. 18).

(ii) (b) Woolgar's response to Slezak

Woolgar's (1989c) response to Slezak does not overtly engage the issues of the possibility of 'asocial discovery' by computer and its implications for SSK. It does, however, consistently draw on his overall methodology and goals in its attempt to 'transcend the debate' and to illustrate how the SSKer should respond in such situations. At the same time, his writings elsewhere[23] do directly bear on the issues of the possibility of 'asocial discovery' by computer and its implications for SSK. In this section, I will first use these other writings characterize what I take to be Woolgar's position vis-a-vis Cognitive Science and the possibility of 'asocial discovery' by computers. I will then discuss how his response to Slezak can be understood given that position, and draw some connections between his position on the issues, his response to Slezak and his general methodological commitments as discussed in the previous Chapter.

(1) Cognitive Science and 'asocial machines':

Recall from the previous Chapter that Woolgar assents to what he takes to be the general phenomenological thrust of

> the extent and scope of the concept of the 'social' as used by SSK... when action is construed as a move in a language game then all actions are 'social'... 'social' no longer constitutes a contaminating influence, since in this view it makes no sense to conceive of the presence of the social as an 'influence', let alone a contaminating one (Woolgar (1989c), p. 660).

The prospects of an 'asocial discovery' on this view would then be nil almost by definition: actions involve agents and are moves within language-games', specific representations and interpretations of these actions no less so.

How does Woolgar develop this SSK theme in the context of the Cognitive

[23]. See Woolgar (1987), Woolgar (1990) and Woolgar (1985c).

Science/discovery by computer case? To the extent that he considers all actions to be 'social' his view is very much like that of Collins. However, Woolgar differs from Collins in insisting that there are no distinctions to be drawn between human beings and machines from the analyst's point of view,[24] as insisting on an a priori restriction of the 'social' to human individuals or groups might amount to the identification of an essentially human attribute. He thinks that

> an examination of sociological critiques of cognitivism [like that of Collins] reveals the crucial role of these sociologists' commitments to particular methods for constructing the nature of man... [and] the sociological commitment to particular methods for representing the character of human action and behaviour [which]... crucially underpins preconceptions about what man is (Woolgar (1987), p. 312).

This is against the spirit of SSK in general, as

> SSK doesn't say there is a difference [between human beings and machines]... it has no wish to appropriate a warrant for specifying either the existent or nonexistent character of a difference. SSK argues that whether or not there is a difference is the upshot of what certain groups take to be the case (Woolgar (1989c), p. 663).

Woolgar therefore sees the Cognitive Science/AI debate as providing another opportunity for SSKers to push their (unrestricted) conception of the 'social' into the domain of what are clearly regarded as 'nonsocial' phenomena (the behavior of certain kinds of machine in this case). Drawing a parallel between the development of the SSK approach to the study of science in the post-Kuhnian era and the projected role for SSK in the case of Cognitive Science, he wishes to develop a strategy by which the distinction between the 'cognitive' and the 'social' might be overcome by SSKers:

---

[24] In brief, where Collins agrees with Slezak that machines are 'asocial' but disagrees with him on whether they can be said to make discoveries, Woolgar disagrees with Slezak in insisting machines are social, but agrees with him that machines could come to be seen as making discoveries.

We need to recognize such distinctions as the achievement of science, as a resource for their characterization of behaviors and practices, and as deeply ingrained in a discourse that sustains its own practice as scientific (Woolgar (1985c), p. 559).

Research in Cognitive Science and AI is marked by the same rhetoric of progress typical of any Science, and symptomatic of the three problems that an adequate sociological understanding of scientific practices must deal with: (i) the dichotomy between the 'scientific' (in this case 'cognitive') and the 'social', (ii) the dichotomy between underlying realities and surface signs (in this case 'cognition' or 'intelligence' and its outward manifestations in behavior), and (iii) the failure to recognize the interpretive flexibility of key concepts (e.g. 'intelligence', 'thought') and the 'occasioned nature' of their use.

Given the above flaws of Cognitive Science from the SSK viewpoint, Woolgar believes that uncritically adopting the categories of Cognitive Scientists and AIers could potentially restrict the scope of sociological inquiry: for example,

sociology is left uncertain as to the intelligent character of its subjects, and has to wait upon the outcome of what (currently) seems an interminable research 'progression' (Woolgar (1985c), p. 565).

Instead, Woolgar proposes two mutually reinforcing ways in which SSK could proceed with its investigation of Cognitive Science theories and research practice. The first would be the development of an autonomous sociological approach to the 'same' phenomena that Cognitive Science takes as its subject matter (e.g. the phenomena of reasoning, thinking, seeing, knowing, understanding etc). The second would be developing an account of 'Cognitive Science research practice'. This would

take as topic the dichotomies and distinctions which characterize this discourse [of Cognitive Science]... [investigating things like] the relationship between the pronouncements of spokesmen on behalf of AI [or Cognitive Science] and the practical day-to-day activities of AI researchers (Woolgar (1985c), p. 567).

While the 'sociology of cognition' would attempt to dissolve the distinction between the 'cognitive' and the 'social' (including the boundary between 'human beings' and 'machines', see below), the accounts of Cognitive Science practice would reveal the constructed and accomplished nature of the distinctions produced by the practitioners to complement their collapsing.

What is Woolgar's rationale for taking this approach? Insofar as Cognitive Science and Cognitive Science inspired AI-research is informed by a particular 'theory' or 'model' of cognition (e.g. computationalism), Woolgar thinks that it is committed to an 'essentialist' view of cognition. This essentializing amounts to the creation of a reified object of inquiry (viz. 'cognition'), and leads to a research program in which the goal is to produce an 'explanation' of the relationship between an object and its putative manifestations on the basis of the model. Accepting the Cognitive Scientist's 'explanations' of cognition allows for the setting up of distinctions between human beings and machines in the following way: the community of cognitive science researchers get to pronounce judgement on the nature of the similarities and differences between human beings and various sorts of machines. As such, they also appropriate authority to dictate the nature of the changes in these distinctions. Now, while Cognitive Scientists are committed to some form of cognitive essentialism, so do certain sociological critics of their work seem to be. Though approving of certain aspects of Coulter's (1983) critique of cognitivism[25], Woolgar points out that

> in accepting as unproblematic the notion that human behaviour has a physiological basis, Coulter implicitly endorses the distinction between physical and social phenomena, and questions only the nature of the

[25]. Coulter (1983) is a 'neo-Wittgensteinian' critique of the Cognitive Scientists' attempts to reconstitute ordinary attributive mental predicates (e.g. 'thinking') as 'scientific' ones via postulation of 'inner processes' like 'brain' and 'central nervous system' functions to explain them.

distinction between them (Woolgar (1987), p. 317).

Furthermore, by insisting that human actions like 'thinking' that have been appropriated by the cognitivists be redesignated as 'social', Coulter runs the risk of replacing, say, computationalist 'explanations' of thinking with a sociological alternative, and getting needlessly embroiled in the debate over what 'cognition' or 'intelligence' really is (i.e. when Coulter argues that Cognitive Science/AI research has no bearing on the explanation of human action or intelligence and that the technical achievements and functional successes of AI be delinked from their 'theoretical' baggage, he might not only be (re)asserting a strong sociological version of the human-machine distinction, but also suggesting that alternative, i.e. sociological, explanations of these concepts *can* be given).[26]

Both cognitive scientists and their sociological critics therefore embrace a version of the human-machine distinction that draw boundaries around the domains of the 'cognitive' and the 'social'. Since the 'social' is all pervasive on Woolgar's conception of SSK, acknowledgement of the existence of contrast categories by sociologists is a mistake: "The revelation of social regularities is in danger of merely supplying new facets of 'intelligence' [etc.] for further axiomatization" (Woolgar (1987), p. 326). Instead of attempting to provide definite descriptions of social behavior, SSKers should take seriously the indeterminacy (and the flexibility of construal and occasioned nature of use) of all 'explanations' (causal or otherwise) of behavior. As a result, the SSKer should be content to merely examine the construction of explanations of human action and agency without pronouncing upon their merits. At the same time, she should also be concerned with the construction of explanations of machine 'agency' and 'action'. In

[26] In this connection, Woolgar points out that to the extent that Coulter provides detailed criticisms of Cognitive Science research, these would be used to 'improve' it, while his more global concerns would be ignored.

this regard, presupposing the machines to be social beings and refusing to acknowledge any distinction between humans and machines would allow the SSKer *access* to the machines without needing to rely on the expertise of cognitive scientists 'explanations' to understand them.

Presupposing machines to be agents would also give the SSKer the necessary 'anthropological distance' to be able to account for the construction of the cognitive scientists' explanations of the machines' 'intelligent' actions.[27] This would allow the SSKer to construct a story about the practices of the cognitive scientists in their construction of a hierarchy of explanations (or 'order of representation') of human and machine behavior and activities. In light of the above, Woolgar's response to Slezak's argument about the refutation of SSK in view of computerized scientific discoveries could perhaps have been something like this.[28] He would perhaps begin by identifying Slezak as a strong computationalist, and Slezak's use of discovery programs such as BACON as an attempt to legitimate the computational thesis/model by 'demonstrating' the essential similarity in the cognitive behavior of BACON and human beings. He would supplement this with a detailed ethnographical study of the activities of cognitive scientists and their machines. He would document the rhetoric of progress underlying Slezak's description of the cumulative accomplishments of the successive programs in the BACON series, and how Slezak used these 'impressive' results as evidence for the increasing success of the model and the need for further research in various areas. He would then establish the

---

[27]. By giving machines agency 'up front', the SSKer would be able to 'find strange' the representations of its doings constructed by the cognitive scientists. The contrast provided by this radically 'alien' viewpoint would apparently supply a perspective from which the constructed explanations of the cognitive scientists could be analyses without recourse to any of their presuppositions. (Compare Garfinkel's 'troublemaking' ethnographical experiments).

[28]. This is a construction on my part.

elusive nature of the 'real meanings' of key concepts by revealing Slezak's flexible construals of these terms in various contexts, and the occasioned nature of their use in actual cognitive science research. He would also come up with alternative fictions about the construction of explanations of these key concepts by the cognitive scientists, and the corresponding attributions of various abilities and properties to humans on the basis of say the BACON programs.[29]

(2) Woolgar's actual response to Slezak:

Woolgar would therefore see the need to take BACON to be a social agent *because* BACON was being represented as an asocial entity in Slezak's moves within the 'language game' of cognitive versus social argumentation. Would he say that BACON was making scientific discoveries? Perhaps he would not take a position on this, claiming that such a question would be settled by the various social groups battling it out over the issue while his task would be to narrate how various groups on various occasions and for various ends granted and withheld the attribution of 'discoverer' to BACON.[30] In his 'official' response to Slezak, however, Woolgar does not set out any of this. There, he takes the view that being locked in conventional forms of discourse like academic debate with cognitive scientists cannot bring about changes of mind- they merely serve to perpetuate the dichotomy between the opposed positions and result in a pointless set of exchanges where the participants end up talking past each other. In order to 'transcend the debate' he attempts a different way of articulating the SSKer's special, constitutive sense of the 'social' by way of a rhetorical strategy based on

[29] As well as the constructed attributions to the BACON programs of some hitherto essentially human abilities and properties and the non-attributions of others pending 'further research'.

[30]. In this, Woolgar represents himself as the quintessentially uncommitted neutral observer, but see the next Chapter.

paradox.[31]

Woolgar (tongue in cheek) asks his readers to suppose that a BACON-like program (he calls it 'COLLINS') could, using abstract heuristics, reproduce certain aspects of SSK 'discoveries' about the social basis of scientific knowledge. Would this lead us to say that the arguments made by Slezak about BACON had been 'decisively empirically refuted'? Furthermore, would we be prepared to say that what we had on our hands was an instance of the 'scientific' discovery of the 'asocial' essence of the fundamental SSK tenets, which after all claim that scientific discovery is always influenced by social factors? These puzzles are supposed to highlight the fact that it is not so much what the machines are doing 'by themselves' as what their human interpreters are claiming for them (e.g. how the BACON discoveries are being represented by both sides) that should be the focus of the discussion. The paradoxes are meant to underscore the futility of debate to SSKers given the different senses of terms such as 'social' and 'discovery' employed by SSKers and cognitive scientists, and to suggest that instead of *arguing* with Slezak over his occasioned pronouncements thereby producing alternative pronouncements of their own, SSKers should simply *assume* their constitutive sense of the social and proceed to investigate how such representations are created and sustained in practice (including through such debates). With respect to the debate itself, Woolgar suggests that the most fruitful way out of the paradox would simply be to consider both BACON and COLLINS[32] to be social beings. In that case, the astute SSKer would be free to remain above the fray to 'simply document' the construction of representations of these entities by their

31. This is clearly meant not so much to appeal to Slezak as to the other SSKers.

32. While 'COLLINS' does not exist, what Woolgar is perhaps suggesting here is it could be taken to be the reified expression of all representations produced by SSKers (e.g. Collins' EPOR).

respective spokespersons, without contributing any representations of her own.

(3) Connections to Woolgar's methodology:

To sum up, Woolgar's opposition to Cognitive Science (especially the computational variant seemingly endorsed by Slezak) follows from his anti-essentialist methodological stance. The need for an alternative, phenomenologically oriented, constitutively social conception of such phenomena as 'understanding' and 'thinking' is linked to the rejection of the essentialist Cognitive Science models. At the same time, he takes attempts to construct 'social essence' explanations (even if based on noncausal 'social rules' for attribution of such phenomena) to be flawed in the same way (recall his critique of SP as described in the previous Chapter, and his critique of Coulter above). This is also related to his rejection of the social-cognitive distinction, as the postulation of 'essences' (whether social or cognitive) is what sustains the distinction. His insistence that the analyst ought instead to limit herself to observing the attribution of these characteristics and the variability of the attributions derives from his ethnographical commitment to the 'presuppositionless' investigation of the active construction of 'explanations'. This in turn leads Woolgar to suggest abandonment of the human-machine distinction in the context of the present debate (as here this is what permits the cognitive-social dichotomy to be sustained).[33] This is Woolgar's way of 'interrogating the ideology of representation' (see previous Chapter) in the context of the Slezak case.

---

[33]. As noted above, I take this to be a strategy adopted by Woolgar in the context of this particular debate: since both sides (albeit for opposing reasons) draw this distinction, its rejection provides a standpoint from which the analyst can better recognize the constructed and occasioned nature of the distinctions actually drawn.

Chapter Three: A Critique of Collins and Woolgar

I now use the material of the previous two Chapters to criticize Collins' and Woolgar's methodologies for the study of scientific knowledge development, using their analyses of Slezak's 'refutation' to delimit and illustrate my critique.

(i) Collins' methodology and response to Slezak

An entry point for my criticisms is best secured by initially drawing a parallel between Collins' methodology for the study of scientific developments as described in Chapter Two and his response to Slezak's arguments as described in the previous Chapter. While sketching the development of Collins' EPOR in response to TRASP explanations of the scientific development process in Chapter One, I noted that Collins divided these into two types: 'hidden-hand' and 'rational-actor'. From the previous Chapter, it should be clear that for Collins, Slezak is the very embodiment of TRASPness: his two construals of Slezak's BACON-based arguments, the 'encapsulated community' and 'scientist-mimic' interpretations, take Slezak to be claiming that BACON vindicates hidden-hand or rational-actor explanations of the scientific *discovery* process respectively. I will use this connection to explore the links between Collins' analysis of Slezak and his methodological commitments for the purposes of my critique.

Recall that hidden-hand TRASP explanations were supposed to account for the process of development of scientific knowledge by appeal to 'abstract principles of rationality' (or the like) that need not have been straightforwardly apparent to the community of scientists immediately involved in day-to-day research. Collins rejected these abstract principles of rationality because (beside the lack of agreement over what they were) they failed to account for the actual process of scientific development.[1] They were thought to be inadequate in this regard because by abstracting their principles of TRASPness from the sociohistorical context, they did violence to the actual course of scientific development in specific instances. Given the obvious analogy between the abstraction of principles of rationality from actual cases to test normative models of scientific development (as attempted by philosophers of science) and the abstraction of principles of discovery (heuristics) to test normative models of scientific discovery (as suggested by Slezak), Collins thinks that the same objections apply to BACON interpreted as reconstructing the history of scientific discoveries on the basis of abstract heuristics- it just fails to capture the complexity of the actual processes, given the diversity of the actual cases of discovery and the subsequent difficulty in identifying the 'heuristics' used in particular instances.[2]

Rational-actor TRASP explanations were rejected by Collins because of the 'open-ended' nature of the process of scientific-knowledge development, and the genesis of novel scientific-knowledge claims through a social negotiation process which inevitably includes 'nonscientific' or 'social' considerations. This was because this process involves the use of rules that are embodied and institutionalized in (social) forms of life. Since these rules are never fully articulated and their deployment is a matter of contingent social skill, appeals by individuals to explicit principles of method or rationality never suffice to explain the process by which new scientific claims are produced-they always leave something out. In the same way, the claim that the BACON heuristics are all that e.g. Kepler needed to discover his laws fails to do justice to

---

[1]. Collins has in mind here large-scale models of scientific change of the Lakatos (1978) or Laudan (1977) kind.

[2]. These reasons are bolstered by Collins' arguments against the rational-actor case- i.e. if BACON cannot even capture the process at the micro-level of individuals, how is it supposed to simulate the discoveries of a society?

whatever must have actually been involved in the discovery process.

Let me begin my critique of Collins by probing the status of his proposed alternative explanations of such situations. As a motivating question I will use the familiar dilemma: does Collins' EPOR analysis provide incompatible and alternative accounts of what really happens when a new scientific discovery is made, or does he merely try to suggest or show that TRASP explanations are inadequate when taken by themselves and need to be augmented by incorporating the 'social' aspects of the process?[3] If the former is the case, then of course Collins seems to be in the familiar relativist's quandary: his own account need be no more compelling than the TRASP ones, given that it too arguably relies on similar sorts of rules embodied or institutionalized in social forms of life- what then would guarantee its completeness or veracity? If the latter, then the claim from the TRASPers is usually that they are freely abstracting from historical or current cases to produce normative principles to guide future practice, and are hence not interested in explaining the *whole* process but only those features deemed relevant or essential from their present perspective. Collins' response is to take the first horn and insist that he is indeed giving an account meant to supplant the TRASP ones.[4] As noted in the first Chapter, he rejects the reflexive criticism by sharply demarcating the social and natural sciences, and insisting that EPOR takes a stance of epistemological relativism toward the explanations produced in the natural sciences but not to its own explanations.

How does Collins justify this demarcation and his asymmetrical stance toward the

---

[3]. This is a rehashing of the standard reflexive criticism of the Strong Programme and its variants- see Laudan (1981) and the response by Bloor (1981). In Collins' case, some different lessons can be drawn from it, however.

[4]. The tack taken by TRASPers in the second horn of the dilemma is then rejected by Collins on empirical grounds: thus he claims that even though TRASPers might claim to provide normative principles to guide future inquiry, the actual use of these principles will remain a matter of social negotiation in various ways.

epistemological status of natural and social science claims? His phenomenological and hermeneutical leanings point toward an answer. Collins believes that the two sciences have different objects of inquiry- natural scientists deal with the natural world and social scientists deal with human beings.[5] He suggests at various points that human actions are embedded within the "taken-for-granted-realities" and "ways-of-being-in-the-world"[6] of the groups to which they belong and that understanding these actions must involve a different method and be judged by different standards than the ones used by natural scientists to study the natural world.[7] The task of the sociologist is then to "take on the ways-of-being-in-the-world of different groups" and interpret them with the goal of sorting out their interrelationships.

What makes the sociologist uniquely suited to perform this task is her ability to "alternate" between these various ways of being-in-the-world and discern the interrelationships. This ability derives from the expertise which sociologists acquire through their training, and in turn justifies the non-relativistic stance taken by the sociologist toward her own explanations. Because she apprehends the 'social reality' she describes in a direct and certain way, there is no question as to its status as the correct version of events and an adequate replacement for the incomplete accounts of the TRASPers.[8] In addition, these accounts must be accepted by the misguided

---

[5]. Collins also thinks that there is something distinctive about "human beings" understood as a category that can support this distinction. I discuss Collins' views on this below.

[6] For details, see Collins and Yearley (1992)

[7]. I have called Collins' method 'phenomenological' because it involves naive and direct experiencing of the social realities of various groups, and 'hermeneutical' because their interrelationships are subsequently interpretively accounted for.

[8]. In the same way, Collins suggests that the natural scientist has (or ought to have) a non-relativistic attitude toward the natural world. Collins doesn't go as far as Bloor in claiming that social factors influence the very *content* of theories- after all the content of theories is whatever the relevant community takes them to be. He is willing to accept the results of scientific inquiry and scientists' explanations of their results insofar as they don't involve appeal to 'social' considerations, but not their explanation of the *process* by which those results were obtained, especially if it excludes

TRASPers as the correct ones because they clearly lack the expertise required to judge or appreciate their merit, just as the sociologist accepts the results of natural scientific inquiry without demurral, given the expertise of the natural scientist in this area.

This way of responding to the dilemma is, I believe, open to several objections which Collins fails to meet successfully, thereby being pushed onto the second horn. The first concerns the possibility that Collins might think the appeal to 'expertise' sufficient to ground the alternative explanations proposed by the sociologist. But the EPORer is hardly likely to persuade by claiming superiority for her explanations on the basis of her expertise and letting it go at that.

In the first place, this begs the question against hidden-hand TRASPers, who maintain that *their* explanations are designed to eliminate the influences of the negotiation process.[9] Collins therefore needs more than a reiteration of that expertise plus principled skepticism toward the possibility of hidden-hand explanations to persuade the TRASPer of her error. As noted in the first Chapter, his rejection of the 'encapsulated community' interpretation of BACON amounted to little more than a reiteration of his methodological commitments. While the idea that BACON is supposed to be a substitute for entire communities of human beings is indeed worthy of caricature, Collins need not have given it so radical an interpretation. If instead BACON were construed as a *tool* in the search for abstract heuristics to inform the making of scientific

---

social considerations.

[9]. The macro-perspective that large scale TRASP models such as Lakatos (1978) draw on are pertinent here. TRASPers point to the fact that the normative component of their models is meant to eliminate these social considerations and the models can consequently even diagnose those cases in which the failure to eliminate them has led to the wrong results. Insofar as Collins' methodology fails to provide an adequate alternative macro-perspective, there is even less reason for his alternative explanations to be accepted at face value.

discoveries pursued as an enterprise in its own right on the behalf of the scientific community,[10] Collins would have to do more than dismiss BACON as attempting to simulate an 'encapsulated community'.

Secondly, what Collins calls the 'scientist-mimic' interpretation of the Slezak's BACON argument actually challenges the very basis for Collins' alleged expertise (i.e. the distinction between human beings and the natural world), and hence again requires more than an invocation of the very same expertise in response. As I noted in the previous Chapter, Collins has attempted to justify this distinction between human beings and the (rest of the) natural world by drawing on the notions of 'action' and 'behavior'. Putting aside for the moment the sociologist's alleged special ability to interpret the actions of members of social groups, how should this distinction be understood? Collins (1990, Chs. 3 and 4) holds that only 'human beings' can *act* (while the rest of the world only exhibits behavior) because the other members of the social collectivity to which they belong have the ability to recognize what would otherwise only be behavior as action, and respond in appropriate ways. Following Winch (1958), he believes that since actions can only be *recognized* by reference to a particular social group, what counts as an action remains a matter of group consensus. Beside the circularity, this has the bizarre consequence of reifying 'cultures' by compartmentalization.[11] As a conceptual point the distinction can surely be made for specific purposes (and it is very effective when used to draw out and fortify one's intuitions in the case of BACON-type machines), but

---

[10]. The claim made by Langley et al. (1989), for example is that BACON-type programs model or simulate *features* of the 'scientific community'- specifically, the rules used by the community to get the results obtained on the basis of accepted community procedures in order to see how they might be improved.

[11]. See Whittaker (1978) for a criticism of this view.

it is very hard to sustain as an a priori proclamation predicated on the 'essential nature of human beings' (i.e. as an intensional definition). However, in the way that Collins draws the distinction it is a very general and contingent fact about human beings that they are said to 'act' and not merely 'behave', and there are borderline cases where it would be difficult to appeal to intuitions about 'human essence' to settle the issue.[12] In this connection, it must also be said that Collins' apprehensions about the dehumanization that would follow if machines came to be regarded as 'actors' seem ill-founded. Given the enormous range and diversity of human activity, it is a puzzle as to why this would be a problem 'in the large', especially if such attributions were natural in certain circumstances or met felt needs. On the whole, it is my belief that Collins cannot pin down a timeless category of 'human action' that he can identify as the sociologist's special object of inquiry. As such, the action-behavior distinction cannot underwrite an a priori demarcation of the natural and the social 'worlds' either.

Can Collins maintain nonetheless that the sociologist has special expertise in understanding and relating the actions of groups as they presently stand? He almost appears to be suggesting that the properly trained sociologist is the only competent interpreter of actions and that members of other groups lack any ability to do so. This is obviously false, and the case of the scientific community and the EPOR researcher is no exception- as Collins himself points out, scientists deliberatively use the 'social' strategies of e.g. 'creating contradiction' and 'using prior agreements' to negotiate results. Given that those who cleave to TRASP explanations can also profess to have acquaintance with the social dimension of the knowledge-production process and yet unequivocally dismiss its influence in determining the results, Collins' appeals to

---

[12]. The action-behavior distinction is surely itself open to negotiation given Collins' methodological commitments- whether or not an animal or a machine can be said to be 'acting', and the sense and extent to which it can be said to be doing so is impossible to answer generally.

expertise have the status of preaching to the converted.[13] Collins' alternative explanations must therefore rely on the cogency of his methodology and results, not on the appeal to expertise.

What of his methodology and case studies then? Let me begin with the issue of the status of Collins' explanations. Collins often seems to suggest that his goal is to merely offer detailed descriptions of the actual process of scientific knowledge development. These descriptions are supposed to be 'impartial' and provide the complete account of what occurs when a novel discovery is made or a new experiment is replicated. Collins believes that in this way the 'false picture' that science proceeds on the basis of TRASP principles that informs its public image (including perhaps the scientists' own) can be dispelled. If this is the case, then Collins is clearly doing more than mere description of the process and should recognize the normative status of his analyses. This is especially evident where his 'impartiality' is called into question by Slezak's 'refutation', which reveals that his methodology is hardly presuppositionless, and needs to be defended when called into question. In short, Collins is forced to take note of the fact that he is not 'above the fray'-in terms of his 'network model of society' he must account for his own position in the network- the rejection of reflexivity, as well as the claim to impartiality are therefore both untenable.

That said, I find several other aspects of Collins' methodology also questionable. One reason that the TRASP rejections of Collins' alternative explanations cannot be countered easily is that he *reinforces* an internal-external distinction with respect to the study of science by his sharp demarcation between the natural and social 'worlds'. Far from making science "one with

---

[13]. *Does* the sociologist have any special authority to explain the process of scientific knowledge development in virtue of her expertise? Clearly, this depends on the credence granted her ability to 'alternate' within the society. Despite himself, Collins seems to ignore the obvious- that there is no privileged vantage point for his own pronouncements either- the point here is that Collins must find ways of extending the appreciation and perceived importance of his expertise within the society.

the rest of our cultural enterprises", this global reification leads to corresponding reifications at the level of Collins' analyses of particular cases by means of the 'core-set' concept. At the global level, this might result not in alternative explanations of the same phenomena as the ones given by the TRASPers as Collins would like to think, but explanations that can be construed as disjoint from the TRASP ones altogether, and which TRASPers would be free to ignore on the grounds that their aims were different. At the level of core-sets, such reifications would lead to the attribution of 'oversocialized' group characteristics to their members to explain their different positions on the issue being examined, as suggested by Collins' metaphor that the groups have different 'taken-for-granted-realities'. The 'linkages' between the core-set members therefore become either impossible to provide, or are at least subject to challenge from competing alternatives which must then be shown to be in error.[14] These difficulties only snowball when one moves from the 'internal' linkages of the core set members to one another to their 'external' linkages to the rest of the scientific and wider societal networks. Whereas TRASP explanations attempt to provide a macro-perspective which seeks to be independent of such networks, Collins needs to provide one which incorporates these linkages directly into his explanations. If such a rigid internal-external or natural-social distinction is to be maintained at the macro-level of 'society', Collins would need to provide an alternative 'theoretical' basis (of the sort suggested by Bloor) to explain the linkages of the core-set members to the external societal network.[15] Finally, Collins believes that his methodology provides explanations which

---

[14]. In his case studies, Collins does not attempt to identify the 'taken-for-granted-realities' in any systematic way, but only eclectically. Why insist on them as a methodological tenet then?

[15]. This part of Collins' framework (Stage 3) is less developed than the others; at Stage 1, one has some idea of the sorts of explanations that are to be given from the case studies, but Collins hasn't, to the best of my knowledge, attempted to provide these 'larger links'.

do not seek to intervene in the process of science itself. But surely if the conceptual and social orders were related by 'joint entrenchment of concepts in forms of life' as he suggests, his expertise should enable him to participate in the negotiations.[16]

Collins' analysis of Slezak therefore turns out to be inadequate because his methodological assumptions prevent him from developing a substantive response- he ends up having to defend his methodology instead of investigating Cognitive Science.

(ii) Woolgar's methodology and response to Slezak

In the previous Chapter, I speculated on some of the ways in which Woolgar's methodology might be applied in a response to Slezak, and also noted the details of his actual response and its connections to his methodology. I now develop a critique of his methodology and its application to the Slezak case as portrayed there.

Let me begin by considering Woolgar's conception of the 'social' in a little greater detail. Woolgar has insisted throughout that the appropriate SSK sense of the social should be a *constitutive* one, as opposed to the factor/variable or 'social component' versions endorsed by the Edinburgh School. The difference here is supposed to be that whereas the orthodox Strong Programmers think that scientific development has a separable social dimension which can itself be an object of inquiry[17], Woolgar thinks that the 'social' cannot be definitively separated out from scientific activity itself, and his conception is therefore much more closely aligned in this

---

16. Collins' motive has been to dispel the false image of science engendered by the TRASP accounts in the realms of science pedagogy and the public image of science (Collins (1985), Postscript). These are important but limited objectives.

17. Orthodox Strong Programmers believe that this component can itself be studied 'scientifically' (i.e. by giving a naturalistic, causal analysis of social influences), whereas Collins (at least in the way I have characterized his position in the previous two Chapters), while endorsing this component sense of the social, thinks that it cannot be analyzed via causal connections, but requires the development of a suitable alternative methodology.

respect to the Actor-Network (A-N) or 'translation' theorists- for example see Callon, Law and Rip (1986)- than to Collins. However, there seems to be a significant difference between Woolgar's conception of the 'social' and theirs in the following regard. Whereas the A-N theorists see their work deriving from the French Structuralist and semiotic tradition, where the world is conceptualized as a collection of signs for the purposes of theoretical analysis, Woolgar is much more ambivalent toward such 'theoretical' treatments.

Woolgar appears to be much more an ethnomethodologist in the Schutzian tradition of directly experienced commonsense reality (and in this respect his views have considerable resonances with those of Collins as portrayed in the previous Chapter).[18] Again, while Woolgar's aim of 'interrogating the ideology of representation' has some resonances with Latour and Callon's assaults upon immutable mobiles -see Latour (1988), I do not think that Woolgar would endorse such studies as e.g. Latour's (1992) 'sociology of door closers'. The door-closer analysis is a 'theoretically' informed one, in the sense that Latour takes doors and automatic closers to be 'actants' which constrain and channel the behavior of other objects in the world including human beings. On the theory of 'actants' (as a generalized form of A-N theory), the goal is to eventually provide a macro-perspective on how things stand (even if the account is ultimately a 'fiction' and not an 'explanation') by mapping out in detail the complex and heterogenous micro-links that stabilize into macro-chains of mutually enabling and constraining actants in agonistic fields. For Woolgar, the idea of creating new representations to interrogate the ideology of representation is problematic. He is more interested in

---

[18]. While Woolgar sometimes does refer to Garfinkel's attempts to theorize the practical management of indexicality and reflexivity, he is not so much interested in a theoretical understanding of 'practical managing skills' as he is in using the notions of indexicality and reflexivity to dismantle representations. (See Collins, R. (1988), pp. 275-89 for a discussion of Garfinkel, and Lynch (1992) for a theoretical attempt to use Garfinkel's framework to analyze scientific practices).

interrogating existing representations with minimal use of representations of his own.[19] He therefore opposes his reflexive project to theoretical approaches to questions of knowledge development.

Recall that Woolgar takes acts of representation to be ideological because they potentially involve 'false consciousness': all acts of representation take for granted the preexistence of objects underlying them, whereas it is only in virtue of these acts of representation that these objects are constituted. As a consequence, the role of the 'agent' or 'self' in creating these representations is systematically ignored or discounted. Woolgar thinks that the nature of explanation is such that it forces a separation or dichotomy between subject (i.e. the one making the representation) and object (what is represented) so that the active role of the agent in constituting the object is pushed into the background. Woolgar's reflexive project seeks to highlight this 'backgrounding' by a general examination of how such a misunderstanding of the nature of representational practices continues to prevail.

The importance of interrogating the order of representation might be brought out by thinking of Woolgar's conception of representational activity as a Foucauldian 'technique of power' which potentially has long term repressive effects.[20] The creation of representations is then essentially the creation of new categories and dichotomies in institutional 'discourses'[21] which become constitutive of the discourse itself (first locally, and once stabilized by extension

[19] As I noted in the previous Chapter, his position in the Slezak case should be seen as an effort to 'interrogate the [present] order of representation' of Cognitive Science discourse.

[20]. See the discussion of Foucault on 'forms of power' in Rouse (1990), pp. 213-26.

[21] "The concept of a discourse should not be reduced to talk. Discourse also includes structures of explanation, systems of categorization, and modes of conventional practice" (Woolgar and Grint (1991), p. 377).

of the discourse itself from its institution or discipline into society at large). These representations then become the means by which people understand themselves and their actions and organize their relationships. Through many such interactions, 'reality' itself becomes the outcome of such deeply ingrained representational practices. Since the 'objective' character of the representations can be potentially used to disempower the represented and those who are obliged to take these representations at face value, Woolgar thinks that a principled skepticism toward representations becomes a necessary strategy for the purposes of redressing the imbalance: the SSKer must adopt the dictum that there is no underlying reality outside of the representations themselves for the purposes of her analyses of scientific practices.

The interrogation proceeds by means of ethnographic examination of representational practice in existing discourses each of which supports an order of representation- that is,

> several discrete classes of entities... perceptions of the character and nature of each class, the rights and obligations that accrue to each class, the similarity and distinction between classes and the nature of the differences between them... (Woolgar (1989b), p.210).

Woolgar's sense of the constitutively social therefore includes the actions of all such entities in these existing discourses, and takes as task the examination of how the order is maintained (or rather to show rather that such orders are only sustained to the extent that the 'objectivity' of certain privileged representations is not challenged).[22] In the case of Cognitive Science, Woolgar notes that its discourse supports a particularly malignant order of representation, providing as it does a locus for the determination of the natures of such entities as 'human beings' and 'machines'. Woolgar maintains that the actions of the cognitive scientists, their

---

[22]. Unlike Collins, who might be interpreted as saying that what Woolgar calls the 'present order of representation' is the set of rules relating concepts embodied in institutional forms of life, Woolgar isn't concerned to provide a 'model' of it.

sociological critics and the machines and programs involved in this discourse are all responsible for maintaining its associated order of representation. He will therefore have no truck with either side in the dispute, preferring instead to engage in an examination of how the order itself is sustained.

I will criticize Woolgar's methodology and application in the response to Slezak by questioning some aspects of the interrogation of representational practices qua project and the implications of the interrogation in the Slezak case. Consider Woolgar's insistence on a principled skepticism with respect to representations as a heuristic device for getting at the fundamentals of representational practice, and his tactic of always taking existing dichotomies as topic for the purposes of understanding the nature of the construction of representations. As Woolgar realizes, there is no 'neutral-observer' standpoint from which such a project could be carried out, as he himself ends up somehow 'representing' the process that he is examining, thereby asserting his own authority over others to the extent that they are compelled to take his characterizations of the process as revealing of some 'underlying social reality'. That is, his own project could be taken as the creation of a reified object of inquiry (viz. 'the nature of representational practices') and his recommendations about the modification of existing representational practices might be taken to be subtly coercive forms of imposition of authority. Such a prospect is sufficiently worrying for Woolgar that he attempts to avoid the deleterious consequences of his own representations by attempting to interrogate or deconstruct his own accounts in the very act of presenting them.

What is questionable about this however is that Woolgar's 'principled skepticism' about representations inevitably becomes a 'selective skepticism' in his reflexive examination of his own accounts of representational practices. Obviously, any attempt to characterize the 'reflexive position'- e.g. Woolgar (1988b)- is to (re)present it in opposition to other positions, and any attempt to criticize other viewpoints requires that they be (re)presented for these purposes. Why then should Woolgar's characterizations be taken to be unproblematic just because he has undertaken to interrogate them himself? Woolgar's response to such criticism is usually that such problems arise because of the conventions of academic discourse that force the issue to be conceptualized in forms of 'positions' and 'debates'- deeply ingrained practices constituting the very ideology of representation that he is concerned to question. Woolgar looks forward to the day when such constraining conventions will no longer be necessary. What is odd is that while Woolgar wishes to blur the distinctions between representation and object and win recognition for the occasioned and situated character of the construction of their relationships, he feels it necessary to pursue his study of representational practices as a self-contained project that seeks to find alternative forms of representation by distancing himself from actual representational practices.

As an illustration, take Woolgar's call to other SSKers to transcend the debate with Slezak and concentrate instead on an examination of the representations sustaining the discourse, which runs something like this: 'we must recognize our role in perpetuating the ideology of representation and develop reflexive ways of interrogating our representational practices even as we engage in them'. The question, however, is: who is the mythical 'we' here that must reconsider our representational practices? After all, there is no 'we' to be found except in the context of 'our' occasioned actions in the discourse. How, then, can one transcend the debate and concentrate instead on the way that actions are constructed within it?[23] This suggests that the very distinction between pursuing 'reflexivity as a project in its own right' and 'reflexivity as critical self understanding in the course of the pursuit of other ends' should be collapsed.[24] Another way of putting the point is to note that Woolgar's project of interrogating representational practice seems to result in judgments of adequacy of the representational practices of various discourses without sustained engagement. The worrying thing here is that the professed 'principled' skepticism and the actual 'occasioned' skepticism of Woolgar's reflexive project might lead to arbitrary criticisms of other representational practices.[25]

To elaborate further with respect to the Cognitive Science case, Woolgar worries that such representations as 'human intelligence as essentially a decision-procedural mechanism' or machines as 'essentially asocial entities' which get constructed within the discourse, and are only 'situated' and 'occasioned' characterizations might end up affording special authority to Cognitive Scientists, who continue to

> define their nature and character..., establish that these are proper objects for investigation and claim to be uniquely competent in speaking on behalf of these objects. The rest of us have to defer to what these privileged spokesmen have to say (Woolgar (1985), pp. 565-6).

Woolgar's particular way of responding to this is not to enter into the debate itself but to step

---

[23]. Woolgar has criticized SSK because "sociological analyses of scientific knowledge need to constitute themselves (in the course of argument) as harder than scientific knowledge since Self (the analyst) has to assume a greater degree of hardness than scientific knowledge in order to effect its deconstruction". I think that Woolgar's deconstruction/interrogation of 'representational practices' is all of a piece with this- the analyst of representational practices (Woolgar) has to assume a greater degree of hardness than 'representation' in order to interrogate it.

[24]. In this regard, see Fuhrman and Oehler (1986), who criticize the narrow conception of reflexivity within Discourse Analysis, which Woolgar seems to have taken over.

[25]. The problem here is not with the 'abstraction' involved in characterizing representational practices (which is unobjectionable) but with the assumed homogeneity of the object of inquiry, 'representational practices'.

back and take as topic or "critical target our own ability to construct objectivities through representation" (Woolgar (1988a), p. 93). But surely, Woolgar is not suggesting a decontextualized investigation of representational practices drawn from different discourses while denying their occasioned character! Again it appears that interrogation of the representations produced must proceed by engaging specific ones and not by pursuing the reflexive project as an end in itself. In other words, while Woolgar feels free to charge e.g. such sociological critics of cognitivism such as Collins and Coulter of 'tacit sociological essentialism', he himself is apparently engaged in an examination of the 'essentials' of representational practice. Yet, his disclaimer is that his goal is not to explain representational practice but only to interrogate it. But it is hard to see this interrogation as being an effective one when attempted in an insular fashion from existing representational practices.

Yet another way of putting this problematic aspect of Woolgar's methodology as applied in the response to Slezak is as follows. Woolgar is quick to point out that Cognitive Scientists are uncritical perpetuators of the ideology of representation and critics like Collins are only successful in reinforcing this ideology by producing representations of their own in order to challenge them. But if the 'objectivity' of representations arises solely because of the privileged status of their creators and lack of resistance offered to their constructions resulting in unwanted imposition of authority, then surely the point would be to 'offer resistance' in exactly the manner that Collins and Coulter have attempted. Standing back from the debate itself to interrogate representations and diagnose flaws of current representational practice to design alternatives could lead to ones that might be marred by the very lack of resistance incurred in creating them. Instead of actually changing any current representational practices, Woolgar runs the risk of ending up with yet another reified object of inquiry (viz. 'representational practices') on his hands and a bunch of unrealistic proposals for change in specific situations.

Consider some of the other features of Woolgar's response to the Cognitive Science enterprise. Woolgar is ready to identify the rhetoric of progress underlying Slezak's pronouncements about the successes of AI as uncritical acceptance of the ideology of representation and a failure to recognize (or deliberate deception regarding) the historicity of the disparate collection of 'achievements' that are retrospectively identified as the successes of Cognitive Science or AI research. And yet, he is ready to dismiss his own reconstructions of the 'progress' made in the social studies of science from its beginnings with Merton to his own constitutive position as an irony which shouldn't be taken too seriously. But no doubt what is not a matter of irony where Woolgar is concerned is that he does think that his criticism of SSK leads to an alternative position which is in some sense better than traditional SSK analyses. This simply suggests that from the point of view of changing current representational practices, it is not enough to merely document the situated and occasioned use of representations within them without allowing for the interrogation of one's own documentation by others.

Another aspect of Woolgar's strategy of interrogating the representations of Cognitive Science- that of granting agency to the machines involved in the discourse[26]- is also objectionable in this regard. The question of whether the machines in question can be granted agency is surely an empirical one that must be capable of settlement by means acceptable to both the SSKer and the Cognitive Scientist. Otherwise Woolgar simply assumes the role of representative for the machines and the agency question again becomes a matter of a clash between Woolgar's representations of the machines (based on his ethnographic study) on the

[26]. Woolgar has even undertaken an 'ethnography of computers' in order to accomplish this.

one hand and the Cognitive Scientists' on the other, despite Woolgar's claims to have merely examined the construction of the representations of the Cognitive Scientist without actually passing judgment on the character of the representations themselves. Such a proposal for using 'empirical' means to settle the agency issue would be no doubt unacceptable to Woolgar given his anti-naturalistic bent, but I submit that the alternative is simply to perpetuate another dichotomy between the scientific and sociological approaches to the study of science in the name of dissolving all such distinctions. Woolgar thinks that such appeals to 'experimental evidence' or 'method' inevitably set up an aura of objectivity that puts distance between the representations produced and those responsible for their creation, or that objectivity' can be used to ill-effect. Skepticism regarding these 'objective' assessments claiming independence of their active constitution by specific agents is supposed to be an antidote to just such a possibility. As I noted above however, such principled skepticism must give way to an occasioned skepticism in one's own case in practice, and such occasioned skepticism must be deployed in actual contexts of investigation.[27] While Woolgar thinks that a naturalistic approach in such investigation would inevitably be pernicious, I would suggest that there is no need to prejudge the issue on this score or unreflectively endorse the distinction between naturalistic and social approaches in offering alternative accounts of the representational practices of the discourse being examined.[28] The better way to interrogate representations

[27]. Woolgar could always freely admit after the fact if criticized that his own accounts were also situated and occasioned ones and should be treated as fictions. In confronting the role of the agent in the creation of representations, 'Self'-refutation becomes an epistemic virtue! (But not when used in the arbitrary way Woolgar seems to be proposing).

[28]. The fear that a 'scientific' examination of scientific practice would be a de facto acceptance of the ideology of representation or preclude a constitutively social view of science seems misplaced. There are no special reflexive problems with 'naturalistic' approaches to science, and nor do these deny that the representations that would result would be anything other than 'occasioned' or 'situated'. See Barker (1989) on the possibility of a psychology of science.

would be by holding their originators accountable (and taking responsibility for one's own). That way, the 'objectivity' of the representations could be firmly tied to their creators without precluding naturalistic means of assessing them. The appropriate place for intervention would then seem to be in the process of creation of the representations themselves.

Finally, the sort of presuppositionless inquiry suggested by Woolgar's ethnographic method must be questioned to the extent that it precludes more 'theoretical' approaches to the study of scientific practices.[29] Woolgar has stressed throughout the need to interrogate existing representational practices rather than to explain them. But as I have noted above, the distinction between interrogating 'representation' as a project in its own right and the examination of the use of specific representations cannot be sustained indefinitely, unless Woolgar doesn't see his project as having any consequences ultimately for actual representational practice in the discourses that he examines. The virtue of a 'theoretical' approach in the latter instance is that it helps make presuppositions explicit (or rather results in their joint construction) so that the recommendations made by the analyst of scientific practices can be evaluated for their adequacy and incorporated into existing practice. To take the example of Coulter's critique of Cognitive Science research, Woolgar thinks that Coulter's specific criticisms merely provide "new facets of 'intelligence' for further axiomatization" because he is pessimistic about the prospects of Coulter succeeding in the long term in having any effect on Cognitive Science practice. But the construction of new forms of representations does have the effect of transforming existing practices. The development of 'new literary forms' and 'reflexive ethnography', for example, it is hoped will transform SSK or anthropological practice without

---

[29]. At least, his methodological edict of never engaging in any existing discourses but taking each of them as topic in order to find out how they are sustained seems to suggest this.

at the same time providing alternative 'foundations' for them. Why then might Coulter's efforts to transform Cognitive Science practice not do the same? This is yet another instance of Woolgar's selective skepticism in action.

Along these lines, consider Woolgar's criticisms of the work of Rob Kling (Kling (1991), Woolgar and Grint (1991) is the response). Kling's attempts to gauge the 'effects of computer technology on specific social relationships'[30] is challenged by Woolgar because such efforts take as unproblematic the existence of an object such as 'computer technology' without realizing the essentially indeterminate nature of such technology, and the fact that "its nature and capacity arises through the discourse of which it is a part" (Woolgar and Grint (1991), p. 374). Thus instead of taking the computer as a well-bounded technological artifact and then examining its 'impact' on society (which means drawing a distinction between 'technology' and 'society') Woolgar suggests that sociological analysis concentrate on wider-ranging topics such as using the computer as a means to interrogate the construction of representations of our basic assumptions about human nature (given that the computer is inextricably bound up with our idea of ourselves). But again it would be misleading of Woolgar to claim that his investigation of 'how the nature of objects such as computers gets constituted in practice' doesn't itself make *some* assumptions about the boundedness of the entities (and hence their representation) that he is investigating, or to deny that his investigations too have implications for the reconceptualization of human nature. No doubt Woolgar would take Kling's research also to be 'topic' rather than 'resource' in his own investigation into 'the construction of representations

---

[30]. To talk about the 'social' significance of AI research in quite another sense, Kling identified the reluctance of AI researchers to allow assessment of the 'social implications' of their research because of their substantial military funding.

involving basic assumptions about human nature'. But one finally realizes what Woolgar's own account of 'the construction of representations' eventually turns out to be, namely 'topic' for a reflexive examination of his own construction of representations, and so on ad infinitum.

Woolgar's analysis of Slezak therefore turns out to be inadequate because he doesn't actually engage in the debate, and while he urges an interrogation of CS representational practices as an alternative to the debate, he does not actually carry one out.

Chapter Four: An Alternative Analysis of Slezak

Since this penultimate Chapter recapitulates the previous ones, is long, and is drawn upon in the final one, I begin by briefly setting out the structure of its argument. I begin by summarizing and extending the analysis of the previous Chapter by indicating how my response to Slezak differs from Collins' and Woolgar's in the first section. Here I conclude that 'models' or theories are to be considered as conventional modes of representation having determinate applications in the contexts of activity in which they are used. This in turn makes these models 'social' in the sense that it is their use in various activities that gives them their significance.

In the following two sections, I consider the issue of CS as a coherent research programme, for as I noted in the first section of Chapter Three, Slezak's 'empirical refutation' of SSK relies not merely on the computer models of scientific discovery that he cites taken by themselves, but on these models of discovery considered as part of the larger research programme of Cognitive Science. I examine the use of the 'Turing machine model' which Slezak invokes in each of the areas identified by him. I draw on a number of sources to argue that the allegedly coherent research programme turns out to be a motley collection of activities that are at best tenuously related to one another. As a consequence, I argue that Slezak's claims to a decisive refutation of the social views of scientific discovery cannot be sustained. On the contrary, it becomes clear that if anything, it is the 'social' considerations of *use* of models such as the Turing machine model in seemingly disparate contexts which would lend to its overall credibility than anything else.

Having shown that the Turing-machine model is not as extensively implicated in the various 'subdisciplines' of CS invoked by Slezak as he claims, I go on to consider, in the final section, the extent to which its use within the subdiscipline of AI research considered independently could serve to refute the social views of discovery. Here too, my contention is that what Slezak identifies as 'abstract heuristics' of discovery implemented in various programs, like the Turing machine model itself, are also rules used in specific social contexts and for specific purposes, and hence not 'asocial' at all. I conclude from this that if a BACON-like system were taken to be making discoveries (on the basis of such heuristics), then such a system could well be itself regarded as a social being rather than as 'proving' that discovery is essentially an asocial process.

(i) My approach versus Collins' and Woolgar's:

Slezak has presented 'Cognitive Science' as a project informed by work in philosophy, psychology, AI and neuroscience, unified by a shared conceptual model (or theory) of cognition (the Turing machine model) whose goal is the naturalization of epistemology. However, those who favor social and contextual construals of scientific knowledge and activities (from Strong Programmers like Bloor to constructivists like Woolgar to philosophers of Science like Rouse, to repeat some already mentioned names) are united by their suspicion that 'abstract models' or 'theories' are much more than tools- 'theories' are themselves to be understood as conventions having various uses within scientific activity, and not 'real entities' (linguistic or otherwise) located anywhere from a platonic 'third world' to within the heads of individuals.[1] I also take 'theories' to be conventional representations expressed in different ways and for varied ends, none of which allows for their independent existence from the activities where they are formulated and used.

[1]. While this seems to suggest that Slezak has lined up a 'unified' opposition as his target, it is apparent that a characterization 'SSK position' is as elusive as that of Cognitive Science turns out to be (see below).

On the face of it, this seems to be a direct subscription to Woolgar's notion of representations being inextricably embedded in the discourses in which they occur insofar as 'theories' or 'abstract models' are considered to be representations. However, there are some significant divergences between Woolgar's discourse-and-representations view of theories (if it is indeed such) and mine. Woolgar's conception of representations-within-discourses and his insistence that 'there is nothing beyond the representations themselves' seems to make the material and 'social' (by which I mean relational and communicative) aspects of 'discourses' parasitic upon the concept of 'representation', and Woolgar seems unsympathetic to the view that particular aspects of discourses might be separable from the representational acts within them for the purposes of analysis.

On Woolgar's view, it seems to follow that the analyst is methodologically constrained to limit herself to an examination of the representations created and maintained within existing discourses. I believe, however, that the analyst is not necessarily obliged to examine (at least the material and social aspects if not the conceptual ones of) the 'discourses' under consideration on the basis of the representations of the practitioners, but is free to conceptualize them on the basis of resources drawn from elsewhere as well.[2] There is another related feature of Woolgar's view of discourses that I would question. Woolgar takes all acts of representation within the discourse under examination to be acts of deliberative distancing by the agent of her Self from the 'object' represented.[3] Woolgar overlooks (or perhaps intentionally chooses not to make

---

2. Nor does this require the analyst to deny the 'constitutively social' nature of either the practices under consideration or her own activities. What would be furhter required here, however would be some sort of reconciliation between the divergent views of practices by a mutual 'translation' process.

3. Which effects the impression that the agent is not responsible for the character of the representation produced- it attains 'objectivity' and 'independence'.

explicit) the possibility that a large proportion of the representational activities of any discourse are unpremeditated, so that their 'deconstruction' amounts to the creation of alternative representations where (perhaps) none had existed before. In that case the radical constructivists' reconstructions of 'existing' ways of representing clearly have normative ramifications, whether Woolgar is willing to admit this or not. Consequently, there is need for explicit recognition of this aspect of the analyst's activity.

In any event, Slezak's assertion that 'Cognitive Science' is unified by a shared model of cognition is to be interpreted not as suggesting that there is an 'abstract model' that all Cognitive Scientists abide by, but rather that a common representational convention is implicated in the activities of the practitioners of each of the disciplines that Slezak sees as contributing to the overall project of CS. The first question that arises then is one of the extent and degree to which 'Cognitive Science' is indeed committed to such a common representational convention in what appear prima facie to be a diverse set of practices. To put the question in terms of Collins' framework (cf. Chapters 2 and 3),[4] to what extent is the 'network of Cognitive Science research' stabilized by the institutionalization and entrenchment of the 'Turing machine' concept in the 'forms of life' of CS researchers?

The first difference, between Collins' (purely descriptive) approach and the more normative one I would suggest is as follows. By taking the 'Turing machine' concept to be institutionalized in the Cognitive Science 'form of life' on the authority of practitioners, a Collinsian analysis would grant at the outset perhaps that there was some 'link' between the variety of enterprises that Slezak identified as belonging to the 'Cognitive Science research programme'. Collins

[4]. I do this in order to point out the differences between Collins' approach and mine.

would perhaps take adherence to the 'Turing machine concept' to be part of the Cognitive Scientists' 'way of being-in-the-world' and responsible for their 'beliefs' and actions in the somewhat insulated institution of Cognitive Science research. I would suggest that such imputation of common representational conventions prior to investigation by the analyst might potentially prejudge their status and signigicance. On this point, I concur with Woolgar (in a sense) on the need to establish some 'anthropological distance' and not explain the activities of the tribe under investigation in terms of their own categories (see below). The analyst therefore assumes the heterogeneity of what are claimed to be activities informed by the 'same' model of cognition (and as a working assumption expects to find differences in the use of the model in different circumstances).

Secondly (again, I hazard a guess) Collins would perhaps take developments within Cognitive Science (say by the settlement of current controversies at least partially on the basis of commonly accepted current concepts and procedures) to proceed by means of a negotiation process where the 'open-ended' concepts in use by the CS community (including the 'Turing machine' concept) were extended in various ways for use in novel contexts. Collins would then take his task to be that of the description of the changes in the overall network of CS brought about by these extensions. Contrary to this, I would urge the kind of constructivist perspective where anything worthy of the label 'concept in use' would not be considered 'open-ended' but would be presumed to be quite determinate and as having explicit associated criteria for use, and implicated in quite definite ways in the CS 'form of life'. Any change in such concepts would then essentially result in a *new* 'form of life', the relation to its predecessor of which need not be one of 'straightforward extension'- it could implicate material and social features and considerations of a significantly different character even if the 'conceptual change' involved could be said to be a simple one, or vice versa).[5]

Thirdly, an adequate analysis would attempt to offer less impressionistic accounts of the relationships between scientific practices than the ones that seem to follow from Collins' model and methodology. I noted earlier that Collins' methodology reinforced the internal-external distinction by separating out the natural and social worlds at the global level, and then by setting up several mini internal-external distinctions at the level of the core-set in specific cases of scientific knowledge development, where the agents involved in (say) the settlement of a controversy and hence the modification of the overall 'network of concepts' are each said to have greater knowledge of 'their own parts' of the network, and differing perceptions of their actual position within the network. Collins' suggestion is that the sociologist has the best overall picture of the network, and is therefore capable of providing the most complete explanation of the changes within it.[6] However the accounts of these changes are then necessarily given ex post facto and they are largely narrative in form (see e.g. Collins (1985), Chapter 4). Consequently, as I pointed out earlier, they are not convincing alternatives to those who also accept a sharp natural-social or internal-external distinction but give more credence to the natural world in determining the course of scientific knowledge development.

By contrast, an adequate analysis would reject such distinctions, especially at the level of 'natural versus social worlds'. Insofar as the analyst provides a 'network' model of the the CS research programme and its relationships to other networks 'outside' CS, it should be

---

[5]. Clearly, existing 'forms of life' which Slezak takes to be informed by the 'same' abstract model could also turn out to be quite different upon closer scrutiny in this regard.

[6]. Even though Collins is willing to profess more or less complete ignorance with regard to the 'conceptual' network of science following his demarcation of the natural and social worlds and location of his own work within the social, he nonetheless holds that he has a better overall picture of the network because the 'conceptual order' is subsumed by the 'social order' and that changes in the former are reflected in the latter which he has access to.

interpreted not as a depiction of a real state of affairs with respect to the 'social order' of CS, but as an 'instrumental' one, the role of which would be to help the analyst in the assessment of the significance of these practices for each other. In the spirit of Collins' suggestion that 'changes in the conceptual order are at once changes in the social order', the interrelationships between network nodes are taken to be coterminously scientific as well as social, but for this very reason there is no need to separate them out or label them as such. Furthermore, to the extent that such interrelationships are postulated by the analyst, the nature of these relationships would have to be explicitly indicated and argued for on the basis of appropriate evidence when put into question. This would certainly not preclude the use of theoretical resources when understood as conventional modes of representation. Let me turn to the details of the alternative analysis of Slezak on this note.

(ii) CS as the naturalization of epistemology

Slezak has argued that CS should be understood as part of the project of 'naturalizing epistemology'.[7] Broadly speaking, the call to naturalize epistemology has been interpreted as urging a move from an autonomous and purely conceptual inquiry into questions such as 'What is knowledge?' and 'How ought we arrive at our beliefs' to one which is informed by the results of investigations into how in fact beliefs and knowledge are actually arrived at (Kornblith (1985), Chapter 1). Variants on the naturalization of epistemology run from Quine's proposal that epistemology be replaced by a descriptive empirical psychology to others that contend that while epistemology must be informed by empirical, scientific work, its task is still ultimately to prescribe norms for arriving at beliefs that in some sense go beyond the empirical results

[7]. See Kornblith (1985) for an overview of the general project and some positions taken on the central issues by analytic philosophers.

(which might be useful for testing or constraining epistemological theorizing, but are certainly not its goal). How does Slezak construe CS as naturalizing epistemology? It is clear that CS is at least an attempt to naturalize epistemology in the descriptive empirical sense because it hypothesizes a mechanism for explaining the acquisition of beliefs (again the Turing machine or computational model). One way that CS 'naturalizes' epistemology therefore, is by empirically ascertaining how the process of belief acquisition actually works (for human beings) on the basis of the model. At the same time, Slezak sees the CS naturalization of epistemology as also having a normative dimension. This is apparent from Slezak's use of computer models of scientific discovery to demonstrate that scientific discovery is guided by heuristics or methodological rules applied by individuals, so that while all beliefs are acquired in the same way, the ones that properly count as knowledge ought be acquired on the basis of appropriate heuristics (which are also recovered by the modelling process). It therefore appears that Slezak takes the CS naturalization of epistemology to proceed in the latter of the above ways. I will argue in this section, though, that there is little reason to think that Slezak's naturalization of epistemology in this way amounts to a decisive refutation of what he calls 'socially contaminated' views of knowledge production.[8]

I will first try to show that generally 'social' considerations cannot be regarded as extraneous to the naturalization of epistemology, and then go on to indicate how such a naturalization might proceed. In order to do this, I will use David Bloor's response to Slezak (Bloor (1991), pp. 167-70) to establish the first point and then indicate how my conception of

---

8. It will also become clear in the next section that even within the various subdisciplines that Slezak takes to be contributing to the overall Cognitive Science project, there is little agreement over how epistemology is to be naturalized, or even if this is the eventual goal of these subdisciplines.

naturalized epistemology would differ from Bloor's version. Bloor is all for a naturalized epistemology, but unlike Slezak he doesn't see SSK and CS as opposed positions in this endeavour, but as complementary ones. Bloor suggests that CS (naturalistically) studies the processes of belief/knoweldge acquisition in the individual. He calls this the study of the 'background natural rationality' that is generally taken for granted by SSK analyses, so that CS becomes a background theory of individual reasoning capacities. SSK then (naturalistically) attempts to understand the process of "how a collective representation of the world is constituted out of individual representations" (Bloor (1991), p. 169). Clearly, then, SSK would benefit if a naturalistic (causal) theory of the processes occuring in individuals was developed within CS because such an account would help the SSKer to assess the import of of individual reasoning in interactive knowledge development- but a complete account would require the sociologist.

Slezak would be right to distrust Bloor on this issue however, because the SSKer is the one who determines, in the final analysis, to what extent 'social factors' are causally responsible for even the theories of individual belief acquisition provided by CS. Bloor's version of naturalized epistemology calls for an overarching 'social theory of knowledge' which would therefore ultimately causally explain the acquisition of beliefs by individuals as well as the certification of knowledge by communities.[9] I will elaborate my differences with Bloor on this issue below, but let me note here how Bloor's response indeed suffices to show that a naturalized epistemology need not be one that is insulated from a consideration of social factors. Bloor suggests that CS is a naturalistic account of belief acquisition in individuals 'in the state of nature', and proceeds to add the social dimension involving the interaction of individuals by

[9]. While reflexively accounting for itself in the process.

means of his social theory. The implication really, is that the 'individual in the state of nature' is a myth- Bloor drives a wedge between the individual and the social aspects of belief acquisition to grant the cognitive scientist a separate domain of inquiry and then takes it back by proceeding to subsume it into his more comprehensive account.

The key point to note here with respect to Slezak's version of naturalized epistemology, however, is that Bloor makes it clear how Slezak fails to preclude 'social considerations' from having a role in the belief acquisition process. The descriptive part of Slezak's account, which identifies the belief acquisition process in the individual does not contribute to the determination of what specific beliefs the individual comes to have. The normative part seeks to identify the heuristics that should be utilized by the processes in the descriptive part to reproduce paradigmatic cases of knowledge. Bloor neatly inserts the 'social' in between these two parts by claiming that what count as paradigmatic cases of knowledge are the upshot of interactions between individuals and these must be taken into consideration in providing a naturalized epistemology.[10] It then becomes quite clear that the abstract model of cognition that is supposed to unify CS itself does no work in supporting the normative part of Slezak's naturalized epistemology, as the normativity comes from elsewhere- it is derived from the nature of the heuristics, and not the model itself. That the heuristics themselves are ones we ought to use *because* they are socially uncontaminated remains to be established, and indeed would seem to require justification on independent grounds;[11] Bloor's contention that the

---

[10]. It seems to me that Bloor need not take a stand on the whether the computational model is descriptively adequate or whether the computer itself could be said to be a social being to say this much.

[11]. I will consider corroborating evidence for the heuristics from the various subdisciplines of CS that could reestablish their link with the abstract model in the next section; and the possiblity of establishing the independent existence of such normative heuristics untainted by any social factors whatsoever in the one following.

heuristics themselves are the upshot of conventional agreement, far from being decisively refuted at this stage, is unscathed by the invocation of the computational model alone.

In order to illustrate how I think a naturalized epistemology should obtain, let me take issue with Bloor on three features of his 'social theory of knowledge'. Firstly, I accept Woolgar's criticism of Bloor's factor/variable notion of the 'social' (typically 'interests') and 'natural' components (his distinction) of knowledge production. Bloor takes each of these factors to provide a causal contribution to the overall production of scientific knowledge (the causally determined entities are usually 'theories' or 'beliefs'). The 'naturalized epistemology' that results is usually underwritten by a single overarching 'theory'. Woolgar's criticism might then be construed as a demarcation problem- how can Bloor identify those beliefs/actions of scientists which are informed (caused) by their social interests from those that are informed by natural causes? Woolgar's countersuggestion is that the analyst must restrict herself to merely examining how, in particular cases, the scientists themselves create and sustain these distinctions by engaging in ethnographic study. I would stake out a middle ground between these two extremes. *Contra* Bloor, I would suggest that a single theory cannot explain scientific 'beliefs'- indeed, beliefs alone are the wrong sort of thing for the analyst to concentrate on in her examination of scientific knowledge production- this should be broadened to include the conceptual, material and social aspects of the scientists' activities. Once one moves from 'theories' or 'beliefs' to this broader conception, the single theory account becomes hard to sustain, and the theoretical and methodological resources that one deploys for the purposes of analysis (for *pace* Woolgar, I do not think that all the analyst can do is describe) need to be tailored to the 'situated' nature of the activities under consideration.

I certainly do not want to suggest that the analyst cannot use 'interests' to explain aspects of scientific activity, for at the very least descriptive ethnographic studies substantiate the plausibility of such ascription. But while the analyst must be reflexive here (in the sense of being willing to accept examination of her own accounts), I do not think that this aspect of the analysis need necessarily aspire to strictly 'causal' status, and this is my second point of difference from Bloor.[12] For example, what would count as a causal explanation of the 'belief' of cognitive scientists that the Turing machine model was appropriate for investigating cognitive processes? We might usefully give a historical account of how the model came to be utilized for these purposes, but taking this historical account to be a causal one doesn't appear to add anything to the explanation or help in assessing the significance of current CS practice.

This brings me to the third problem with Bloor's naturalization of epistemology. Bloor seems to take his account of scientific knowledge to be of the purely descriptive naturalistic variety. I would urge a more normative conception. Bloor's goal seems to be to provide evidence (largely on the basis of historical case studies) for his thesis that social as well as natural factors must be regarded as constitutive of scientific knowledge production. We might take Bloor's case studies as underlining the fact that attention must be paid to the social dimensions of current scientific practice as well. At their best, the theoretical and conceptual resources developed by Bloor[13] would then perhaps help in conceptualizing and describing aspects of current scientific practices. The point, however is that such resources must be deployed in specific situations, and the consequences of such analysis could well be specific

[12]. I don't see how such a causal connection could be consistently made. Bloor ((1991), p. 166) suggests that "apparent exceptions to covariance and causality [by the social interest component] may be merely the result of other natural causes apart from social ones" which are nonetheless present in the background-but short of the actual identification of particular interests in specific cases, this seems to be a 'mere' conceptual point.

[13]. By which collective representations are created out of individual ones (e.g. his grid-group method-see Bloor (1983)- of analysis of the social organization of 'language-games').

proposals for change- the theories employed would be resources for analysis, and the adequacy of the analysis would hinge on the applicability of the theoretical resources to specific situations, not on some preselected theory.

Thus on my view, a 'naturalized epistemology' would involve the prescriptions for changing specific activities in the course of analysing them, and not for the acquisition of beliefs in some general sense. One might well ask to what extent this would be a naturalization of epistemology, which is usually understood as seeking a general 'scientific' determination of what should count as justified true beliefs. My response would be that what count as 'justified true beliefs' are determinate only in the context of the concrete situations analysed. Beyond this, the analysis need not be 'nonscientific' in any programmatic sense (without trying to provide a single global theory).

(iii) CS as a unified research programme

In this section, I consider the issue of the extent to which CS can be understood as a coherent research programme unified by the computational model of cognition along two lines. I look at each of the subdisciplines that Slezak sees as contributing to the overall CS project to determine, firstly to what extent use is made of the 'computational model' in them; and secondly to what extent 'corroborating evidence' can be plausibly seen to accrue to both the computational model and the abstract heursitics from them.

Let me begin by noting just how disparate CS appears as a field. Fuller (1992b, Chap. 5) notes that the group of people originally credited with launching the 'Cognitive Revolution' against behaviorism comprised a "veritable Chinese encyclopedia... [hence] how might one characterize- not to mention explain- the relevant sense of resemblance between... [their various projects]?" Slezak's candidate for such an explanation is the Turing machine model. To what extent can this model be said to have guided (and continue to guide) the work in each of the subdisciplines identified by Slezak? Fuller's use of Bloor's grid-group method of analysis of the 'cultural cartography' of CS to illustrate that those who pursue the 'cognitive' as an object of study tend to conceptualize it in vastly differing ways suggests that common use of such a model is highly implausible. I now examine in turn the extent to which the model is utilized in each of the subdisciplines invoked by Slezak.[14]

(a) Philosophy: Slezak cites the work of Thagard (1988) and Holland et al. (1986) as that of philosophers utilizing the computational model to explain the process of scientific knowledge production. But on the one hand, as Thagard (1989) makes clear in his response to Slezak, even he is not willing to totally discount the importance of 'social factors' influencing cognitive ones in the scientific discovery process. Furthermore, Thagard notes that his project is intended to account for the (internal) course of scientific development as it actually occured:

> the BACON program is ahistorical... showing how scientific developments *might* have come about falls short of explaining how they actually did (Thagard (1989), p. 654).

It is apparent that Thagard is more interested in a descriptive-historical rather than a normative account of scientific discovery, and it appears that he takes the 'computational model' as a tool for this purpose rather than as a literal description of actual human cognitive processes and mechanisms.[15]

Slezak then goes on to invoke, in a review of Collins (1990) among others, the names of Stich and Dennett as philosophers who have informed views of CS. But to reiterate a point

---

[14]. I consider philosophy, linguistics, neuroscience and psychology in this section and AI in the next, where I look at the independent plausibility of the 'abstract, socially uncontaminated heuristics'.

[15]. See also Giere (1989) for a similar view of 'computer modelling', though Giere appears to give more credence to the 'interests' of individual scientists in the causal determination of their beliefs than Thagard is.

made by Fuller (1992), neither of these philosophers seem to have much confidence in the 'computational model of mind' that he subscribes to. Thus Ramsey, Stich and Garan (1990) call for the elimination of folk-psychological concepts such as 'beliefs' from a mature CS, and also (guardedly) question the usefulness of a 'Turing machine model' of mental processing, preferring a 'connectionist' model instead.[16] Similarly, Dennett argues against a 'computational functionalist' account of cognition (of the sort that Slezak seems to be proposing) for the reason that it fails to account for the 'intentionality' of systems that can be said to have beliefs (i.e. the fact that their beliefs are about something, cf. the discussion of Dennett's views in Bechtel (1988), pp. 70-78). As some of the very philosophers Slezak cites seem to more or less explicitly reject the 'computational model', it can hardly be said to be the unifying basis for philosophical ruminations on the nature of cognition. Equally, no incontrovertible evidence seems to accrue to the model from work in this area.[17]

(b) Linguistics: Slezak takes the work of Chomsky and his followers as both being informed by a computational model of knowledge acquisition (in this case language acquisition) as well as providing evidence for it. The two questions that arise here concern the relationship of Chomskyan linguistics to the sort of CS research that Slezak invokes in his refutation of SSK, and its relationship to other work in linguistics. With regard to the first question, it is by no means clear what the conceptual, material and social contexts of convergence between CS and Linguisitics are, and Slezak doesn't offer any clues. Perhaps the direct link between Chomsky's

[16]. In the 'connectionist' model, particular beliefs cannot be individuated in a network of 'neural activation', and indeed the processing following 'input signals' resulting in 'output signals' itself becomes intractable to algorithmic determination. The status of connectionism is itself the subject of considerable disagreement.

[17]. Philosophers might be construed as providing largely 'conceptual resources' to CS. A survey of the diversity of positions here suggests that taking CS to an activity unified by its reliance on the computational model would be a mischaracterization.

lingusitics and the 'computational model' can be made on the basis of Chomsky's claim that human beings have an innate, 'psychologically real' linguistic competence resulting from a special 'language organ' (Chomsky (1980), p. 229). Then Slezak could take this organ to be a computational mechanism (with similar parallels being drawn between language acquisition and other cognitive activities and their embodiment). But there certainly doesn't seem to be an established research area here. Indeed, even computationalists such as Fodor (1975), who take an internal 'language of thought' to underlie all specific human languages do not see the need to posit 'psychologically real' mechanisms to explain linguistic competence, preferring to conceptualize language as an abstract object rather than theorize about actual mechanisms and language organs. Therefore there does not seem to be evidence of the sort Slezak claims here.

Regarding the issue of the relationship of Chomskyan linguistics to the field of linguistics in general, Botha (1989) offers some sense of the diversity of the discipline and the immense range of conceptual distinctions that demarcate 'Chomsky's view' from that of others. Thus, some linguists reject the very fundamental initial distinctions that Chomsky draws for the purpose of the study of languages. For example, his requirement that "a grammar, as a description of a particular human language has to be perfectly explicit" (Botha (1989), p. 2)[18] is rejected by 'empiricist' (as opposed to Chomsky's 'rationalist') approaches to the study of human languages and linguistic reproduction which suggest that entities such as 'linguistic universals' and even 'languages' are more socioculturally and temporally malleable creatures than the mythic entities that Chomsky's explicit grammars seem to require (cf. Lakoff (1987) and Muhlhausler (1986) for alternatives). Given that 'evidence from linguistics' could equally

[18]. This is part of the definition of a 'generative grammar'.

be seen to vindicate the 'conventionalist and contextual' views Slezak is out to disprove, invocation of Chomsky cannot unproblematically be part of a univocal and decisive refutation of SSK.

(c) Neuroscience: Slezak thinks that neuroscience research provides evidence for the computational model of belief acquisition. But generally, philosophical defenders of neuroscience research[19] have taken the view that neuroscience research lends itself best to interpretation by that position on the mind-body problem known as 'eliminative materialism', which is incompatible with the belief based mentalistic account that Slezak apparently thinks neuroscience is both based on and also provides evidence for. Hence Churchland's 'neurocomputational perspective' too takes a 'connectionist' or 'parallel distributed processing' approach to the study of cognition (see Churchland (1989), Chapter 10). Beyond this, the sort of naturalized epistemology that Churchland foresees as deriving from his computational neuroscience is of the descriptive rather than normative variety. Neither does Churchland suggest that neuroscientific research can help deliver the sort of 'socially uncontaminated heuristics for scientific discovery' for which Slezak sees evidence forthcoming from it- in fact Churchland doesn't see social factors as being irrelevant to the process of scientific knowlegde development on his account of neuroscience- "the character of the social pressures will have a vital role to play in any adequate account of learning in scientific communities" (Churchland (1989), p. 248). And with fitting irony, Churchland seems to go on to actually approve Bloor's 'Strong Programme' in its factor/variable version of the social and natural causal determinants of scientific knowledge (Churchland (1989), p. 248)! Again Slezak's unsubstantiated assetions

[19]. Cf. Bechtel (1988), pp. 102-6, and also Churchland (1989), another philosopher who Slezak credits with an informed view on CS, in contrast with the uninformed SSK meddlers.

about neuroscience research either supporting or using the Turing machine model appear too strong.

(d) Psychology: Baars (1986), at the end of his history of the 'Cognitive Revolution'[20] in psychology goes on to make the prediction that psychology

> may be moving toward a first theoretical integration... the coming theoretical integration will probably be based on some sort of information processing foundation... (p. 415).

But for Baars this is a chancy prediction, on the details of which he isn't willing to speculate. Nor does he take a stand on whether information-processing metaphors will provide merely a theoretical language for psychology or a literal model for processes actually occuring in human beings. This again seems to suggest that the decisive evidence from cognitive psychology that Slezak thinks vindicate his abstract heuristics has not yet materialized. In fact, at least some cognitive psychologists (e.g. Heyes (1989), p. 127) have expressed reservations about collaboration between themselves and epistemologists who rely solely on work in information-processing 'cognitive psychology'. Heyes also expresses uneasiness about the attempts of such epistemologists such as Goldman (1986) who criticize the research of (also 'cognitive' in some sense) psychologists on the 'rationality' of human beings, as this is taken to be the exclusive preserve of philosophers.[21] Unsurprisingly, Slezak and other epistemologists searching for 'abstract principles of rationality' construe relevant work in cognitive psychology very narrowly.

Even those psychologists who are specifically interested in using the sorts of 'computational

---

[20]. By which he means replacement of strictly behaviorist (purely 'observational') psychology by a cognitive one (allowing for the postulation of 'inner menatal processes').

[21]. The work of Kahneman et al. (1982), and Nisbett and Ross (1980) on actual 'cognitive processes' is in fact eminently amenable to 'social factors' based interpretations.

models' Slezak has in mind- e.g. Tweney (1989), (1990)- do not see them as providing either the unifying basis for CS or as instrumental in discovering the sort of abstract heuristics that Slezak takes them to provide. Thus Tweney's (1989) 'interpretive framework for the psychology of science' has an overarching 'level of goals and purposes' within which other levels such as that of 'cognitive style' and 'heuristics' are embedded. Tweney's research on Faraday doesn't seem to support the claim that the heuristics used by Faraday were 'abstract principles of discovery' detached from the other levels in which they were embedded, and hence 'context-free' in the sense Slezak seems to require. Elsewhere, Tweney emphasizes that in his work, computational models provide a useful tool to study the intricacies of scientific thinking, but not

> the larger context of our developing themes of our scientific thinking. The fact that a particular model can do something is not, in and of itself, much help in assessing its value for theory (Tweney (1990), p. 481).

Similarly Gholson et al. (1989), in their programmatic statement declare that

> a major goal for the cognitive psychology of science is to provide a theory that can account for how the *working practices* of scientists lead to developments in scientific knowledge... (pp. 267-8, my emphasis) Unlike similar models previously presented by philosophers which describe scientific developments only abstractly, we specifically call attention to the subject matter for which a cognitive psychology of science must account: the practices of working scientists... [hence] the discipline must articulate closely with the social psychology of science... cognitive structures... are influenced by how groups are organized internally and in relation to the broader society (Gholson et al. (1989), pp. 271-2).

While Gholson et al explicitly recognize the importance of social factors in the study of scientific cognition, at least one psychologist contends that there is a challenge mounting against those cognitive models (i.e. models of the brain etc. in which the putative Turing machine like mechanisms envisioned by Slezak are instantiated) which postulate

> an executive (or centralized) organization or control... [in which]... the coordination of sensation and action is accomplished by 'executive processes' or 'executive routines' that somehow intelligently 'connect' input

and output (Mahoney (1989), p. 150).

Mahoney points to alternative

> constructivist theories of human neural/mental organization... [which]... with their emphasis on holistic complexities, interacting systems and ongoing processes have frequently challenged the reductionistic, closed system accounts that have characterized twentieth century psychology (Mahoney (1989), p. 151).

Whatever the implications of these alternative theories for the 'social' accounts of knowledge, it is quite clear that the computational model cannot be unproblematically invoked in a decisive refutation of them here.

The above four subsections were intended to show that four of the five subdisciplines that Slezak took to either utilize or vindicate the computational model and abstract heuristics do not in fact appear to do so. I now turn to the question of whether AI research and the abstract model and heuristics suffice to refute the 'social' accounts on independent grounds.

(iv) <u>AI and the 'abstract heuristics for discovery'</u>

Fleck (1982), in his sociohistory of the origins and development of AI research, notes that AI does have some distinctive sociocognitive characteristics, which include the common use of the general purpose digital computer and list-processing languages, craft-skill or knowledge of these languages acquired first hand via apprenticeship, and the deployment of these resources in the construction of computational models of aspects of intelligent activity (pp. 170-71). But Fleck also notes that these basic elements have been utilized by AI practitioners for a variety of goals, and that beyond the mutual reliance on the abovementioned general tools and techniques, AI work has not been particularly methodologically or theoretically driven- rather, practitioners have

> become involved with the subject matter of many other disciplines, and [AI] has therefore developed as an interdisciplinary area. This interdisciplinary character of AI has induced many mutually competing divisions

> *within* the area, as well as leading to many external views on the status of the field, which has consequently been very much a matter of negotiation (Fleck (1982), p. 209).

These divisions include some along such basic lines as whether it is 'human intelligence' or 'intelligence in some abstract sense' that is being modelled, and whether the goal of AI is to produce 'machines that are *really* intelligent' or just 'simulations of aspects of intelligence good enough for specific purposes'. Suffice it to say here that even adepts in the 'discipline' of AI (insofar as it is a discipline), just as in the others cases noted above, would no doubt offer differing opinions on the significance of the 'Turing machine model' for their work.

Where on the map of AI research can the computer models of scientific discovery be placed? In his original paper, Slezak (1989) seems to suggest that computer models of scientific discovery are meant to be literal accounts of (apsects of) cognitive processes (occuring somewhere, in the brain perhaps) of individual human scientists (of the sort where 'social factors' play no role). The development of these models was then considered by Slezak as contributing to the project of developing an asocial naturalized normative epistemology within the philosophy of science understood traditionally (i.e. as supplying 'principles of rational discovery' to scientists). Slezak is therefore interested in AI as applied philosophy of science. Slezak (1989) thinks that the project is sufficiently advanced that

> the issue of social determination of theory content can be decided directly by asking whether scientific theories may be produced [discovered] in the absence of social factors (p. 565) [so that]... our test is designed to capture what, on the traditional view are thought to be the essential features of scientific discovery and the production of theories... [by abstracting] science from its contingent embodiment in a ubiquitous social context (p. 566) [and that]... work on BACON and AM is relevant to our concerns since they are part of the empirical enterprise of cognitive science- which after all has something relevant to say concerning human behavior and its causes (p. 571).

We have already seen above that 'what the empirical enterprise of CS has to say' concerning human behavior and its causes is, to put it mildly, quite inconclusive vis-a-vis the question of whether 'social factors' have or do not have bearing on the process of making scientific discoveries.

So we are left to consider the extent to which the computer models can be said to be simulating the 'asocial aspects of discovery processes' on independent grounds, and here too things are by no means as clear cut as Slezak thinks. For our purposes, we may distinguish two questions which arise here, which relate to parts (v)-(vii) of Slezak's argument as it was outlined in Chapter 3. The first is the question of the validity of the computer models as hypothetical constructs of human cognitive processes construed literally (and hence of the descriptive adequacy of the accounts of discovery produced). The second I take to be the independent issue of whether the computer models (whatever their relationship to actual human cognitive processes and actual scientific discoveries) do in fact constitute an 'existence proof' to the effect that there are abstract heuristics for making scientific discoveries that can be isolated from their contexts of use.[22]

On the first issue, it seems that Slezak virtually acknowledges the potential irrelevance of the computer discovery systems to questions about actual human processes and actual scientific discoveries when he echoes the views of Zytkow and Simon (1988) to the effect that the computer discovery systems are meant to provide a paradigm for a normative theory of discovery (Slezak (1989), p. 573), and not necessarily a descriptively accurate one. In the same vein, Langley et al. (1989), while assuming that

[22]. I will also consider here the matter of the 'social' status of these discovery systems.

a computer is quite a general symbol processing system... [which makes it] particularly convenient for the simulation of cognitive processes, [as] the human mind is also a symbol-processing system (pp. 32-3)[23]

go on to add that

we have focused, in constructing our programs, on the level of heuristics, trying to incorporate in them the kinds of selective processes that we believe scientists use to guide their search for regularities in data... [but]... in testing our simulations we will usually have to be content with a sufficiency criterion (p. 33).

This 'sufficiency criterion' usually amounts to the 'demonstration' that a computer model has succeessfully reproduced a specific discovery (usually drawn from the history of science) on the basis of the information available to the original discoverer using appropriate heuristics. Clearly the extent to which such a 'reproduction' is a descriptively accurate account of the 'actual cognitive processes' of the discoverer or 'actual discovery' remains an open question.[24] If even we are to assume that the computer models are in some sense simulating processes of the sort that occur in human beings, Langley et al. believe that "the chief conclusion we will be able to draw from our experiments with our simulation programs concern the nature and efficacy of search heuristics" (Langley et al. (1989), p. 34). Again, from the point of view of descriptive adequacy, the question arises of the extent to which the scientific discovery process either is (or has a separable component which can be equated with) the search of a problem-space for a solution on the basis of heuristics by the discoverer.

Downes (1990) has given a number of reasons why computational models of scientific discovery fail to be descriptively adequate accounts of the scientific discovery process. He notes

---

[23]. The relationship between actual human cognitive processes and their computer simulation is never specified in greater detail than this.

[24]. Langley et. al. (1989) do not care to speculate on this matter.

that minimally, "discovery is part psychological activity, part sociology of group acceptance and part historical accident and timeliness" (p. 23). Moreover in describing the process of scientific discovery as it actually occurs, these parts are not neatly separable. Downes notes that the spottiness of the historical 'data' (usually written records) that Langley et.al. (1989) rely on to determine what 'information' was available to the original discoverers (e.g. Boyle) to which their program heuristics might be applied cannot be equated with sorts of rich actual environments in which discoveries actually occur. Langley et al. (1989) contend that

> at the level of resolution we are interested in, it will not be hard to match initial conditions, equating the data and knowledge provided to our programs with the data and knowledge available to the scientists who made the historical discoveries under study (p. 34).

The confidence with which 'data' and 'knowledge' are ascribed to e.g. Boyle would make any historian wince.

With respect to the lack of descriptive adequacy, Downes also notes that the computer models do not account for the interactive aspects of both making and accepting discoveries. In this connection, Slezak has argued against Brannigan's (1981) 'attributional model' of scientific discovery- which holds that the acceptance of a discovery by the relevant community at least partly involves social recognition of it- that Brannigan's concern is really with "how and why something might be so characterized... (hardly more than a definitional matter)... rather than an explanation of a phenomenon in the usual causal sense" (Slezak (1989), p. 577). Presumably this means that Slezak takes *BACON* to provide a causal explanation of Boyle's discovery and that of the others. But at the 'level of resolution' at which the program operates this hardly sounds like an adequate explanation. What Slezak might be suggesting is something as qualified as the claim that BACON provides a possible causal explanation of some aspects of the essential processes involved in making the discovery. But as an 'explanation' of the discovery this counts for very little from the standpoint of descriptive adequacy without some specification of the context in which the discovery is recognized as such.[25] This would be so even if we grant (as Brannigan does not) that students in laboratory classes regularly 'discover' Boyle's Law. As putative causal explanations of the various discoveries, BACON and the other computer models seem to leave far too many loose ends to be convincing- or is Slezak willing to claim that the several different computer chess-playing systems available each provide possible causal explanations of the 'essential cognitive processes' involved in playing chess?

Thus the claim that human cognitive processes are essentially computational mechanisms is not substantiated, and the computer models do not appear to be descriptively adequate accounts of the discoveries they simulate.

This brings us to the second issue. Perhaps 'descriptive adequacy' was neither the goal of these programs, nor the basis for Slezak's empirical refutation- ultimately, therefore, claim (vii) of Slezak's argument (as described in Chapter 3) can be regarded as saying that the computer models of scientific discovery utilize asocial heuristics which (whatever their precise relationship with actual human cognitive processes or discovery procedures), because of what the programs are able to do on their basis (i.e. manifestly perform some of the recognizably important activities involved in 'discovering'), demonstrate that the contingent social and contextual factors inevitably invovled in actual discoveries can be disregarded as inessential.

Langley et al. (1989) are in fact quite explicit that their goal is not descriptive adequacy but to provide "a set of criteria for judging the efficacy and efficiency of processes used to discover

---

[25]. It is ironic that Slezak criticizes Bloor for falling short of the strict standards required for 'proper' causal explanations when this is the best he can do in their stead.

scientific theories" (p. 45). Their procedure is to try to recover from actual cases the "relatively general methods [that] play (and have historically played) a significant role in scientific discovery" (p. 46). It is noteworthy, however, that they stress that the recovery of these general methods neither exhaust all available ones nor provide a complete theory of discovery, as certain methods are quite domain specific. In fact Simon (1991) has gone on to say that the heurisitics used by BACON are not context-free at all: "BACON has the same dependence on social context as do human beings" (p. 147). Simon thus grants the point that Slezak doesn't, but goes on to add that

> the individual/social issue is not the same as the intrinsic/extrinsic issue. Computer simulations are compatible with all sorts of social influences. The actual programs under discussion assume that these influences are overwhelmingly intrinsic- that is, related to the content of science (p. 147).

But Simon regards this assumption as open to test, and in any case not essential to utility and the purpose of the programs.[26] And in a more recent work on computational approaches to scientific discovery, Shrager and Langley (1991) hold that "embedding and embodiment... have significant bearing on science but have not been addressed by existing computational models" (p. 15). They go on to urge that

> the social organization of science in the laboratory and in broader contexts has a major influence on the nature of science, and future modelling efforts should move toward incorporating aspects of this structure (p. 18).

It appears therefore that the very creators of the models Slezak cites give interpretations of them which are quite antithetical to his claim that they are standing refutations of the 'social' views.

Slezak's 'refutation' seems to be based on the following interpretation of the models and

[26]. Which is to help discover normative heuristics, be they intrinsic or extrinsic.

heurisitics. The heuristics used in the models are intended or designed to capture only the essential features of the discovery process by abstracting from a variety of actual contexts (as Simon was quoted above, the 'intrinsic' nature of what was being recovered was a guiding assumption). Slezak thinks that the 'asocial' character of the heuristics derives from here- as noted in Chapter 3, his argument is that the abstraction of the heuristics from a particular context eliminates (or perhaps reduces) the dependency on contextual features of the actual situation, and that generalization across contexts guarantees the elimination (or further reduces the dependence). The 'asocial' character of the heuristics is then confirmed (or the 'intrinsic'-influences-only assumption is vindicated) where Slezak is concerned because the heuristics actually work or are successful at reproducing discoveries. As a consequence, the 'social' views are empirically refuted by this existence proof.

Now it might appear that one might question this conclusion on the basis of the descriptive inadequacy of the computer simulations. Thus, one could argue that even granting that the heuristics used in the models were truly 'asocial', since *all* aspects of the discovery process were not covered by the model, the claim to a decisive refutation at least appears too strong- after all, 'social factors' may well be involved in the aspects not modelled. How can one claim then that the heuristics 'work' if the models are descriptively indadequate?[27] But Slezak would not be impressed by this line of argument. His response would be that BACON did simulate recognizably important aspects of the process, and that it did so on the basis of asocial heuristics. He would take this alone to suffice the 'social views'. And admittedly, such a rebuttal from descriptive adequacy would be unsatisfactory inasmuch as whether the models

[27]. This is very much like Collins' argument to the effect that BACON-like machines cannot possibly be 'scientist-mimics'.

would in the future be able to reproduce further aspects of the process (as Slezak would doubtless claim) is indeed an open question.

My own view would be that what Slezak takes to be the 'asocial abstract heuristics' used by BACON are not 'asocial' at all. They are used by human beings in concrete situations and for specific purposes, and this very fact makes them thoroughly 'social'. Slezak's response to such a claim is that of course the heuristics must inevitably be used in specific circumstances, but this does not detract from the fact that they are of universal applicability and hence independent of the context. After all, it is surely the *same* heuristic that Boyle, BACON and I use to recover the law from the data- how else could we all recover the law unless we all 'possessed' (internally stored) the heuristic (as BACON explicitly does)? To this I would have to respond that the judgment of 'sameness' itself presupposes some context, and the seeming independence of the 'content' of the heuristic from that context is illusory. (It would be as difficult to 'empirically demonstrate' this though as it is for Slezak to 'empirically refute' it). As to the question of what BACON (and like discovery systems) can actually do on the basis of the heuristics (and the relationship of what the 'machine does' to the 'heuristics it possesses' need not be a straightforward one), my response would be that if what the machine does is recognizably (even significant aspects of) making discoveries, then the machine ought to be regarded as a 'social being'.

Chapter Five: Science as 'Practice'

In this final Chapter, I first set out a 'practice' conception of scientific activity. I then use this conception for the purpose of tying together the material of the previous four Chapters-I develop a contrast with the methodologies of Collins and Woolgar, reexamine my alternative analysis of Slezak, and suggest avenues for further development.

(i) <u>A 'practice' conception of scientific activity</u>

In this section, I derive two metarequirements on scientific practices and four metaconstraints on the analysts thereof. I begin by listing them here so that it is easier to see how they are subsequently motivated and utilized.

Metarequirements on practices:

(1) Determinateness: There is agreement between the practitioners over actions and procedures within the practice.

(2) Analysability: Practices must be 'open' to piecemeal analysis.

Metaconstraints on analysts:

(1) Explicitness: The aspects of the practices under consideration must be identified to the satisfaction of practitioners.

(2) Theoreticality: The analyst must make use of appropriate theoretical resources in her accounts.

(3) Reciprocity: The analyst must allow for the continual scrutiny of her accounts.

(4) Normativity: The analyst must recognize the normative status of her accounts and attempt to gauge their consequences for the practices in question.

In the introduction to a recent collection of Science Studies essays, Pickering (1992) provides a set of contrasts between 'science as knowledge' and 'science as practice' with which will be useful to begin. The 'science as knowledge' view takes the analyst's subject matter to be the *products* of scientific activity- usually the products are understood to be 'scientific theories' or 'scientific knowledge'. The 'science as practice' view by contrast stresses the *process* of scientific activity- the analyst in this case pays attention to what scientists actually do and what resources they use to do it (Pickering (1992), pp. 2-3). The 'science as practice' view therefore involves a broadening of the analyst's vision in looking at scientific activity. For elucidatory purposes, Pickering suggests that one might extend one's analysis to address at least three distinct aspects of scientific activity rather than looking at the products of science as 'knowledge'- namely the conceptual, the material and the social (though in the actual course of scientific activity these 'aspects' are obviously intertwined in complex ways). Let me build upon these distinctions in what follows.[1]

(a) Conceptualizing and analysing scientific practice:

The conception of a 'practice' developed here is a 'constructivist-objectivist' one which takes a practice to be a purposive, rule-governed (collective) activity. I have labelled it 'constructivist' because the social, material and conceptual resources that are implicated in practices, though drawn from the 'background pool' of the larger culture, are themselves not specifiable in their entirety for the purposes of the analysis. This ensures that practices are *autonomous*; that is, from the standpoint of the analyst the features of the practice she is concerned with cannot be taken to be determined by any 'underlying reality'.[2] The conception takes practices as

[1]. I have freely used here the ideas of Wittgenstein (1953).

[2]. Thus one might say that practices are 'constructed' out of the available cultural resources, but just in case one did not misleadingly assume that all features of the practice are *actively* 'constructed' from scratch in this way.

'objective' because the analyst assumes that the activities of the practitioners are quite determinate- that is, the analyst assumes that the practitioners can give reasons for their actions within the practice- this is why practices were taken to be 'purposive, rule-governed activities'. This does not mean, however that the analyst expects the reasons for the practitioners' activities (i.e. the rules governing the practice) to be straightforwardly apparent insofar as she is unfamiliar with the practice (i.e. not a practitioner herself). For the relationship between the 'rules' governing the practice (i.e. the reasons that the practitioners would give for their various actions) is an *internal* one- that is, the reasons given by the practitioners for their actions within the practice (or the 'rules' they formulate to explain them) ultimately give out, and then they simply 'act blindly'. The objectivity of the rules correlates with the *lack of doubt* that accompanies the actions of the practitioners.[3] 'Objectivity' in this sense is then taken by the analyst as a *metacondition* for something to be a practice- she expects that there is a (working) consensus among the practitioners which is the result of their prior agreement in (a preponderance of their) actions.[4]

(b) The possibility of analysis of scientific practices:

How then does the analyst proceed with her investigation of scientific practices (the 'rules' in the above sense of which she is presumably unfamiliar with)? The answer to this question derives from the fact that the analyst is not obliged to account for 'whole scientific practices'

---

3. I would like to thank Prof. Peter Barker for clarifying this point. This notion of rules being 'internal' to practices (or equivalently, of practices being prior to rules) is not the same as the 'internal-external' distinction with respect to 'science' however. *That* 'internal-external' distinction cannot be sustained, see below.

4. If these are the features of practices in general, what distinguishes a 'scientific' practice from other kinds? The answer to this would have to be- a variety of criteria (e.g. subject-matter, methods used, activities involved, goals, historical circumstances)- and societally grounded paradigmatic judgments on their basis.

(i.e. all at once, and in their entirety) in practitioners' terms (or 'actor's categories'). The analyst is not so obliged, and can yet proceed with her analysis because she can be expected to have *some* familiarity with the 'background pool of resources' (material, conceptual and social) which particular scientific practices draw on.[5] Essentially this amounts to the belief that even 'autonomous scientific practices' are in some degree 'permeable'- that is to say, practices have *aspects* or *features* (and again, these may be material, conceptual or social) which are readily apparent to the analyst and tractable to examination on the basis of what are strictly speaking 'external' resources (i.e. resources that the practitioners themselves might not use or be familiar with). This in turn implies a rejection of that sense of the 'internal-external' distinction which takes a practice to be not just autonomous (where it has its own 'rules' as described above) but free-floating (where these rules are beyond non-practitioners in principle)- all practices derive from a 'shared background of activities' which makes a purely 'internalist' (in the sense of rigidly demarcated from the 'rest of the culture') stance with respect to them untenable. This is in effect the second metarequirement for something to be conceptualizable as a 'practice'- it must be amenable to piecemeal investigation in this way.

(c) Metaconstraints on the analyst:

While the analyst assumes the 'objectivity' of the practices she examines and their tractability to piecemeal analysis as described above, there would be certain constraints placed on her analysis as well- which would be meant to enhance the adequacy as well as the utility of the accounts given. Firstly, the analysis must aspire to explicitness. This is to say that the aspects

[5]. As stated above, these cannot be expected to be given a determinate and exhaustive 'listing' in any sense- the 'background' always remains only partially specified, and the practices remain autonomous.

of the practice under consideration must be clearly indicated.[6] This in turn means that secondly, the use of 'theoretical resources' by the analyst is desirable. The virtue of such 'theoretical' accounts is not that they are ipso facto justified, or acquire independence or 'objectivity' from the analysis itself.[7] Rather, it is that it helps makes the analyst's presuppositions explicit and renders the account given more accessible. This brings us to the third requirement, that of reciprocity. The analyst must be reciprocal in the sense that she must both herself examine as well as allow for the scrutiny of her own account (especially by the practitioners of the activity being examined). The reason for this is that the analyst does not suppose her account to be 'impartial' or 'unbiased', but believes that its acceptability needs to be arbitrated in a public setting. Fourthly, the analyst should explicitly recognize the normative status and possible consequences of her accounts- the analyst's examination of particular features of scientific practices may result in proposals for their reorganization. This should lead to an attempt on the part of the analyst to transform practices examined in a process involving mutual renegotiation between the analyst and the practitioners. All the above metaconstraints emphasize the need for reciprocal engagement with the practice being analysed.

(ii) <u>Collins and Woolgar's methodologies in perspective</u>

To further draw out the above conception of a practice, let me draw some contrasts with Collins' and Woolgar's methodologies- to what extent can Collins' and Woolgar be said to be taking a 'practice' view of scientific activity, and how far compatible are their methodologies

---

[6]. This does not mean that the analyst would have to give universal 'rigorous definitions' or 'precise identifications' (as I noted above that these would not be forthcoming in any case), but contextually appropriate ones.

[7]. For the analyst's theories too are to be understood as 'representational conventions' having various analytical uses. (Note that nothing precludes the theories appealed to by the analyst from being 'scientific' or 'naturalistic' ones).

with the metaconstraints proposed above?

The salient features of Collins' model of scientific change can be summarized as follows. 'Society' is to be understood as a 'network of concepts' together with institutionalized rules for their use. The 'institutions' here are identifiable groups who associate these concepts with one another in the course of their activities (and it is these associations that 'link' the concepts together to create the network). The overall societal network is then taken to be comprised of a number of subnetworks, each linked to the others: thus the 'network of science' consists of scientific concepts linked to one another by such 'joint entrenchment' in the activities of scientists. Changes occur when particular concepts in the scientific network (whose meanings/uses are taken to be 'open-ended', i.e. their associations and range of applicability is taken to be subject to unpredictable change) are 'modified' by changing their links to other concepts, or 'new concepts' are introduced which are linked to specific existing ones. The 'mechanisms' for these changes are negotiations between members of the various institutions whose own concepts are potentially affected by putative changes, and the results of such negotiations are new links between concepts.

The 'practice' conception is different from this network model in the following ways. The network model is firstly one-dimensional in that it theorizes exclusively on the level of concepts and their linkages. On the practice conception, the emphasis is on scientific *activity*, and creation and use of concepts is only part of that activity. Now admittedly Collins incorporates other aspects of science by talking about the use of concepts within various activities, but the primary analytic focus remains concepts (and the goal remains the documentation of the creation/extension of concepts). Second, Collins preconceives 'scientific concepts' to form a network that is distinct from the overall societal network, and linked to it in specific ways.

Taking this as a guiding analytical assumption, however is to assume the very 'internal-external' distinction that the practice conception is at pains to dissolve. It leads directly to the 'different worlds' ontology that was criticized earlier, and implies a stronger than desirable 'incommensurability' of the subnetworks to each other by insisting that the various institutionalized groups can only know 'their own parts' of the network. The mysterious thing on this view becomes how it is that Collins qua analyst is able to understand all the parts of the overall network. The practice conception by contrast assumes that such 'incomensurabilities' between practices, though real are surmountable, and just not through establishment of 'conceptual links' alone.[8]

This brings us to the second issue- that is the compatibility of Collins' methodology with the proposed metaconstraints on the analyst. With respect to explicitness and the desirability of theoretical descriptions, it seems to me that to the extent that Collins relies on the 'alternation' technique as a source for his accounts of science and 'expertise' as their legitimating basis, he has failed to satisfy these metaconstraints.[9] Next, insisting that the sociologist of science's accounts of scientific activity not be subject to the same standards that the analyst applies in looking at science, Collins explicitly rejects the reflexivity requirement. It might appear that he doesn't entirely do this, because he does leave room for an examination of his accounts, only insisting that the question of what standards are appropriate for their examination remains an open question. But it seems to me that by insisting on 'appropriate standards' in this way, Collins preempts criticism until such a time as the appropriate standards are found- this amounts

---

[8]. See Fuller (1988), Chapter 3.

[9]. Admittedly, this is more of a 'conceptual' quibble than anything else, for Collins' case studies are quite developed in these respects, though one might complain that the negotiation mechanisms might be better developed.

to 'internalism' in the sense criticized.

Finally, Collins' descriptivist methodological stance toward the network of concepts analysed is questionable with regard to recognition of normative status. The purely descriptivist approach can be interpreted as an assumption on the part of the analyst of an independently existing network which can be unproblematically accessed. It also assumes that the analyst's activity in doing the describing itself has no effect on the network, or the negotiations of the new conceptual links that arise. But I would contend that on Collins' model and methodology, it would be wrong for him to think that all he was doing was post-hoc description- after all, the very act of participant observation, the conceptual clarifications sought by him from the core-set members in a controversy could affect the outcome of the negotiations. It would also be a mistake for him to suppose that he was giving 'impartial' accounts in this sense, as the Slezak case makes clear. This is why the practice conception urges the explicit recognition of the normative status of the analyses and an assessment of their possible consequences (conceptual, material and social).

Woolgar, unlike Collins, doesn't propose a full-blown model for scientific change given his commitment to presuppositionless ethnographic inquiry. The closest one gets to a 'theoretical basis' for his work remains the discourse-and-representations view of the sort presented in Woolgar (1986). Woolgar conceptualizes a 'discourse' as an *institution* which incorporates 'structures of explanation, systems of categorization and modes of conventional practice'. What is categorized are the 'classes of entities' with which the discourse is involved and their character and interrelationships. Discourses therefore come with their own 'orders of representation', which are the accepted systems of categorization and the hierarchies of relationships between entities within them. How does Woolgar's notion of a discourse compare with the practice conception sketched above? Woolgar's notion of a 'discourse' seems to be analogous to Collins' 'network of concepts' insofar as the one-dimensionality of their accounts goes. Woolgar seems to lay excessive stress on the 'representational' aspect of discourses as the appropriate object of analysis, and fails to acknowledge the 'other aspects' of discourses that are also implicated in representational practices- again, the primary commitment is to the 'conceptual' aspects of the discourses examined. And just as Collins does consider these other features via the 'institutionalized use' of concepts, Woolgar too takes the notion of 'discourses' to include 'modes of conventional practice' within institutions.[10] But again, analytic access to these other features is constrained by the preemphasis on representations.

Unlike Collins, on the other hand, Woolgar does not preanalytically differentiate 'scientific discourses' from 'other discourses'. In this regard, his view of the 'constitutively social' or 'actively constructed' nature of the representations within discourses rejects the internal-external distinction which Collins seems to sustain by his separation of the 'network of science' from the rest of society. Woolgar's notion of discourses is in this respect very much like the minimalist conception of a practice as a 'purposive, rule governed autonomous activity' which was proposed above- further characteristics not being specified outside the context of concrete cases. However, two further distinctions need to be made with respect to Woolgar's notion of discourse and the practice conception of the previous section. Woolgar's insistence on the 'active construction of representations' might overemphasize the extent of the 'construction' of representations within discourses- as I noted above, at least some representations are taken for granted, both by practitioners as well as the analyst in the course of their activities. Woolgar's

[10] Which is an expansion from purely conceptual considerations to activities more broadly construed.

'radical constructivism' therefore involves the danger of 'overanalysing' practices. Secondly, while Woolgar does reject the internal-external distinction with respect to 'science' and 'society', he concedes too much (or so I believe) to the second sense of internalism- the internalism arising from the autonomy of discourses. Thus Woolgar does not impose the second metacondition described above on his 'discourses', holding that the analyst takes on the entire 'order of representation' associated with a discourse at once.

But rejecting the need for this metacondition on discourses also implies the rejection of some of the metaconstraints on the analysis itself, which I now turn to. With respect to the requirements of explicitness and use of theoretical resources, Woolgar's methodology seems to come up short as follows. His 'radical constructivist' commitment to a presuppositionless ethnographic method assumes that the analyst is free to examine any aspect of the discourse she desires, because it is assumed that the accounts produced are themselves equally open to subsequent evaluation.[11] But then Woolgar's interrogation of representations can then take any of the representations within the discourses examined to be problematic on what might turn out to be completely arbitrary or 'occasioned' grounds- no further justification would be required because of the above promissory note. Woolgar's principled skepticism about representations is therefore capable of selective application in the course of analysis, which the more sustained engagement required by the use of theoretical resources would not allow.

This brings me to the reflexivity requirement. It might appear that Woolgar can hardly be accused of failing this, given his near obsession with the many-levelled examination of his own accounts. But all the same, Woolgar does not allow for the cross-examination of his own

[11] This is partly why Woolgar rejects the second metacondition of discourses- he assumes that representations within discourses are straightforwardly available in this sense.

accounts by others in any sustained way, especially by those 'represented'. In this regard, the avowed principled skepticism and the achieved occasioned skepticism toward his own accounts amounts to a preemption of their criticism by admitting their occasioned character.

Finally, there is the question of the acknowledgement of the normative status of one's accounts. To the extent that Woolgar takes the goal of his project to be the finding of alternative modes of representation that will not involve undesired imposition of authority, it appears that he is looking for representational practices that will allow him to give purely descriptive accounts of discourses. But there is some question as to how successful we can take Woolgar's project to have been in this regard. As I noted earlier, there is no 'neutral-observer' standpoint from which Woolgar can give his accounts of representational practices with impartiality and the alleged freedom from biases. Again to some extent Woolgar professes impartiality and agnosticism because he is interested only in studying the 'representational aspects' of discourses as a project in its own right, so that he does not need to interfere in the discourses he examines in order to carry out his own project.[12] But again, he can neither assume that he has no immediate effect on the discourses he examines, nor no long term effect via the recommendations made concerning representational practices. This much, then should be acknowledged.

(iii) Illustrating and reevaluating the Slezak analysis, and beyond

It is now time to assess my assessments of Collins and Woolgar, as well as my alternative analysis of the Slezak case given the practice conception I have sketched in the first section.

[12].I believe this is why Woolgar rejects the second metaconstraint on practices stated earlier- he needs unproblematic access to 'whole discourses and their associated orders of representation' at once to be able to identify their representational practices unproblematically.

My alternative analysis of Slezak was intended to avoid what I took to be the shortcomings of Collins' and Woolgar's analyses. To what extent have I succeeded in doing so? On the plus side, I can say the following on its behalf. I made no claim to impartiality in the analysis- I began with the statement that I did not admit the possibility of the 'decontextualized abstract heuristics' that Slezak claimed BACON et al to be providing. Unlike Collins, I attempted to provide an explicit reply to each of the points raised by Slezak and to provide alternatives to his 'refutation' claims without assuming the prior distinction between the natural and social worlds that prevented Collins' analysis from getting into the details of Slezak's claims themselves. I also took my replies to Slezak's claims to be accountable to the same sorts of standards that I had applied in the analysis. And in contrast with Woolgar's response, I attempted to explicitly engage the issues of the debate, and did not attempt to merely agnostically pass judgment on the nature of the representations and their construction from the 'outside'. I also tried to assess Slezak's 'representations' without supposing that my own were so thoroughly 'occasioned' that no criticism of them was possible.

To what extent does my own analysis of Slezak correspond to the guidelines I have proposed in the practice conception sketched in the first section? Here the answer is less satisfactory. I began examining Slezak's 'refutation' by trying to determine to what extent Cognitive Science could indeed be considered a 'practice', a collective rule-governed activity with recognizable conceptual, material and social features. I took the position that where Slezak saw apparent uniformity, there were at best a variety of distinct practices in the various subdisciplines that Slezak invoked to prove his case. However, my account too was 'one dimensional' in the sense I criticized Woolgar's and Collins' accounts of being in the above section: I took up the issue of whether Cognitive Science research could indeed be considered a coherent practice by considering only the question of its shared *conceptual* basis- namely whether the activities of cognitive science researchers were unified by their shared use of the 'Turing-machine' model. This was an adequate starting point for the analysis because it was one of Slezak's central claims and I used the second metacondition on practices above, beginning a piecemeal analysis on its basis. But admittedly, the analysis is far from complete, as I did not consider other possible bases for a coherent cognitive science practice- material, social, and alternative conceptual ones. Of course, I would still claim that pursuing the analysis would fail to secure the sort of refutation which Slezak had in mind, but some further reconciliation between my 'social' conception of knowledge and some parts of Cognitive Science, or their further modification could at least be expected.

Next, I criticized Collins and Woolgar for their failure to use 'theoretical resources' in their analyses of Slezak. But my alternative analysis of Slezak did not really use alternative theoretical resources either- I merely indicated where theoretical resources were to be found which rendered Slezak's claims to a decisive refutation problematic. Clearly, whether the theoretical resources that I mentioned are indeed all compatible with the view of scientific activity that I have begun to develop, and with each other, is a matter for more detailed investigation as well. What is apparent at this stage is this- if an account of CS practice that is relevant to the issue of understanding scientific practices is to be found, this will take the sort of sustained dialogue and transformation that I identified as underlying the metaconstraints on analysis of practices above- this I cannot claim to have more than begun. Finally, related to this point, I criticized Collins and Woolgar for failing to make their presuppositions explicit in the course of their analyses. I have attempted to begin to do that for my own account here, but that they would need to be made more explicit in the process of the production of a more complete account of Cognitive Science is obvious, and the extent to which this needs to be done is clearly not entirely my prerogative.

Bibliography

Baars, B.J. (1986). The Cognitive Revolution in Psychology. New York: Guilford Press.

Barker, Peter. (1989). The Reflexivity Problem in the Psychology of Science. In Gholson et al. (1989).

Barnes, Barry. (1978). Interests and the Growth of Knowledge. London: Routledge and Kegan Paul.

Barnes, Barry. (1983). Social life as bootstrapped induction. In Sociology, 4, 524-45.

Bechtel, William. (1988). Philosophy of Mind: An Overview for Cognitive Science. Hillsdale, NJ: Lawrence Erlbaum.

Berger, Peter and Thomas Luckmann. (1967). The Social Construction of Reality. New York: Doubleday.

Bijker, W.E., T. Pinch and J. Law (Eds.). (1987). The Social Construction of Technological Systems. Cambridge, Mass.: MIT Press.

Bijker, W.E. and J. Law (Eds.). (1992). Constructing Networks and Systems. Cambridge, Mass.: MIT Press (forthcoming).

Bloor, David. (1976/1991). Knowledge and Social Imagery (2nd Edition, 1991). London: Routledge and Kegan Paul.

Bloor, David. (1981). The Strengths of the Strong Programme. In Philosophy of the Social Sciences, 11, 199-213.

Bloor, David. (1983). Wittgenstein: A Social Theory of Knowledge. New York: Columbia University Press.

Botha, R.P. (1989). Challenging Chomsky. Oxford: Basil Blackwell.

Brannigan, A. (1981). The Social Basis of Scientific Discoveries. Cambrdige: Cambridge University Press.

Callon, M., J Law and A. Rip (Eds.). (1986). Mapping the Dynamics of Science and Technology. New York: Macmillan.

Chomsky, Noam. (1980). Rules and Representations. New York: Columbia University Press.

Churchland, P.M. (1989). A Neurocomputational Perspective. Cambridge, Mass: MIT Press.

Cole, D.J., J. Fetzer and T. Rankin (Eds.). (1990). Philosophy, Mind and Cognitive Inquiry.

Dodrecht: Kluwer.

Collins, H.M. (1979). The Investigation of Frames of Meaning in science: Complementarity and Compromise. Sociological Review, 27, 703-18.

Collins H.M., (Ed.). (1981a). Knowledge and Controversy: Studies of modern natural science. Special issue of Social Studies of Science, 11 (1), 1981.

Collins, H.M. (1981b). What is TRASP?: The Radical Programme as a Methodological Imperative. In Philosophy of the Social Sciences, 11, 215-224.

Collins, H.M. (1985). Changing Order: Replication and Induction in Scientific Practice. London: Sage.

Collins, H.M. (1989). Computers and the Sociology of Scientific Knowledge. Social Studies of Science, 19, 613-34.

Collins, H.M. (1990). Artificial Experts. Cambridge, Mass.: MIT Press.

Collins, H.M. and S. Yearley. (1992). Epistemological Chicken. In Pickering (1992).

Collins, Randall. (1988). Theoretical Sociology. New York: Harcourt Brace Jovanovich.

Coulter, J. (1983). Rethinking Cognitive Theory. New York: St. Martin's Press.

Downes, Stephen. (1990). Prospects for a Cognitive Science of Science. Ph. D. Dissertation, Virginia Polytechnic Institute and State University.

Elias, N. and Richard Whitley (Eds.). (1982). Scientific Establishments and Heirarchies. Sociology of the Sciences Yearbook. Dodrecht: Kluwer.

Fleck, J.A. (1982). Development and Establishment in Artificial Intelligence. In Elias and Whitley (1982).

Fodor, J.A. (1975). The Language of Thought. New York: Crowell.

Fuhrman E. and K. Oehler. (1986). Discourse Analysis and Reflexivity. In Social Studies of Science, 16, 293-307.

Fuller, Steve. (1988). Social Epistemology. Bloomington: Indiana University Press.

Fuller, Steve. (1992a). Philosophy, Rhetoric and the End of Knowledge. Madison: Univesisty of Wisconsin Press.

Fuller, Steve. (1992b). Talking Metaphysical Turkey about Epistemological Chicken. In Stump and Galison (1992).

Fuller, Steve, M. de Mey, T. Shinn and S. Woolgar (Eds.). (1989). The Cognitive Turn.

Dodrecht: Kluwer.

Garfinkel, H. (1967). Studies in Ethnomethodology. Englewood Cliffs, NJ: Prentice-Hall.

Gholson B., W. Shadish, R. Neimeyer and A. Houts (Eds.). (1989). Psychology of Science. Cambridge: Cambridge University Press.

Giere, R. (1989). Computer Discovery and Human Interests. In Social Studies of Science, 19, 638-43.

Goldman, A. (1986). Epistemology and Cognition. Cambridge, Mass.: Harvard University Press.

Hesse, Mary. (1974). The Structure of Scientific Inference. London: Macmillan.

Heyes, C. (1989). Uneasy Chapters in the alliance between Psychology and Epistemology. In Gholson et al. (1989).

Holland, J., R. Nisbett, J. Holyoak and P. Thagard (Eds.). (1986). Induction: Processes of Inference, Learning and Discovery. Cambridge, Mass.: MIT Press.

Kling, R. (1991). Computers and Social Transformations. In Science, Technology and Human Values, 16, 342-67.

Knorr-Cetina, K. and M. Mulkay (Eds.). (1983). Science Observed. London: Sage.

Kornblith, H. (Ed.). (1985). Natuaralizing Epistemology. Cambridge, Mass: MIT Press.

Kahneman, D., P. Slovic and A. Tversky (Eds.). (1982). Judgment Under Uncertainty: Heuristics and Biases. Cambridge: Cambridge University Press.

Lakatos, Imre. (1978). The Methodology of Scientific Research Programmes. Cambridge: Cambridge University Press.

Lakoff, G. (1987). Women, Fire and Dangerous Things. Chicago: University of Chicago Press.

Langley, P., H. Simon, G. Bradshaw and J. Zytkow. (1989). Scientific Discovery. Cambridge, Mass.: MIT Press.

Langley, P., G.L. Bradshaw and H.A. Simon. Rediscovering Chemistry with the BACON system. 1980. In R.S. Michalski et al eds., Machine Learning. New York: Springer-Verlag.

Latour, B. (1988). The politics of explanation. In Woolgar (1988b).

Latour, B. (1992). Where are the missing masses? In Bijker and Law (1992).

Latour, B. and S. Woolgar. (1979). Laboratory Life. Beverly Hills: Sage.

Laudan, L. (1977). Progress and its Problems. Berkeley: University of California Press.

Laudan, L. (1981). The Pseudo-science of Science? In Philosophy of the Social Sciences, 11, 173-98.

Lawson, H. and L. Appagnanesi (Eds.). (1989). Dismantling Truth. London: Weidenfield and Nicolson.

Lenat, D.B. On Automated Scientific Theory Formation. (1979). In J.E. Hayes et al eds., Machine Intelligence 9. New York: John Wiley.

Lynch, M. (1992). Extending Wittgenstein: the pivotal move from epistemology to the Sociology of Science. In Pickering (1992).

Mahoney, M.J. (1989). Particapatory Epistemology and Science. In Gholson et al. (1989).

Muhlhausler, P. (1986). Pidgin and Creole Linguistics. Oxford: Basil Blackwell.

Nisbett, R. and L. Ross. (1980). Human Inference. Englewood Cliffs, NJ: Prentice-Hall.

Pickering, A. (Ed.). (1992). Science as Practice and Culture. Chicago: University of Chicago Press.

Ramsey, W., S. Stich and J. Garan. (1990). Connectionism, Eliminativism and the Future of Folk Psychology. In Cole et al. (1990).

Rouse, J. (1990). Knowledge and Power. Ithaca, NY: Cornell University Press.

Shrager, J. and P. Langley (Eds.). (1991). Computational Models of Scientific Discovery and Theory Formation. San Mateo, Calif.: Morgan Kaufmann.

Simon, H. (1991). Comments on the Symposium on Computer Discovery. In Social Studies of Science, 21, 143-56.

Slezak, P. (1986). Actions, Cognition and the Self. In Synthese, 66, 405-35.

Slezak, P. (1989a). Scientific Discovery by Computer as an Empirical Refutation of the Strong Programme. In Social Studies of Science, 19, 563-600.

Slezak, P. (1989b). Computers, contents and causes. In Social Studies of Science, 19, 671-95.

Slezak, P. (1990). Man not a fit subject for science? In Social Epistemology, 4, 327-42.

Slezak, P. and W.R. Albury (Eds.). (1988). Computers, Brains and Minds. Dodrecht: Kluwer.

Stump, D. and P. Galison (Eds.). 1992. Disunity and Context: Philosophy of Science Studies. Stanford: Stanford University Press, 1992.

Thagard , P. (1988). Computational Philosophy of Science. Cambridge, Mass.: MIT Press.

Tweney, R. (1989). An Interpretive Framework for the Psychology of Science. In Gholson et al. (1989).

Tweney, R. (1990). Five Questions for Computationalists. In Shrager and Langley (1991).

Whittaker, J.L. (1978). Language Games and Forms of Life Unconfused. In Philosophical Investigations, 1, 39-48.

Winch, P. (1958). The Idea of a Social Science. London: Routledge, 1958.

Wittgenstein, L. (1953). Philosophical Investigations, tr. G.E.M. Anscombe. Oxford: Basil Blackwell.

Woolgar, S. (1981). Interests and Explanation in the Social Study of Science. In Social Studies of Science, 11, 365-94.

Woolgar, S. (1983). Irony in the Social Study of Science. In Knorr-Cetina and Mulkay (1983).

Woolgar, S. (1985). Why not a sociology of machines? In Sociology, 19, 557-572.

Woolgar, S. (1986). On the alleged distinction between Discourse and *Praxis*. In Social Studies of Science, 16, 309-17.

Woolgar, S. (1987). A Note on some recent sociological critiques of cognitivism. In Bijker et al. (1987).

Woolgar, S. (1988a). Science: The very idea. London: Ellis Horwood/ Tavistock.

Woolgar, S. (Ed.). (1988b). Knowledge and Reflexivity. London: Sage.

Woolgar, S. (1989a). The Ideology of Representation and the Role of the Agent. In Lawson and Appagananesi (1989).

Woolgar, S. (1989b). Representation, Cognition and Self. In Fuller et al. (1989).

Woolgar, S. (1989c). A Coffeehouse conversation on the possibility of mechanizing discovery. In Social Studies of Science, 19, 658-68.

Woolgar, S. and D. Pawluch. (1985a). Ontological Gerrymandering: the anatomy of social problems explanations. In Social Problems, 32, 214-27.

Woolgar, S. and D. Pawluch. (1985b). How shall we move beyond Constructivism? In Social Problems, 33, 159-162.

Woolgar, S. and M. Lynch. (1990). Representation in Scientific Practice. Cambridge, Mass.: MIT Press. (Special Issue of Human Studies, 11, 1988).

Woolgar, S. and K. Grint. (1991). Computers and Social Analysis. in Science, Technology and Human Values, 16, 368-78.

Zytkow, J. and H. Simon. (1988). Normative Systems of Discovery and Logic of Search. In Synthese, 78, 65-90.

RANJAN CHAUDHURI
(10 June, 1967)

8700F Foxridge Apts.
Blacksburg, VA 24060.
703-552-0686

Center for the Study of Science
in Society, VPI & SU
Blacksburg, VA 24061
703-231-7687

EDUCATION

M.S., (1992). Virginia Tech, Graduate Programme in Science and Technology Studies. Thesis title: "Science as Practice:A Methodological Critique and Case Study". Advisor: Prof. S.W. Fuller.

B.S., (1990). Mathematics.
A.B., (1990). Philosophy. Lafayette College.

HONORS AND AWARDS

Virginia Tech: Graduate Assistantship, Partial Tuition Waiver (merit based).

Lafayette College: Honors in Philosophy. Charles A. Dana Scholar.

R. Chaudhuri

Ranjan Chaudhuri
May 1992