5-1995

The Social Structure of Science

Natasha L. Nummedal

The Social Structure of Science

In partial fulfillment of upper division

in

philosophy

by

Natasha Nummedal

May 1995

Dr. Husain Sarkar - thesis advisor

Table of contents

Chapter I

Introduction

Within the field of the philosophy of science, there has always been an emphasis on rationality. In fact, the terms "science" and "rationality" simply seem to go hand in hand. But most traditional philosophies of science have emphasized the individual rationality of scientists. Husain Sarkar introduces a new idea: group rationality. In the past, philosophers of science have overlooked the importance of the scientific community, having focused instead on the behavior of the individual scientists. But if the aim of science is to advance scientific knowledge, philosophy of science must take not only the individuals but the community as a whole into account. For just as an individual may have not only epistemic but personal goals for achieving some end, so too the group of scientists is striving to achieve, if not some precise goal or end, some general advancement in the general body of knowledge. It is the role of the philosophy of science, and a theory of group rationality in particular, to provide suggestions and options that help science work toward its goals. Specifically, the philosophy of science provides methods for scientists to use in their research: methods which suggest which theories of science to believe (such as the most-corroborated theory), and which theories offer the greatest degree of promise for either solving particular problems in science or leading to newer and better theories.

But how is someone in science to sort through all of these methodological options posed in the philosophy of science? Sarkar

emphasizes the role of heuristic advice in a method, claiming that a method can be tested and evaluated based on its heuristic advice, in a way similar to the testing of theories in science. Briefly, the heuristic component of a method advises which scientific theory or theories ought to be examined and tested in the hopes of developing new and better theories. Theoretical advice, on the other hand, ranks the theories of the time in order of their degree of corroboration.

A. *The importance of group rationality*

This division between the heuristic and theoretical advice of a method suggests that the best theory available to a field of science is not necessarily the only theory a scientific community should engage in. As Philip Kitcher mentions (a distinction made by Lakatos), there is a difference between belief in a theory and pursuit of the theory. A conflict arises: would a scientist agree to work and further a theory he thinks is wrong and a waste of time?

The importance of group rationality is that it calls for a division of labor amongst theories (or, as Sarkar suggests, methods). Kitcher mentions some dangers of individual rationality. If all of the members in a scientific community evaluate two theories and each scientist decides theory T_1 is the best, then everyone in that field will be working on T_1, without distributing labor and giving another theory T_2 a fair chance. The same problem could happen with methods. Such situations lead to an irrational group structure.

Group rationality considers the structure of the community. While science would certainly advance quickly if

scientists only studied the theories that are true or verisimilar, such knowledge is available only in retrospect. Through a division of scientific labor, the community can advance science faster by ensuring that the underdog theory of the day is not overlooked. Either by studying a problem in light of another theory (even if it is wrong) or by giving a competing theory or theories an opportunity to solve such issues can a scientific community be sure (or relatively sure) that it has arrived at the best possible theory or solution at that time.

B. *Goals of a scientific community*

Karl Popper is one of the first philosophers of science to emphasize the important role that goals play in the advancement of science, and claims that certain goals are better than others. Just as the theory choice of a scientist is dependent on the method he uses, so too the choice of which method to use is dependent on the goals of the scientist. In addition, the goals of the entire scientific community affect the choices made by its individuals and the rational structure of the community.

Sarkar suggests the existence of ideal or "objective goals," which other goals "aim" toward. Depending on the goals of a given method, we may determine if it is better or worse than its rival methods. Judging by the standard of an absolute goal, may we ask if the goals of a scientific community determine if that community is better or worse than its rivals? And do the ends (achieving these goals) justify the means of the community?

Chapter II

On *group rationality*

In the last chapter of A Theory of Method, Husain Sarkar introduces his own suggestions and proposal for the philosophy of science. His main goal is to present his view of group rationality, a view of science that recommends for a scientific group to divide its labor amongst not only competing theories, but more interestingly, amongst competing methodologies. By proliferating methods, a scientific group in turn proliferates its theories, goals, and aims, leading the group to an atmosphere ripe for criticism, testing, comparisons, and proposals of new theories. Science is intricately tied to the concept of rationality, but most theories of method until this point have focused on the problem of the individual's rationality. Sarkar's theory of method proposes that one also look at the group's rationality as a whole as a way of determining how science should or should not be practiced. Sarkar proposes that "a group is rational if, and only if, it is organized in subgroups, each of which adopts a distinct method for pursuing its goals" (148). These goals interlace with each other within the entire group to tie together the scientific group as a whole.

Sarkar presents four theses that he further develops throughout the chapter. But briefly, he first suggests that methods ought to give heuristic advice. Second, as a way of evaluating methods, scientists should experiment with methods in the way that they would experiment with competing theories, rather than to rely on the history of science as sole arbitrator between methods. As a

scientist would likely choose the theory most adapted to his concept of the world or the reality of what he is (or believes he is) looking at, so too he would choose the method most adapted to his concept of truth/verisimilitude or his personal goals. Third, Sarkar suggests evaluating his theory of method in a way similar to the evaluation of theories in ethics. And lastly, he announces the need for a method that elucidates group rationality, and that considers what *makes* a group rational.

A. *Heuristic advice*

Heuristic advice, given by the scientist's chosen method, is the advice which a scientist employs as a guide for selecting a particular theory or several theories as the most appropriate for both experimentation purposes and development of the theory or theories in the hopes of furthering the body of scientific knowledge and of producing newer and better theories. Ideally, all methods of science can or should provide three forms of advice: theoretical, practical, and heuristic. Theoretical advice differs from heuristic advice by selecting which theory is "best" at the time, and practical advice recommends the theories that are applicable to a particular problem in science. Heuristic advice is unique, therefore, because it recognizes that if scientists are aiming at verisimilitude, the method should provide not only for the possibility (and likelihood) of future, more "truth-like" theories, but should also help the scientist in choosing the theory or theories most apt to *lead* the scientist to such future theories. The present theories a scientific group may be operating with may all be deemed weak (or strong), but it is only

with pushing different theories that the possibility of a new theory will evolve which best satisfies the group's goals. The heuristic advice of a method points the scientists to the most fruitful theories available to develop. In other words, the heuristic component of a method is that part most directly responsible for the growth and development of scientific knowledge.

The heuristic advice of a method helps the scientist "rank-order" the theories at hand into a list of most to least fruitful for leading to future theories. Without this rank-ordering, a method could lead to an "anything goes" approach, which surely is not the aim of (most) methods, let alone a feasible or productive approach to use in research. But by critically selecting the theories to be used, the method contributes to the growth of knowledge. In addition, heuristic advice cannot and should not be viewed as algorithmic, for no scientist or methodologist would claim that theories can be created by the method or its advice, and no one would even venture that the best of theories suggested by the heuristic advice will produce another or better theory. Heuristic advice simply allows for a wider range of theories to be used (not necessarily limited to the best theory available according to the theoretical advice of the method, although at times it may be the only theory being used) to provide the scientists with a larger and richer framework of theories to draw from in doing their research.

Heuristic advice is not only for ranking theories in a body of scientific knowledge; it can also be key to choosing methods themselves. Sarkar suggests testing methods in the same sense as testing theories, since "a method is defined by its objective and

normative components" (161). Part of the normative component in a method is rational decision, in other words, following the heuristic advice of the method. Hence, one can test a method by testing its heuristic advice. If its advice proves itself useful in the expansion of knowledge and aids the scientist in discovering new theories, one could say the advice is "true" and the method followed is "true." If the heuristic advice is useless, however, the method also would show itself useless and warrant being replaced. Since methods and their heuristic advice are so crucial to scientific progress, it follows that science and methodology should focus on and integrate the issue of heuristic advise as the advice that provides science with an arena in which to develop and further scientific knowledge. After all, "...it is the methodological commitments of the scientists, and especially the heuristic advice such a method gives, that will ring in, or profoundly affect, the scientific tradition" (141).

A scientist, when choosing theories he deems to be worthwhile for further research and experiment, is said to be in "reflective equilibrium" if the decisions "dovetail" with the heuristic advice of a method (167). The scientist may then employ that particular method in future choices or in conducting his work, especially when confronted with difficulties in his field. The chosen method is not necessarily the final method that the scientist or the group will choose, of course, but provides a reference for research and decisions. Just as with scientific theories, if the method proves itself unsatisfactory, it can be modified or changed. It is important to remember that "the correctness of truth of heuristic claims is not

based on what even the ablest members in the field believe, think, or do. It is based on the objective methodological reality" (170).

So how does a scientist arrive at a reflective equilibrium? First, he must realize that he needs a method to help in evaluating decisions in his research, even if theory-choice is not always done in conscious, full light of the advice given by the method. Next, he considers the plausible methods available to him and judges the competing methods according to their success rates both inside and outside his given field of research, putting their heuristic claims into practice or studying how these claims benefited or hindered other fields. The scientist then makes a choice based on the results of his findings. This is only a model, however, and different scientists will of course evaluate methods in different ways, based on their experiences and their particular judgments. After all, "it is most unlikely that *all* the scientists in a given group or society will have more or less the same preanalytic judgments that will drive them to select or arrive at *one* method which they all regard as the best one" (168).

To illustrate the use of heuristic advice in testing methods, Sarkar draws from an example in moral philosophy. In A Theory of Justice, John Rawls creates a hypothetical situation in which rational, moral agents are placed in the "original position," a position where the agents know general information about their society, but otherwise they are unaware of any particular status or role they may have as an individual in the society, such as race, property, sex, creed, etc. Rawls then asks, "By what principles ought they to govern and regulate their society" (173)? After considering a

variety of governments, Rawls argues that they must inevitably end up at his theory of justice.

Robert Nozick extends Rawls' example. Nozick proposes that the moral agents start in Rawls' original position, but they arrive at another principle of justice, other than that of Rawls. Adhering to that chosen principle, they enter society and begin to conduct their normal affairs. After a while, however, this principle will slip out of reflective equilibrium as more and more problems begin to arise. The moral agents will then return to the original position with the tacit knowledge they gained, and reevaluate the possible theories of justice. They will chose another theory, and the cycle continues. Eventually, Nozick claims the moral agents will arrive at Rawls' theory of justice, but only after actually living in a society and developing their tacit knowledge and moral attitudes.

Based on Nozick's hypothetical situation, Sarkar poses an example in which scientists choose a method in the original position, and then step out of the original position to proceed with their research. Once problems begin to arise with the heuristic advice provided by the method (assuming they did not choose the ideal or true method from the beginning, of course), the scientists return to the original position and reevaluate methods. But it is because of the scientists' work under a particular method that there is brought to light the specific problems with the heuristic advice and that they realize the need for reform. It is this form of testing that leads scientists to determining which method provides them with the most fruitful heuristic advice. In addition these scientists, or knowledge-seekers, "develop some more their sense of rationality [and] gain

further tacit knowledge of the correct conditions under which theories should be accepted or rejected" in a way they could not have done before with a formal, methodological discussion in the original position (175). The knowledge-seekers are also able to locate problems with the chosen method's heuristic advice faster and more forcefully than without operating under that method.

There is a significant difference between the moral situation and the methodological situation, however. Mainly, in the moral situation the resources are scarce. In a moral society, an individual cannot simply have another person's property without proper legal issues resolved. But in methodology, a methodologist or a scientist has access to the methods and theories proposed by others within his group since the aim in methodology is knowledge. "The knowledge-seekers will agree that pieces of knowledge should be treated unlike pieces of property.... [they] know that a lack of free, easy, and open access to scientific knowledge and discovery can seriously hamper scientific growth" (176). Paradoxically, a scientific group is rational for distributing knowledge and resources amongst all its members, but within our western society it would be deemed irrational in the moral sense, and vice versa. The goal of a scientific community or group is not, however, the moral sense of justice and equality, it is rather the pursuit of knowledge.

The above example shows the importance of the heuristic component of a method, and its significance over theoretical and practical advice. In fact, heuristic advice is one of the points that leads Sarkar to proposing a proliferation of methods. Since the best way to test methods involves applying their heuristic advice, he

suggests that a group of scientists be further divided into subgroups that each employ a unique method of science. Through this proliferation not only will different methods be tested, but scientific theories will also be examined using various approaches due to the differing goals of the subgroups.

B. *Review and criticism of Popper's theory of method*

Karl Popper's theory of method focuses on a critical tradition, where bold scientific theories are proposed and tested, resulting in either falsification or corroboration of the theory. Only through such rigorous practice can science advance. Yet strangely enough, despite Popper's desire for severe testing of scientific theories, he does not support testing of methods, especially of his own. For Popper, methods are conventions, unlike scientific theories, and thus have no truth value. Methods, according to Popper, may be judged by their "fruitfulness" and their "distance from absolute rightness," but cannot be tested by scientists nor tested against the history of science since they lack any normative component which could be true or false.

To illustrate his view of a critical tradition, Popper uses the Ionian physicists to exemplify what he considers the *first* and foremost critical tradition in the history of science. In Popper's view of the history of science, the importance of this scientific community stems from Thales, Anaximander, and Anaximenes each posing bold theories about the world, and each in turn creating a new theory that could strike at the core of the theory that came before it. To begin, Thales proposed that the primary element of the universe was water,

and all existing matter was simply water in one of its phases (gas, liquid, or solid). Next, Anaximander contested Thales' theory by pointing out that the earth could not be supported by water, for such a theory would lead to an infinite regress, for what was supporting the water, and so on. So Anaximander suggested that the earth was in balance in the very center of the universe, and being in this perfect center, the earth could not possibly want to move. He thus conjured that the primary element of the world was the *apeiron,* introducing the notion of the infinite. Bothered by Anaximander's theory of an indeterminate, Anaximenes rejected the previous theories and proposed that the primary element was something determinate yet still unlimited: air. To explain different forms of matter, Anaximenes claimed air could exist in either dilated (fire) or condensed (solid) forms.

Of course none of these theories could hold water (so to speak) against observation. But what was important for Popper from this historical example was the climate of criticism that existed at the time. This type of environment encouraged the presentation of such rival theories to be directed at the cores of one another. For Popper, growth in scientific knowledge can occur only through such bold proposals of theories and strong refutations.

But what distinguishes the critical from the uncritical traditions for Popper? After all, if Popper views the Ionians as the first critical scientific community, and such a tradition was not reinvented until the time of Galileo, a long and uncritical history spanned between the two communities. Is it possible to maintain the existence of degrees of criticism in the history of science? In

addition, a critical attitude surely cannot be a sufficient condition for the growth of knowledge, but even if it were sufficient and necessary for progress in one scientific community, say the Ionians, how could we claim history to repeat itself in a different community such as Galileo's? At most, a critical tradition is only a necessary condition - it certainly cannot cause growth in scientific knowledge. No matter how similar two scientific communities may be, conditions will always be different, such as individual interests and personal background, scientific background, and goals. But even if two societies were so similar, to the point of "identity," there is no just reason to think their levels of progress would also be identical, or even similar.

Of all philosophers of science, Popper may be one of the first to emphasize the importance of adopting certain goals and aims in science, and not only concerned with how to choose which theories. Some goals are better than others, according to Popper, in the same way certain scientific theories are better than others. In fact, goals are significant in science because in the same way the theories chosen by a scientist are dependent on the scientist's method, so the method-choice is equally dependent on the scientist's goals. Hence goal's are crucially important, since "arbitrary espousal of goals and aims leaves open the charge of arbitrary acceptance or rejection of theories" (45).

But if scientific theories aim at truth or verisimilitude, what may goals possibly "aim" at? Popper believes that the aim of goals is beyond rationality, but Sarkar suggests ideal goals, or "objective goals." He makes the analogy between scientific theories

and methods. Given a theory, *T*, there are parts or terms of *T* that either "hook onto" certain universal parts and processes or not, in other words, it is an "all-or-nothing affair" (45-6). We can talk about this theory T as being able to capture certain properties more or less than another theory (unless they are equivalent), and thus claim it to be a better or worse theory. The same idea works for methods. Given a method, *M*, there are various goals belonging to that method that either "hook onto" the objective goal or not - an "all-or-nothing affair." Depending on which method captures more of the objective goal or goals, may we claim it to be a better or worse method than its competitors? This argument provided by Sarkar makes aims and goals an objective fact. Since Popper readily admits that certain goals are better than others, could his theory of method accept certain methods be better than others? According to Sarkar,

> One method can be better than another in either of the following ways: other things being equal, the method captures more fundamental goals than do its rivals.... [or] the first method characterizes the nature of the goals correctly, or more correctly, than the other method.... Thus we can account for the growth of fundamental aims and goals. (47)

Popper considers methods conventions - they are neither true nor false, they can only be "good" or "fruitful." Such conventions Popper terms "norms." He distinguishes these norms from natural laws, because natural laws are either true or false (whether we know the answer or not), and if a natural law is true it cannot be replaced. Norms, however, can be changed and replaced, since they can neither

be derived from or reduced to facts. But several problems arise if Popper wants to maintain that norms are neither true nor false, because it is not obvious why such a statement thus leads Popper to believe methods must be conventional.

Sarkar points out that it is not necessary for Popper to maintain that norms lack truth values. First of all, given Popper's claim that norms cannot be derived from nor reduced to facts, one does not necessarily need to conclude that norms must therefore lack a truth value. Sarkar makes the analogy that scientific theories cannot be reduced to or derived from singular observation statements, but that does not leave us to think that these theories have no truth value! Obviously, they must be right or wrong, even if we will never know. In addition, there still exists a relationship between norms and facts: although they cannot be derived from one another, facts nonetheless are *evaluated* by norms. Popper admits that if a norm were to contradict a natural law, it would be useless or non-fruitful. According to Popper, "It is perfectly true that our decisions (norms) must be compatible with the natural laws... if they are ever to be carried into effect; for if they run counter to such laws, then they simply cannot be carried out" (qtd. on 36). Norms evaluate facts, but if the norm evaluates certain facts wrong, providing us with false information about the world, could we not then say that that norm is false?

Sarkar claims that despite Popper's efforts, it would be highly consistent with Popper's method to maintain methods as normative, and not as conventions. "[Popper] can give an account of progress in method that is analogous to the account of progress in

science: one methodological proposal, *M*, is better than another proposal, *M'*, if, and only if, *M* is more like the absolutely correct method than is *M'* " (35). By Popper's claim that methods are conventions, a method cannot give heuristic advice for choosing one method over another when all else, such as methods' ability, are equal. Yet it is the heuristic advice of the method that leads to progress, and without it, a method is rather useless. How, then, is progress in science to be understood by Popper if not by one method being better than another, possibly due to having better goals, if we cannot even evaluate or test methods?

In addition, Popper states that "a student of the logic of science may well take an interest in [the history of science or empirical science], and learn from it" (qtd. on 50). If a student learns from the history of science, which Popper claims cannot be used to test methods, then is it possible for that student to learn how to formulate a better method for testing scientific theories, or learn that a method being used in science at the time is not as successful as it may claim to be, or any number of other reasons to either question, replace, or formulate methods? Is it right to claim that "a methodologist can contribute to the work of a historian, but the historian has nothing to contribute to the work of the methodologist" (49)?

C. *Empirical, conventional, and normative methods*

Empirical methodological statements claim to tell us how scientists behave as a matter of fact. Within empirical methodological statements, there are two positions: lawlike and not

lawlike. The lawlike position, held by methodologists such as Brian Ellis, maintain that scientists behave with very regular behavior. The non-lawlike position, posed by philosophers like Larry Laudan, believe that scientists are without any regulated behavior.

For example, within the non-lawlike position, Laudan claims that historians of science and methodologists try to discover the values and goals of the scientists within the group they are studying, they speculate which method may best accomplish these particular goals of the scientists, and finally they describe the behavior of the scientists in light of this chosen method. Assuming, for instance, that the Ptolemists' goal was to predict the phenomena of the world, they proposed ad hoc theories, creating separate models for each planet, the sun, and the moon. Whenever irregularities arose with their predictions, they used different epicycles to correct the minor flaw. But for the sake of contrast, suppose the Copernicans wanted to learn something about the world and how it works, not to just be able to predict the movement of astral bodies. Due to their different goals, the Copernicans avoided ad hoc theories and used theories that could be empirically tested. "Divergent methods explain divergent practices" (53).

What this example illustrates is that, although the scientists in their respective groups were not necessarily *aware* of their particular method, they were still operating under methods based on their particular goals. The non-lawlike position, therefore, simply reports the history of these groups. It records the history without evaluating the behavior of the group or the individuals, without claiming any lawlike regularity in the scientists' actions, and

without trying to fit the entire history of science into a single method.

On the other hand, the lawlike position of Ellis takes a very programmatic approach to the history of science, evaluating the actions of scientists to figure out the causal processes of these individuals, such as asking how they develop their beliefs and their reasoning. "Ellis is concerned only... with how rational men think" (54). Since this is the science of studying scientists, Ellis views his method as "scientific epistemology." His aim is to find the "fundamental laws of rationality governing an ideal rational agent" (54). Of course, Ellis admits that in scientific epistemology, as in the physical sciences, there are deviations and even ideals are rarely if ever achieved by real phenomena. But these deviations in scientific epistemology, according to Ellis, can be explained away in the same way deviations are explained in the sciences. To Ellis, the laws of rationality employed by the scientists are equivalent to the laws of logic: they are the ideal forms of these individuals' beliefs. Ellis' view is termed lawlike for his claim that rationality is genetically determined, since it is absolutely necessary for human survival.

Ellis' position of lawlike regularity of behavior is useful in helping a methodologist determine the heuristic advice of a scientific group's method, and then testing this method against other parts of history or by applying it in modern-day science. Sarkar presents an analogy to help understand Ellis' position. For a human to develop language, he requires as a necessary condition both a physical structure, such as a brain, nervous system, particular shape of mouth, etc., and the exposure to a language-using society. So as for

science, the scientists need to both "internalize" certain principles so they can conjecture or reject theories as well as have exposure to a science-practicing community that has scientific problems for them to work on. So if, as Ellis states, we depend on our rationality for survival, we must genetically possess the ability to internalize the right things that help us survive, for surely we could not have existed long with just a trial-and-error method.

The problem that arises is whether or not such a method should be used in methodology. For one, it is extremely programmatic. This method does not account for the drastic changes in thought and how scientists suddenly view the world differently when presented with a newer and better theory, such as the transition between Ptolemaic and Copernican science. Second, the laws of logic Ellis emphasizes so strongly can be necessary conditions at the best, for there is no way of deriving heuristic advice or methodological laws from the laws of logic, let alone be able to derive two conflicting methods of science, as in the historical example above. And lastly, to claim humans are "programmed" to act in such specific ways means that not only can we not have methodological mistakes, in history or in present science, but that there is also no room for progress in methods.

Conventional methodological statements, as mentioned earlier, are Popper's position. Sarkar provides an historical example of science, and using two methods (Popper's and his own), shows the importance of the heuristic advice on the development of the rival theories. Does the method a group of scientists follows really affect the progress of science?

The example comes from two opposing views in evolution: the gradualists and the punctuationalists. The gradualists maintained that evolution occurs slowly over a long period of time, whereas the punctuationalists claimed that evolution occurs in quick and abrupt spurts. Sarkar assigns Popper's methodology to the punctuationalists and an opposite methodology that he calls his own to the gradualists. Hence, Sarkar's scientific group is able to propose ad hoc theories, while Popper's group uses only testability, corroboration, simplicity, etc. to judge the appropriate theories to use. The two groups then proceed about their scientific business. The first example in scientific history they run across are the bowfin fishes: fish that barely even evolved over 65 million years and showed a sudden spurt in their evolution. So while the gradualists questioned the fossil record, the punctuationalists moved onto other examples. The groups came across the lungfish, who also evolved rapidly during speciation, and by studying the trace fossils of the first shelled animals they determined that they, too, evolved in spurts. But most interestingly, at one of the final examples from evolution that Sarkar uses, the scientists discovered that in the time tables of evolution, when looking at certain rodent-like creatures that evolved into mammals, there was absolutely no time at all for these animals to become mammals according to a gradualist theory.

So what does this lengthy example show? For one, of course, that the punctuationalists had a far better theory than the gradualists. But the theories, over the course of conducting research, were shaped and honed by their methods. So if both groups, understandably, began with the same goal of verisimilitude, could we

not say that Popper's methodology succeeded far better than his opponent's? By this example, Popper's method had more effective rules and norms than Sarkar's did for reaching their goal. The extreme success as illustrated here of a method, according to Sarkar, cannot possibly be mysterious luck. Rather, the success was, at least partially, due to the method followed, its goals, and the heuristic advice that method had to offer.

Finally, Sarkar talks about normative methodological statements. A normative approach to methodology allows for progress in *method* - both successes and mistakes. Borrowing terms from Popper, Sarkar discusses the objective methodological problem-situation of a scientific group: a situation that every scientific community or society faces, in which they have a methodological past and the solution to present problems in research will call for a method. As with the evolutionary example, Sarkar claims that Popper's method solved the methodological problem of the group better than the method taken by Sarkar.

For Popper, the problem situation of a scientific community must be understood through more than just knowing the history of science. As with Galileo's theory of tides, to understand what Galileo did and how he arrived at his theory, one must know the scientific problem and the problem situation against Galileo's background and the historically theoretical framework. So as with understanding Popper's theory of method, one must understand the problems he faced and the problem situation against his background and theoretical framework.

What Sarkar aims to do, of course, is to show that methods are objective facts, and that the results from following a certain method leads to a truth or falsity. Sarkar constructs an analogy, where he shows that just as the scientific laws are due to the stable physical or biological systems (for without such stability scientists would have to formulate new laws), so too the methods taken by a group depend on a stable scientific community. But then how do methods change? First of all, a change may result from the introduction of a new method, such as when Bacon's method replaced earlier ones. Another change may fall from reasons outside of the field of science itself, such as politics or religion.

So new problems arise for methods on how to evaluate scientific theories and what useful heuristic advice should or could be given. As Sarkar claims, there are conjectures and refutations in methodology in the same way conjectures and refutations exist in science. Methods are not empirical in the same sense, for one cannot make sufficient lawlike claims for methodology, but nonetheless methods are normative, for they are both objective and possess a truth value, no matter how hard it may be to determine.

D. *Group rationality*

In the final section of A Theory of Method, after showing the importance of providing heuristic advice in the form of a normative statement to be evaluated, and the role of goals in science, Sarkar presents his own theory of scientific methodology: the theory of group rationality. To introduce his theory of method, he first discusses the distinction between the classical view of rationality

(held by philosophers such as Popper, Lakatos, Kuhn, etc.) and his own view of multiple methods. Whereas the classical position evaluates rationality on the individual level of each scientist, the multiple methods view considers rationality of the scientific group as a whole. Sarkar's aim is to show why this classical view causes instability in science, why it leads to less scientific progress than his own view of group rationality, and why it is unable to adapt to the changing goals of science.

So first the classical view of rationality. This position focuses on the individual scientist's actions and choices to determine whether that individual is rational or not, according to the method being used. To be rational in one's group is to follow the dictates of the group's methodology. As, for example, in a group operating under Popper's method aiming for the verisimilitude of theories, a scientist is rational if he uses only testable scientific theories, avoids ad hoc theories, and severely tests and pursues the most corroborated theory at the time.

There are two positions that Sarkar evaluates in the classical view of rationality: the single-theory and the many-theory approach. The single-theory approach (Popper and Laudan) allows a scientific community to take on only one theory at a time for testing and research. The many-theory approach (Lakatos and Feyerabend) encourages a scientific community to practice and research as many theories as possible or as allowed by the methodology. Lakatos, for example, suggests a proliferation of research programs so as to avoid committing a group of scientists to just one theory, the advantage being that scientists can compare their theories' abilities with those

of other theories. Or for Feyerabend, the importance of proliferation lies in the fact that one can spot the shortcomings of one's own theory through the presence of its competitors. The problem with their approaches, however, is that "while theories may proliferate in a single domain, the *method* by which the theories in any domain are measured in that scientific community should be the same" (179).

Sarkar draws an analogy from a distinction made by Rawls in his book, A Theory of Justice. Rawls differentiates between the principles that govern individual justice and the principles governing a just *institution*. The two forms must not be confused, for an individual is normally judged according to a particular action during a specific time. An institution, however, must be evaluated looking at the whole state and determining whether it is or is not just. Similarly, Sarkar proposes a demarcation between the principles that govern individual rationality and the principles governing *group* rationality. While an individual scientist may be considered rational or not according to how he chooses or rejects his particular theories, the group must be evaluated as a whole. "A theory of group rationality, such as the theory of multiple methods, will enable us to judge whether a given scientific society as defined, for instance, by its methods and practices of evaluating theories, is rational or irrational" (180). In addition, just as Rawls claims that the primary subject of justice is the basic structure of a society, Sarkar maintains that the primary subject of rationality is the basic structure of the scientific community.

For Sarkar, one of the major problems that arises within the classical view of rationality is that the classical position does not

make such a distinction of rationality. Instead, it focuses on the importance of the individual's rationality and the scientist's choice in theories at either a particular point in time or over a set interval of time, rather than looking at the rationality of the group as a whole. What constitutes a scientist's rationality in any of these views? "A consistent adherence to his adopted method and his ability to change the norms of his method under specific scientific and nonscientific pressures will yield insight into the rationality of an individual scientist" (182). So what, then, makes a *group* rational? To claim that the group is rational in so far as the individuals that make up the group are rational commits the fallacy of composition, for one. But Sarkar goes on to show that the single-theory approach and the many-theory approach both fail on a deeper level when reading these theories of method *as if* they were addressing group rationality.

In the many-theory approach, given a method M, suppose there are p theories (out of q total theories) to be used in scientific research. With the theory of proliferation, it is acceptable for any member of the scientific community operating under M to choose any theory in p. Let T be a theory in p, and assume T is also the best-corroborated, according to M. Suppose every member of the scientific community decided to adopt T. Every member would be considered (individually) rational according to their method M, and yet there would be no proliferation of theories in the community! Such an event would fly in the face of the method, which calls for a proliferation, but still maintains individual rationality to those who

do not research alternative theories. Hence, we have a situation in which the individuals are rational, but the group is not.

A similar argument can be applied to the single-theory approach. In Popper's method, every member is rational if and only if they adopt the theory which is best-corroborated, name it T_1, in the choice of p theories. At time t, the set $p = \{T_1, T_2, ..., T_n\}$, where T_1 has undergone a history of severe testing and done well and T_n is the least corroborated. So a scientist is irrational if he accepts any theory other than T_1 at time t. When at time t' the theory T_1 encounters problems, the set p is reordered: $\{T_2, T_3, ..., T_n\}$ and the scientific community moves on to working with T_2. The problem that arises out of such a method like Popper's is that no new theories may be introduced into the set p until all existing theories have been exhausted and the set of theories is left wanting at some future time.

The failure of these methods, according to Sarkar, is their lack in distinguishing between group and individual rationality. Such a distinction provides one more issue upon which to evaluate methodologies. For if the goal is for science to progress as rapidly as possible with the best of theories available, and the multiple-theory approach considers proliferation as the most certain way to achieve such ends, would it not be beneficial to have a method which effectively safeguards the proliferation of theories within a scientific community? In other words, what would make a scientific group rational?

Sarkar breaks a scientific group down into stable subgroups, each headed by a different methodology. By segmenting the group in this way, the group will possess distinct aims, goals, and

theories, with a lesser probability of operating under the same theory as in the many-theory approach which depends on a single method. The scientific group as a whole is united or defined by a set of shared goals, where the particular subgroups may even have some overlapping goals. A subgroup is considered individually stable if it has a set of scientific problems for its members to work on, if it has the ability to solve some of these problems, and if the subgroup is able to generate new theories.

A subgroup is successful partly due to its method and the heuristic advice it follows. Of course, the success or failure of its members' choices will not necessarily be *due* to the heading method, for there are always some extraneous factors also involved in theory choice. But over time, the subgroup's success will nonetheless be influenced by the method it is following, and will allow scientists and methodologists to evaluate the method's success as the method is engaged in science and theory choice by its subgroup.

A main goal for the view of multiple methods is that the subgroups not only be different, but competing and conflicting as well. Sarkar makes an analogy between the multiple method approach and the many-theory approach. Just as some scientific theories are better than others or have a higher degree of verisimilitude, the heuristic advice given by some methods is better than that given by other methods. Just as Lakatos and Feyerabend suggest that a proliferation of scientific theories allows them to compete against one another to produce better theories, so too when proliferating methods, according to Sarkar, "to let diverse methods compete and find out which method leads to more fruitful results in

the long run" (187). Such testing of methods would lead the scientists and methodologists to discover the strengths and weaknesses of the competing methodologies and to propose newer, alternate methods.

To further explain his view of multiple methods, Sarkar provides two illustrations from the history of science. The first illustration is from the ancient theories of optics. Sarkar quotes a historian of science (David Lindberg in "The Science of Optics") as saying that "these three theories [intromission, extramission, and mediumistic theories] defined the principal battle lines with visual theory" (188-9). Lindberg goes on to say that this conflict between the competing subgroups of optics "was not merely a debate over the direction of radiation, for it was thoroughly intertwined with the basic questions about the aims and criteria of optical theory" (189).

The scientific group studying optics and vision was divided into three subgroups. The first subgroup was that of the atomists, such as Epicurus and Lucretius, who were concerned with the intromission theory: that an object issues atoms to the eye. "While members of this subgroup espoused different variations of the theory, they shared not only a common core of the theory but also a common family of problems" (189). The second subgroup, the mathematicians such as Euclid and Ptolemy, upheld the extramission theory: that the eye sent out radiation that "felt" the object and sent the results back to the eye. The third subgroup (Aristotle and Galen) promoted the mediumistic theory: that there was an independent medium which transferred the perception of the object to the eye.

What is interesting about these different subgroups is that they illustrate how diverse subgroups could hold such competing theories and goals. An explanation of certain phenomena that one subgroup may fall short on, in terms of one of its competing theories, was of little or no consequence to them. For example, the second subgroup was aiming for a mathematical model of vision, whether it took regular scientific observations into account or not. The mediumistic theorists fell short on mathematical explanations, yet here too such problems were nonessential. "There was proliferation of theories in the field of optics because there were multiple methods" (190). Over time, the success of one theory or method over the others developed, but it was during that crucial time period, in which these scientists operated, that such proliferation of method aided in discovering which theories to eventually continue to pursue.

The second illustration is from the history of medicine. Three subgroups, again, compose the scientific group of ancient physicians. The first subgroup, the Methodists, were skeptical about experience, so rather than conducting experiments or research, they relied on inferring possible cures directly from the symptoms of a disease. Due to their methodological foundation, the Methodists avoided the use of drugs or surgery for disease, and tried to control sickness through diet, drink, exercise, etc. The second subgroup, the Empiricists, stressed experience but not theoretical research. For them, a cure was reliable only after repeated experimentation and applications. The last subgroup is the Rationalists, or the Dogmatists. The Dogmatists aimed for progress in knowledge, rather than just

direct cures for every individual case, so they emphasized theoretical work and research, such as anatomical dissection. They were interested in the *nature* of disease, and not in simply proposing ad hoc statements for cures. Eventually the Dogmatists became the predominant school in medicine because of the success of their method and the progress made in science as a result of their goals and aims. As the other schools faded out of the picture and competition between methods disappeared, however, less and less progress was made in the field of medicine for several centuries. As Sarkar claims, "what prompted this sterility in the study and practice of medicine, it might be explained, was the lack of the proliferation of methods" (192).

Next, Sarkar explains "why the subgroups will be in a continuous shifting equilibrium" (192). If stable, all the subgroups within a scientific group will advance forward and make progress in knowledge, although probably at different rates. The equilibrium of a given subgroup can be determined by the stability of its competing subgroups as well as the environment around them. Disequilibrium, then, is when there is movement of scientists between the subgroups. Just as in a chemistry example of gases trying to achieve stability between two or more containers or spaces, so too the members of the subgroups will shift around from one subgroup to another when their original subgroup begins to have problems. If their subgroup is not as successful as its rivals, the scientists will shift to those methods which are more productive and promising. Popper writes that "our system of aims not only changes, but it can also grow in a way closely similar to the way in which our knowledge

grows" (193). Sarkar extends this to the view of multiple methods and conjectures that this theory of method may be a way of determining the downfalls of certain goals espoused by different methods. The failure of certain goals is obviously an internal problem that evolves within the method throughout its practical use in science. But there are external problems that also lead to instability of subgroups and to shifts in members, such as political pressures or social need.

So how does the theory of multiple methods, since it takes group rationality into account, answer the problems that confronted the classical positions earlier? First, consider the problem faced by the many-theory approach: within a scientific community where proliferation is not only acceptable but ideal, it was rational for individual scientists to all choose the most corroborated theory to work with, yet the community itself was irrational for not proliferating theories. Sarkar admits that it is entirely *possible* for the same result to happen under the multiple methods view, but that it is much more *unlikely* than in the many-theory position which operates under a single method. Working under one method, scientists are more likely to choose the same theory because their community is operating under analogous heuristic advice. Different heuristic advice given by contrasting methods is far more likely to lead to distinct theory choice.

But more than just heuristic advice is involved in theory choice. A proliferation of theories is also likely to occur under the multiple methods view because of the competing goals and aims held by the individual subgroups. If one subgroup is aiming for

verisimilitude, it is not necessarily going to accept and research the same scientific theories as the subgroup that is interested in the theories with the most problem solving abilities, for example. As in the illustration from ancient optics, the scientists adopted such divergent theories as a result of their contrasting goals: one subgroup wanted a mathematical model of vision, another one was more interested in explaining the observable phenomena as well as possible. As Sarkar points out, "from the vantage point of Feyerabend, the view of multiple methods would seem best to capture his intentions," for the conflict that arises from two or more competing methods and goals is far deeper at the roots of the scientists than the competition of rival theories within the same methodologies (194).

Sarkar explains the distinction between the multiple methods position and the many-theory approach as being the emphasis on method in his view. It may be contended that his theory of method is just like the many-theory approach under one method, and that the various subgroups are a result of the different emphasis placed on the method's goals by the individual scientists within the group. But Sarkar proposes experimenting with and testing *methods,* in a way similar to scientific theories, to determine which method provides better heuristic advice and has goals closer to the "objective goals." Just as other methodologists may emphasize the importance of scientific theories and the theory choice, Sarkar stresses the role of method choice as being integral in the whole process of the growth of knowledge. Indeed, there is a close relation between the trio of method, scientific theories, and metaphysics. For

Sarkar, however, it is the methods which are the movers and shakers of the growth of knowledge, not independently from theories and metaphysics, but nonetheless intricately there.

Another problem the theory of multiple methods may face is if a large segment of the scientists throng to the best subgroup and method, and leave no proliferation of methods. With such a great shift between subgroups, the group as a whole would seem highly unstable. But Sarkar points out that such an event does not necessarily lead to a drainage of all the other subgroups to a point of their disappearance, nor establishes that the scientists would all gravitate toward the same subgroup. In addition, "the basic goals and aims of the members of a subgroup are not easily relinquished... and so the evolution of a subgroup, and its subsequent extinction, is likely to be a long, drawn-out process" (196). As long as there are as few as two competing subgroups, a proliferation of theories and a faster growth of knowledge are possible.

For Sarkar, group rationality is truly an unquestionable issue that methodologies must address. Science is so familiar with the concept of rationality and continuously claims rationality to be one of the defining traits of science. But surely the rationality of an individual scientist cannot explain how progress in science is fully possible. It is also important to account for group rationality in terms of heuristic advice and goals in science, for conventional methods are unable to resolve the consequences of the issue of group rationality.

In an ending note, Sarkar shows what he imagines to be a remarkable succession falling out of the ideal practice of group

rationality. If the scientific group is rational, the subgroups are rational if they are stable and in that they are following their method and goals. If the subgroups are rational, there are rational members within the subgroup conducting research. Hence, the rationality of the group, due to the structure of multiple methods, guides the rationality of the scientists. Of course not every individual is necessarily rational, but such a predicament does not lead one to think that the entire subgroup is irrational, nor if a subgroup is irrational does it imply the irrationality of the scientific group as a whole.

Lastly, Sarkar suggests that in confronting the issues of how science ought to be conducted, philosophers of science should keep in mind their picture of utopia. Used constantly in ethics for describing ideal morality, utopia provides us with a (hypothetical) visualization of how something *ought* to be done. What would an ideal scientific community look like and how would it operate under a given method? And, is this the portrayal desired? For Sarkar, the ideal scientific community is bound up in the image of an idealistic society, for the two are inseparable in the pursuit of any form of knowledge. As Sarkar claims, "We should focus on the structure of a scientific society as we do, and must, on the theoretical products of that group, for we shall learn more about what a reasonable society should be" (198).

Chapter III

On *the division of cognitive labor*

In his paper, "The Division of Cognitive Labor," Philip Kitcher looks at a scientific community from a perspective of group rationality (partially influenced by the work of Sarkar), but with a twist: he introduces a scientific community that differs from its more idealistic counterpart (a "pure" group) by fashioning a "sullied" group - a group of individual scientists each working for personal gain, or a group of egoists. Kitcher holds that the goal of a scientific community is to achieve progress in science quickly and efficiently, and to test rival theories in science by distributing the cognitive labor of the community as best as possible amongst the rivals to give each theory a fair chance. The rationality of the community is more important than the individual rationality of each scientist, for the goal is to distribute the resources (such as the scientists) appropriately between the theories. A distribution that allows each plausible theory a fair chance at solving the problems at hand is considered rational, whereas an unbalanced distribution leads to what Kitcher calls an irrational scientific community.

Group rationality is significant for its ability to look at the entire scientific scene as it aims toward the advancement of science, rather than each individual in the field aiming for and working on the most-corroborated theory at the time. As a result of these conflicting approaches, individual and community rationality are often at odds with one another. For the progress of the entire

community, it is best that the individual scientists aim at a rational community rather than focusing on a particular theory or personal belief that they might individually hold as true, and Kitcher finds a sullied community more competent for achieving such ends than the pure group.

But in any scientific community, how is a "philosopher-monarch" watching the actions of the group to convince scientists to act in accordance with group rationality? What rewards or incentives are there for the scientists working on the theories they believe in, and on a personal quest for the truth or verisimilitude? For the pure scientists, of course, there is the reason that a fair distribution amongst the theories will lead the group as a whole to verisimilitude faster, ensuring better theories for the future generations of scientists. It is an altruistic approach, however, and unlikely to be taken up under more ordinary circumstances.

Kitcher suggests a scientific community be housed with egoists: those out for personal fame, fortune, or any other selfish reason. Could a scientific society run better with such a crowd, or would it hinder the progress of the group? According to Kitcher, it is possible that the group of egoists, or the "sullied," may operate *better* than the "pure" group. In a pure community, the individual scientists are motivated by their search for the true or the verisimilar. Hence each scientist (assuming they are rational) would want to choose the best theory at the time, leaving no one in the community to work on the lesser-known, untested, or unpopular theories that are available. This situation, according to Kitcher, is an irrational *community*, although made up of rational individuals. The

imbalance is a result of unequal distribution amongst the theories. The sullied group, however, would distribute themselves for want of fame rather than truth. Considering the probability that two separate theories have to solving a given problem in their field, the scientists in the sullied group would spread out between the theories so as to maximize each of their chances at solving the problem and thus reaping the fame. Without consideration of truth, the sullied group would ironically be better and faster at achieving progress in science than the pure group, which was filled with individual scientists who were aiming for the truth all along.

A. *The problem of theory choice*

Kitcher begins with an example from the history of chemistry: the revolution that occurred in the community's switch from phlogiston to oxygen. Although there is no way of determining that crucial point in time that definitively switched the community allegiance from phlogiston to oxygen, the Lavoisier experiment in 1787 was as near as possible to being a turning point. Kitcher suggests that before the experiment the degree of corroboration was 0.51, for example, but after the experiment the degree shifted to 0.49, keeping in mind that the spread of word takes time. After that point, it was not long until all chemists were following Lavoisier. This "sudden jumping of ship" is precisely what Kitcher finds dangerous to the rationality of the community (5). Although each chemist made an individually rational decision, the rationality of the group was risked by such an abrupt shift in allegiance from one theory to the other at a time when evidence to favor a theory was

still limited, rather than distributing the scientific labor between them to lead to the truth. "A community of chemists that responded in the fashion of [this] story is a badly-run community - an irrational community, if you like" (6).

To address concerns one may have with his story and analysis, Kitcher makes a few brief points. To begin with, theory choice in science is not "clear-cut," so one cannot say at the time when rival theories are at work which one will prove itself more useful. In addition, it is impossible to give a numerical assignment for the amount of support that a given theory has within a community. A bigger criticism is that in science, there are often individuals who do not agree with the community's assessment for which theory is best - there will always be those who refuse to go along with the crowd, for one reason or another, and who do not judge a theory and make a theory choice along the same lines as the rest of the scientists in the community. But for what Kitcher is proposing, he stays away from the issue of what type of criterion is used by different scientists and why two or more "rational" scientists would judge a theory differently, even on the same information, and focuses on the fact that "we sometimes want to maintain cognitive diversity even in instances where it would be reasonable for all to agree... and we may be grateful to the stubborn minority who continue to advocate problematic ideas" (7).

For example, in the early twentieth century, Alfred Wegener was having difficulties with the geologic theory of continental drift, such as the problem of not being able to explain where the vast amounts of energy required for the theory could

possibly be coming from (7). It was only through the distribution of the geologists as the time, with scientists such as Alexander du Toit sticking with the continental drift theory, that such a theory received a fair chance at explaining a geologic phenomenon. Was it reasonable or rational to be following such a theory at the time? Today it is easier to say "yes," knowing what has been learned since that time in geology. In addition, Kitcher argues that from the community's perspective it would have been best had the scientists and other resources been more evenly divided from the beginning.

Kitcher mentions the distinction between a belief in a theory and the pursuit of the theory to possibly help in spanning the distance between individual and group rationality. While it is fully possible for a scientist to believe in the best-corroborated theory, and a rational individual most likely will, it does not imply that the scientist must pursue the same theory. While believing one theory true, the scientist can work on a less-corroborated theory - leaving both the individual and the community rational. "What the community cares about is the distribution of pursuit not the distribution of belief" (8).

Sarkar also makes a distinction to help span the gap between individual and group rationality: the split between the theoretical, practical, and heuristic components of a method (the problem of heuristic advice introduced by Lakatos). Given any scientific method, it should provide the scientists with all three forms of advice. The theoretical advice provides for the belief of the scientists - it determines which theory at that point in time is best-corroborated. For a specific problem, the practical advice suggests

which theory is appropriate for its solution. The heuristic advice accommodates for the pursuit of the scientists, for it suggests which theory or theories at the time have promise for leading to better theories in the future. A method's heuristic component offers the theories that need development - whether they be strong or weak in the rank-ordering provided by the theoretical advice.

So Kitcher's distinction between theoretical belief and theory pursuit is nothing new. Such a division has been suggested and recommended before in the philosophy of science, and it has been recognized as an important distinction for the advancement of science. But one might ask if a sullied group is necessary for producing such a distinction in reality, for a pure community may also discern the difference between belief and pursuit and act accordingly in its research.

B. *Choosing a method: the pure*

Kitcher sets up a hypothetical situation in which, with unerring judgment as to the value of given theories, a philosopher-monarch is able to set up a community of scientists from the beginning. This monarch must decide how such a community ought to be run. Suppose he lets the individuals operate on their own rationality within the group. Kitcher claims that this, as with the earlier example from the history of chemistry, would not create an ideal situation for the advancement of science, for "[it is] liable to promote uniformity of opinion when you would prefer to keep your options open" (9).

There is no philosopher-monarch conducting how science should operate, however, but the same questions can be considered when dealing with individual scientists. Kitcher differentiates between two epistemic intentions of scientists: the personal and the impersonal. The personal epistemic intention of a particular scientist may be, for example, a quest for knowledge or the truth in regards to the issue he is working on. But the scientist may also have an impersonal epistemic intention, such as a desire that the community as a whole also work on the best theories in a search for knowledge to be passed down to future scientists. According to Kitcher, the scientist should conduct his work in regards to the latter intentions, that is, work towards the progress of the communities efforts rather than pursuing his own interests or beliefs. He defines such a person as an "altruistically rational scientist," one who sacrifices his own personal goals so as to work on a project or theory that he considers inferior to his own beliefs in the aim of promoting knowledge for the entire group in the long run (9).

Abstractly, given a community of scientists, S, there is a set of cognitive objects such as rival theories, methods, or programs, R (for ease of explanation, let us assume R consists of rival theories). Every member in S has an impersonal epistemic intention that S achieve some particular goal; we will choose one and label it G_i (although these goals may vary from individual to individual, we will suppose that each scientist shares some objective goal G, to help simplify). In addition, each member of S has an evaluation function that operates over the set R and a "cognitive attitude," such as a belief in a theory chosen by the scientist's evaluation function. The

IR (individually rational) distribution is the span of attitudes across S in regards to individual rationality. The CO (community optimum) distribution, however, is the span of attitudes over S that would maximize the community's chances at achieving G. A CO-IR discrepancy arises when the CO distribution has a higher probability for achieving G than the IR distribution.

Kitcher presents an example from the race to discover DNA (or "VIM": very important molecule) to help illustrate the above abstractions. There were two methods available to the chemists to aid them in discovering the structure of VIM: either by X-ray crystallography to study the bonds, or by constructing tinker-toy models (11). It was rational for the individual scientists to believe that the first method was more useful and likely at producing the structure, so all of the members of the chemistry community worked on this method.

Since the community wants to understand the structure of VIM as fast as possible, Kitcher sets up probability functions for each method. The probability (or return) function $p(n)$ represents the chance of a method producing the structure of VIM if n scientists are working on that method over some time t. Where N is the amount of members within the community S, the ideal situation is to assign n scientists to the first method and $N-n$ to the second. Hence, to maximize the probability that the community discover the structure of VIM, is to maximize

$$p_1(n) + p_2(N-n) - Prob_{1+2}$$

where $Prob_{1+2}$ is the probability that both methods find the solution (for ease of computations, assume $Prob_{1+2} = 0$).

Of course, being separate methods, their probability or return functions will also vary according to how fast they respond to effort applied by the scientists from the beginning, or if their rate increases or decreases over time, effort, and the number of scientists entering or leaving the method. Kitcher considers two specific return functions. First, consider the situation where the method's rate of progress increases rapidly at the beginning, with fewer scientists, and slows down as more researchers enter its subgroup[1]. Ideally, a division of labor would be well-suited for a community with a large number of members (N large), a method that quickly responds to effort (k not too small), or if the difference in probability between the two methods is minimal. The second possible situation is one where there is slow growth at the beginning, rapid increase in progress once the group has reached a critical mass of workers, and again a slow-down as the field is saturated with workers[2]. As new members enter the community, they may join which ever method lacks a sufficient amount of scientists, or the method that is not saturated with members to maintain a rational community. If both methods have a sufficient amount of researchers, then new-comers work on the method with the highest chance of delivering a solution.

In both cases above, k is the method's responsiveness to effort. If the methods respond quickly to effort and the intrinsic prospects of both can be realized by the amount of workers

---

1 Suppose $p_i(n) = p_i(1-e^{-kn})$. Then $p_1(n) + p_2(N-n)$ is maximal when $n = [(kN + \ln p_1 - \ln p_2)]/2k$. Note that we want $\ln p_1 - \ln p_2 < kN$ so there is a distribution of labor (in other words, $n<N$).

2 Suppose $p_i(n) = p_i((3n^2-2n^3)/kn]/k^2N^2$ for $n<kN$, and $p_i(n) = p_i$ for $n \geq kN$

presently available ($k<1/2$), then division of labor is ideal. If given the amount of scientists it is possible to realize the intrinsic prospects of one method but not both ($1/2<k<1$), then Kitcher claims it may still be best to divide labor and risk not realizing the intrinsic prospects of either method, provided the difference between p_1 and p_2 is small. But in any other situation where $k>1$, it is best to devote all work to the method with the greatest intrinsic promise.

So how or where do differences arise between IR and CO distributions? If individual epistemic rationality chooses a method according to its promise or ability, chances are there will be no distribution between methods, in other words, yielding a IR-distribution of $\{N,0\}$, where the CO-distribution would rather have members working with both methods. But Kitcher suggests looking at individual rationality from another perspective. Suppose a scientist is more interested in maximizing his personal chances at discovering the structure of VIM. He will then either choose the method he finds most-likely to discover the solution without considering the present distribution of the community members, or he will consider the distribution of the other members, and choose the top method at the time only if the probability of discovery in that method is still greater than the probability of the other method.

Either way, a community composed of such members could end up entirely in the same subgroup, resulting in no division of labor and an obvious CO-IR discrepancy, since they may all decide to work on the top method of the time (simply for it being the top method) or if the other method cannot provide enough probability to seem worthy of an effort. Kitcher suggests that the problem with

such a pure community is that it requires scientists to be altruistic in their work. Hence, he offers what he considers a more realistic scientific community: the sullied.

C. *Choosing a method: the sullied*

So what would be a better option for creating some distribution of labor between these rival methods that are aimed at discovering the structure of VIM? Kitcher introduces the neighbors: the sullied community. The sullied community is comprised of "ruthless egoists" interested in winning a "much-coveted prize," and who are rational in that their actions maximize their chances for winning the prize rather than achieving any idealistic goal of truth or knowledge (14-5). Within the sullied community, each individual has the personal goal to discover the structure of VIM since he personally wants the prize, and each member working in the community is interested in the VIM *because* of this prize that is attached to the discovery. For any member of the community, the probability of him making the discovery in his particular method, supposing the method he is working on will lead to the VIM structure, is equal to the chances of the other members working within the same method (ignoring differences in abilities and interests of individuals to help simplify the example).

What can motivate a scientist to switch from one method to another? Rather than working on the most-promising method of the time, a scientist in the sullied community would take the present distribution of researchers into account. If a scientist in method I were considering switching over to method II, he would want to

know if his chances of finding the structure of VIM were greater in the second method than the first. Analogous to the chances of a lottery ticket winning, the probability of the individual's success is the probability of the method succeeding divided by the number of members working in that method (since we supposed the members have equal chances). In other words, if the current distribution of the community is $\{n, N-n\}$ out of N total members ($n<N$), the scientist would increase his personal chances by moving from method I to method II if

$$\frac{p_2(N-n+1)}{(N-n+1)} > \frac{p_1(n)}{n}$$

To find out where the sullied community is at equilibrium (the point where it is not beneficial for any scientist to move from one method to the other), Kitcher first defines some terms. A community distribution $\{n, N-n\}$ is stable/unstable upward or downward if the following conditions hold:

	stable	unstable
downward	$\frac{p_1(n)}{n} \geq \frac{p_2(N-n+1)}{(N-n+1)}$	$\frac{p_1(n)}{n} < \frac{p_2(N-n+1)}{(N-n+1)}$
upward	$\frac{p_1(n+1)}{(n+1)} \leq \frac{p_2(N-n)}{(N-n)}$	$\frac{p_1(n+1)}{(n+1)} > \frac{p_2(N-n)}{(N-n)}$

In addition, the distribution $\{n, N-n\}$ is bilaterally stable if it is stable both upward and downward (15). In the above example, it would be to the scientist's advantage to switch methods if the current distribution of members were unstable downward.

Kitcher points out the difference between stability and attainability of the distribution of labor. Although once a group is at equilibrium it may be easy to maintain stability, attaining that stability with a community of egoists may not be so simple. Kitcher defines some terms that will help in explaining a distribution's attainability. First, given an unstable upward distribution $\{x, N\text{-}x\}$, $\{m, N\text{-}m\}$ collapses upward to $\{n, N\text{-}n\}$ for all x such that $m \leq x < n$. On the other hand, an unstable downward distribution $\{x, N\text{-}x\}$, $\{m, N\text{-}m\}$ collapses downward to $\{n, N\text{-}n\}$ for all x such that $m \geq x > n$.[3] The zone of attraction for any distribution $\{n, N\text{-}n\}$ is the set of all distributions $\{m, N\text{-}m\}$ that collapse to $\{n, N\text{-}n\}$. So a bilaterally stable distribution $\{n, N\text{-}n\}$ is attainable if its zone of attraction contains all possible distributions. In other words, given any presently existing distribution of a community, if that distribution falls within the zone of attraction for a bilaterally stable distribution (ideally the CO-distribution), the community will tend toward stability. For Kitcher, the community will more likely fall within the zone of attraction if the scientists are motivated by selfish reasons and therefore divide their labor between the two groups to begin with so as to maximize their chances.

Kitcher suggests that attaining stability with a sullied community, although difficult, is more likely than with the pure community. "The very factors that are frequently thought of as

---

3 In terms of mathematical limits, a scientific community's distribution is unstable upward if $\forall\ m \leq x < n$,

$$\lim_{(m,\ N\text{-}m) \to (n,\ N\text{-}n)} (x,\ N\text{-}x) = (n,\ N\text{-}n)$$ where $\{n, N\text{-}n\}$ is any bilaterally stable distribution.

The same goes for an unstable downward distribution, but $\forall\ m \geq x > n$.

interfering with the rational pursuit of science - the thirst for fame and fortune, for example - might actually play a constructive role in our community epistemic projects" (16). In short, a scientific group may take advantage of the individuals' personal desires in the pursuit of the goals of the community, rather than considering such motives as an encumbrance.

For example, the sullied community advances toward its goals by appealing to the scientists' want for a prize and thus forces a division of labor in which each member works to maximize his chances. Kitcher suggests that other personal incentives may also be used by the community to motivate its members, such as personal or political loyalties. He does not, however, clarify the means by which the community should use such incentives.

Is it actually feasible for a scientific community to achieve an ideal distribution of labor? Kitcher reuses the earlier case of rapid progress that slows down as the community reaches the saturation-level of workers, where the probability function $p_i(n)$ is the probability of method i producing the structure of VIM with n scientists, such that

$$p_i(n) = p_i(1-e^{-kn}) \qquad \text{where } k \text{ is large and } p_1>p_2>0$$

Kitcher claims that there is an attainable, bilaterally stable distribution in the neighborhood of $\{[p_1N/(p_1+p_2)],\ N-[p_1N/(p_1+p_2)]\}$. If the difference between the probabilities (p_1 and p_2) is small, the actual distribution of scientists tends toward the CO-distribution. In summary, "there are conditions under which the Hobbesians [or sullied community] do better than their epistemically pure cousins, even conditions under which they come as close as you please to the

ideal" (16). Both methods are given a fair chance and the community progresses as fast as possible toward its discovery.

The case above is a simple formulation, for somewhat ideal circumstances of experimental work in science, since theoretical work would pose more problems for distributing labor and in setting certain community goals. But what if there is no distribution between methods in such a situation, where if the probability of a second method were too low and the community itself were so large (where kN is too big) that whether one scientist is contemplating a move from the top method to the lesser one would make little or no impact in that lesser method's chances? Kitcher admits that "when the community is too big, self-interest leads the community to the same sub optimal state as individual rationality" (17).

Since one scientist may make an insignificant difference in such a case, Kitcher suggests breaking down the scientific community into subgroups, or fiefdoms. These fiefdoms, for example, may be laboratories headed by a local chief (the lab director) who makes the decision whether or not to switch methods, and upon his decision the peasantry (graduate students and assistants) follow. So if within the subgroup of q scientists there are x members (where $x \leq q$) willing to change methods and "jump ship together," then the community as a whole may attain stability provided that from that one subgroup $q > kN$. In brief, "a certain amount of local autocracy - lab directors who can control the allegiances of a number of workers - can enable the community to be more flexible than it would be otherwise" (17).

So far in this modeling of a sullied community, several variables have been left out (such as individual talent), and many assumptions have been made (assuming, for example, that only one of the methods rather than both may succeed) to make the model intelligible. But one problem still remains. What if the sullied community is not fully "sullied"? In other words, what happens when *not* all the scientists within the community are out for the prize or selfish means? It is hard to imagine that such a uniform sullied community could actually exist. How could the community count on manipulating such traits if some of its members had no desire for a prize? In other words, is Kitcher relying too heavily on personal incentives for individuals, and specifically on a uniformity of foibles (where foibles usually tend to be unique)? In addition, how can Kitcher assume that everyone in the community would be able to calculate the probability of their success in one method versus another?

First, Kitcher responds by pointing out that his supposition of the scientists' abilities to figure out the probabilities is equivalent to the assumptions made in traditional theories in the philosophy of science. In response to whether the whole community could be selfish, Kitcher claims that a scientific community may succeed (better or worse than this one, depending) on a number of personal traits, be they virtues or vices. "Not only may vices from greed to fraud play a constructive role, but... perseverance, personal investment, personal and national loyalties, and devotion to political causes may, on occasion, help to close a CO-IR discrepancy" (18). In other words, the rationality of the pure community was not enough -

such psychological incentives motivate scientists to work further than they would without such extraneous reasons.

After seeing how such a community of scientists would decide on the division of labor between methods, how would this sullied community divide labor between competing scientific theories? Before presenting a model, Kitcher points out two difficulties: the change in probability of a given theory's epistemic promise, and that if a theory grows in ability that it will approach the truth. But with some assumptions, Kitcher continues his theory of group rationality with an hypothetical case of rival theories.

Given two rival theories from the same point in time in the history of science, T_1 and T_2, suppose that the probability of T_1 being true is q_1, where $q_1+q_2=1$ and $q_1>q_2$ (although the probabilities of the two theories are close). Moreover, everyone in the community is aware of these probabilities. The community as a whole has the aim of finding which theory is true, eliminating any problems of that theory, and elaborating on both its theoretical and practical advice. Given this community goal, there are two possibilities for the community: (A) the scientists can all work on T_1, or (B) they can divide labor such that n scientists work on T_1 and $N-n$ scientists work on T_2.

Kitcher proposes looking back on the research situation of the community from "the time of reckoning," or the conclusive state, where one of the theories is finally known to be true (19). In the most ideal of situations, all of the scientists would have been working on the true theory and the group would have compiled a utility of u_1. Had the group wasted time on the other theory, however, they would

end up with a utility of $-u_1$. After working on both theories, the community reaches a conclusive state only if one of the theories, say T_1, has resolved all its anomalies while the other one has several anomalies remaining. Assume for simplicity that once a theory has resolved its problems that it is true (no "false positives"). If the community reached the conclusive state through a division of labor, then the group's epistemic utility is u_2, where $0<u_1<u_2$. Otherwise, the utility is 0, and "we are still in the same predicament, although our labor may have given us a clearer view of the problems that each of the rivals faces" (20).

If the community used approach (A) and all members worked on T_1, then the group's epistemic utility can be written as

$$q_1u_1 - q_2u_2$$

In other words, if T_1 is true (q_1=1, q_2=0), then the utility is positive (u_1), and if T_2 is true, then the utility is negative ($-u_2$). To write the utility equation for approach (B), where the community reaches a conclusive state supporting T_1, two matters should be taken into account: (a) T_1 must resolve its present anomalies, and (b) a sufficient amount of scientists must be assigned to T_2 to provide both theories with a fair chance. Let the probability that T_i responds to the work of n scientists assigned to it be called $p_i^*(n)$. Furthermore, suppose n scientists are working on T_1, and $N-n$ scientists are working on T_2, where $N-n>m$ for some m. If m is large enough such that T_2 is given a fair chance, then the probability of matter (b) is 1; if not, then the probability is 0. Finally, the equation resulting from matter (a) is the probability that T_1 is true times the probability that

T_1 responds to effort, $p_1^*(n)$. So the probability equation for the utility of approach (B) is

$$q_1 p_1^*(n) u_2 + q_2 p_2^*(N-n) u_2$$

where $m<n<N-m$. The division of labor, approach (B), is better than approach (A) if there is an n where $m<n<N-m$ such that

$$[q_1 p_1^*(n) u_2 + q_2 p_2^*(N-n) u_2] > [q_1 u_1 - q_2 u_2]$$

Suppose, for example, we choose the probability function from earlier where $p_i^*(n) = p_i(3n^2-2n^3/kN)/k^2N^2$ for $n<kN$, and $p_i^*(n) = p_i$ for $n \geq kN$. For simplification, suppose $p_1=p_2=p$ and $k<1/2$. Accordingly, both theories have been given a chance if n is within the interval of $[kN, (1-k)N]$. Again, division of labor is the best approach if

$$pu_2>(q_1-q_2)u_1$$

So unless the probability that a true theory can displace its anomalies is low, or the utility u_1 of working directly on the true theory (albeit unknown at the time) is higher than u_2, to achieve the CO-distribution means the community must divide its labor.

So how does a sullied community approach the ideal CO-distribution? Suppose the community is already divided into a distribution $\{n, N-n\}$, where $kN<n<(1-k)N$ for $k<1/2$. Then there is an attainable, bilaterally stable distribution $\{[q_1N/(q_1+q_2)], N-[q_1N/(q_1+q_2)]\}$ if $k<q_2$ and $q_1+q_2=1$. In other words, there are situations where, by exploiting the personal incentives and drives of the scientists, the IR-distribution converges toward the CO-distribution.

Kitcher concludes with some further issues that can be extended from his above suggestions. For one, there are personal

motivations other than the drive for fame and fortune that could also lead to the same results, such as strong national ties or personal alliances. Also, in his examples Kitcher assumes that the scientific communities have a fixed number of scientists N, whereas in reality there are constant shifts in number within a field of science and between fields of science. Kitcher suggests that such inter field shifts may lead to distributions that could, like the distribution between methods or theories, either benefit or handicap the world of science. Although such division of labor could happen on a grander (inter field) scale, there are greater barriers, such as the need for retraining.

What is the optimal distribution of labor across and within scientific fields? How do personal incentives and motives affect these distributions: do they lead the scientific community closer to or farther away from the community's goals? Kitcher proposes that these personal factors can be manipulated to the benefit of the community, rather than holding unattainable, idealistic values that such a community must achieve or strive toward utopian ideals to advance science. "How do we best design social institutions for the advancement of learning? The philosophers have ignored the social structure of science. The point, however, is to *change* it" (my emphasis, 22). Maybe so... but what kind or direction of change is in order?

Chapter IV

In summation: *criticisms and proposals*

> *When the late war ended in a thunderclap, it left two noteworthy developments in its wake. Science had become politically interesting, and scientists had become interested in politics.* - J. H. Rush, 1947

This statement was made in the wake of World War II - at the time when the world was embarking on the Nuclear Age. As a closed community of scientists, the members of the Manhattan Project worked on the discovery of fission that eventually led to the creation of the atomic bomb. In today's society more than ever before in the history of science, the power of science and technology is leading to serious moral and social issues. The cliché, "knowledge is power" has much truth to it, and yet what kind of power is being developed and is the world ready for it? Also more than ever before, society is turning to science for answers to its problems, and in turn is receiving more problems that need answering. In the past, society turned to religion for such answers. Now science has become a quasi-religion in itself.

Newspapers today are filled with issues never considered before this century, or even before the last fifty years. Issues such as the AIDS epidemic, rights of embryos, the biological differences between homosexuals and heterosexuals, invitro fertilization, and computer hacking are just a few of such topics that have entered the world scene. For example, society looks to science to find the cures for AIDS, cancer, and other diseases. But as cruel as it sounds, can society afford longer life spans and how? It seems disheartening to

consider such side effects on society, but then issues such as euthanasia among the elderly become the next "hot topics." Our political and justice system can hardly keep up with the speed of technology.

The world of science cannot be abstracted from the whole of society: science and society are closely interwoven. Is the role of science in society simply one of providing technology? Although certainly not an easy task, is the advancement of science (to reword a phrase by Abraham Lincoln) *of* the scientists, *for* the scientists, and *by* the scientists? Scientists can no longer live in ivory towers, for many of them are involved in projects of profound significance (good or bad) to society. Even universities now offer ethics courses designed specifically for medical students and engineers.

Kitcher suggests that a scientific community be fashioned around the scientists' psychological factors. Namely, that to achieve a division of labor within a group of scientists, the group can and should manipulate the desires and personal ambitions of the individual group members, as with his primary example of a group of egoists, or the "sullied" group (although other personal factors could be manipulated just as well). This approach certainly is plausible: that in the fury to discover some theory or solution to a problem the scientists are driven by the want of fame or a prize. The greater the social issue is that is attached to the problem at hand, the greater the fame and reward. Extol the scientist or group of scientists who will find the cure of AIDS, for society will, with good reason, be grateful for their accomplishment. But does society look to science for only such cures and answers to questions of science?

Are moral issues outside of the field of science, and should they be? In the words of J. R. Oppenheimer from 1963,

> In these times, in these years, the atom bomb and nuclear weapons preside over our anxieties. This is an accident. It was, of course, done by design, but it was an accident that it could have been done when it was.... We can only hope that [these weapons] will increasingly appear irrelevant and thus in the end preposterous, that some day we will look back ashamed of how stupid we were.

Should science slow down or take a step backwards? Of course not: progress has its advantages and disadvantages, and society can benefit from both. But should science be reckless in its efforts? What kind of a society should the scientific community be?

A. *An historical model*

By the beginning of the twentieth century, atomic physics was recognized throughout the world of science as one of the newest and most-promising fields. Scientists theoretically understood the power that was possible, associated with the use of atomic physics, and talk about the direction of this power to be used as a weapon was certainly underway by the 1940's. The predominant factor involved in the development of the atomic bomb was the pressure as well as encouragement by the governments in the race to invent the weapon. Certainly it may have been created at a later point in time, but the timing was influenced mostly by the government: in providing endless money, resources, and collecting the top scientists

available in the world in an isolated scientific community at Los Alamos. As one historian comments, "World War II also drew forth a clash of scientific minds. In their respective laboratories, the Allied scientists raced their German counterparts for a wide variety of technical secrets" (Szasz *xvi*).

In the 1930's, popular talk in the news suggested what was then considered science fiction: the harnessing of the energy in an atom. In 1938, atomic physics witnessed the splitting of the uranium nuclei into (approximately) equal pieces after bombardment by neutrons, named "fission" after the biological process of cell-splitting. Several questions arose in the field of nuclear fission that scientists eagerly sought to answer, such as which isotope of uranium was the most likely to split (^{235}U), could a chain reaction result from the nuclear bombardment, and if so could this energy be controlled, or could it only be released in the instantaneous power of a bomb? There also arose the more practical questions for weapon development, such as how much is necessary to produce the amount of energy needed for a bomb (so that it could be transported by plane). In 1938 and 1939, hundreds of scientific findings regarding such issues were published openly in international scientific journals and magazines, and in that time state-run laboratories in the United States, England, France, Germany, the Soviet Union, and Japan began to research the probabilities of developing a weapon with such energy (Badash *xii*).

Los Alamos was chosen to be the site for the Manhattan Project (so named since most of the earlier work was done in New York) in November 1942. Located near the Jemez mountains in New

Mexico, Los Alamos was ideal for both the privacy and security it provided for the scientists and their work. All of the members, the majority being American, British, and Canadian scientists, lived on what they called "the Hill," located on the grounds of Los Alamos near the laboratories. Only two of the scientists left the community half way through the project, for as one member said, "once people entered the gates of Los Alamos, everyone assumed they would stay until the end of hostilities" (Szasz 58).

In 1943, there were twenty scientists stationed at Los Alamos, along with their wives, children, and other workers such as technicians, teachers, and janitors. The director of Los Alamos was J. Robert Oppenheimer, whose major concern for the project was that the scientists have access to as much information as possible within the community, hence the need for such extreme isolation. The overall head of the Manhattan Project was General Leslie R. Groves, and as one historian remarked, "Groves ran the sprawling Manhattan Project as his personal *fiefdom*" (my emphasis, Szasz *xvii*).

So what motivated the scientists in their pursuit of discovery? As one of the scientists, Richard Feynman (just out of graduate school at the time), later commented: "You see, what happened to me - what happened to the rest of us - is we *started* for a good reason, then you're working very hard to accomplish something and it's a pleasure, it's excitement. And you stop thinking, you know; you just *stop*..." (Badash 132). Certainly there were the scientists who "lived totally for the world of science" (Szasz 38). The British Mission present at Los Alamos was also driven by the personal and political motive to stop the Germans, or at least beat

them to the discovery, after witnessing the German bombings on England firsthand. As one British scientist remarked, "we thought that once the bomb had been demonstrated, the world would see how terrible this weapon was and come to its senses" (Szasz 30).

There was also the drive for fame and recognition within their community. "Salaries were generally low.... yet, the reward system seems to have functioned after all, because the peers one cared about also were at Los Alamos and in a position to know of the contributions" (Badash *xix*). The isolated community had to include and provide for everything, from the scientific needs to the social life of its members: in their time off, the members of Los Alamos would go hiking in the nearby mountains, horseback riding, play music, hold political discussions and dances, and even play games such as "twenty questions." As one historian notes, "the *esprit de corps* they developed there played no small role in their success" (Badash *xx*).

One of the scientists that played a significant role in boosting morale was Danish physicist, Niels Bohr. In the 1920's, Bohr's Copenhagen Institute for Theoretical Physics had become a center for physicists from all across the world, and has been compared to Plato's school in Athens (Szasz 74). Since his mother had been Jewish, Bohr fled Sweden in 1942 to escape Hitler's invasion, and went to England in 1943. In fact, Hitler had the Institute searched with the hopes of finding work or research on the development of the atomic bomb left behind. Under an alias, Bohr and his son Aage visited the Manhattan Project at Los Alamos several times as "consultants," although they never lived on the Hill. Although Bohr did help in playing a technical role at Los Alamos as

well, his predominant role was in buoying the morale of the others. Bohr spent a significant part of his time during the war lobbying the governments and political heads of state to warn them of the dangers involved with the creation of such a weapon. Alice Kimball Smith, a historian who had been at Los Alamos, writes that "at Los Alamos... there was no systematic exploration of implications and alternatives and hence little psychological preparation. Toward completion of the bomb the whole laboratory had worked with an absorption that left little time for reflection" (76).

B. *The social structure of science at Los Alamos*

The purpose of this section is to further explore the *social motives* that influenced the scientists in their work at Los Alamos during the Manhattan Project. At possibly no other point in time has a scientific community been so laden with personal motives and incentives to achieve such a specific goal. Although many, if not most, of the scientists were interested in their work of atomic physics for epistemic reasons, other reasons such as politics, citizenship, family ties, religion and ethnicity are, just to name a few, influences that invariably affected the scientists' decisions to either participate in or to leave the work going on at Los Alamos. For regardless of which aspect of the project each scientist specifically worked on, the entire community was working, not to further theoretical understanding directly, although it was a byproduct of their efforts to create the atomic bomb.

How did such motives affect their work in developing an atomic bomb while at the labs in Los Alamos, or earlier in New York,

before the government moved the project? For one, such incentives had direct bearing on the location of the scientists at the time that the war broke out and at the time in which Oppenheimer was coordinating the set up of the labs at Los Alamos. In addition, political and national affiliations influenced at least one of the scientists on the British Mission, Joseph Rotblat, to leave Los Alamos before the end of the project. Another influence that affected the participation in creating the atomic bomb was the individuals' personal views at the time of what the purpose of the atomic bomb was, and several scientists were disheartened in their work on the project as they began to realize the futility of using such a weapon as a deterrent.

How did the scientists view their motives and their work once the bomb was created? What was their opinion of what had happened, once the bomb was dropped in Japan? While some of the scientists had a sense of responsibility from the beginning, others (such as Richard Feynman) admitted that they did not grasp the implications of their work until it was done, and the fact that *science was now a dangerous tool* that could be used by politics awoke in them.

Of all incentives involved, it seems natural that the Jewish scientists involved in the British and American efforts to create an atomic bomb would be highly motivated to help in the development of the weapon to stop Hitler. Many Jewish scientists moved to England or the United States in the years before the war, as well as other intellectuals and scientists as they fled Germany, Italy, and the surrounding countries. Sir Rudolf Peierls, one of the Jewish

scientists who participated in the work at Los Alamos later recalled, however, that "it was largely a force of circumstance, not revenge against the Nazis, that drew the refugee scientists into the fledgling British atomic program" (Szasz 3). And as more and more scientists and intellectuals moved from England and the continent to the United States, there was a shift in the global "scientific center of gravity" away from Europe, the traditional stronghold of science for centuries (Badash *xiii*). Ironically, "Hitler's prejudices had not only deprived himself of talent of the highest quality, but had thrust it upon his future enemies" (Badash *xiii*).

One of the scientists that worked at Los Alamos was Joseph Rotblat, a Polish Jew who happened to be stuck in England when war officially broke out, so he could not return to his country. Rotblat was in Britain on a governmental arrangement to study and work with another atomic physicist, James Chadwick, in Liverpool. Once the war had started, Rotblat began working on the British atomic bomb project with Chadwick, yet already Rotblat expressed doubts about the morality for using such weapons of mass destruction. But when Poland fell to Germany, Rotblat devoted himself entirely to his work on the project. Tensions (and security problems) surrounded Rotblat once he was at Los Alamos, for he retained his Polish citizenship, unlike several other scientists who at the time were either attempting to get British or American citizenship if they were from one of the countries under Hitler's control.

Rotblat also began to doubt the idea of nuclear deterrence, for "nuclear deterrence worked only when both sides

operated on the same rational premises. The plan would be ineffective against a madman" (Szasz 56-7). He viewed the scientific situation at Los Alamos as ideal, yet Rotblat had two main concerns. Although he refused to publicly reveal most of his incentives for leaving Los Alamos until the fortieth anniversary in 1985, Rotblat's first reason was his constant concern for the safety of his family, since his wife and young son were still in Poland, and his brother was fighting in the Russian Army. Second, his doubts concerning the purpose of the project kept increasing. Rotblat finally requested leave from Los Alamos in December 1944 after a conversation with General Groves in which he understood Groves to claim that "the real purpose of making the bomb is to subdue our chief enemy, the Russians" (Szasz 57).

One of the families present at Los Alamos was the Fermis, from Italy. Laura Fermi, a physicist in her own right but not involved in the Manhattan Project at Los Alamos (her husband was in the Theoretical Division), said years later: "If I look at the personal events that took me to Los Alamos I find that I can choose two starting points. One is scientific and the other is political" (Badash 89). The scientific date was in 1934 when Enrico Fermi, her husband, participated in the achievement of artificial radioactivity by bombardment of elements with alpha particles in his lab in Italy. The political date, however, was in October 1922, when the Fascists brought Mussolini to power in Rome. Although they did not move to the United States until 1939, the Fermis considered themselves as part of the intellectual and scientific wave of migration from Europe into to the new world in the 1930's.

Life at Los Alamos was very different than what most of the families or scientists were used to. Housing, as well as maid service, was assigned according to size of family or need, rather than status. A group of old ranch school buildings were turned into houses, and were termed "bathtub row" for the desired amenity that all other housing on the mesa lacked. Meanwhile, the other houses at Los Alamos were mostly army huts, while bachelors lived in the dormitories. Despite Los Alamos being a community of scientists working on such advanced technology, most houses were lucky if they had wood-burning gas stoves for cooking, called "Black Beauties." Yet regardless of the rustic and simple living arrangements, or more likely *because of* such arrangements, the community on the Hill "shared the esprit of a battle unit. The political discussions at dinner and afterwards frequently became intense. Here the scientists and their families debated the probable political fate of postwar Europe" (Szasz 35). One scientist recalled the discussions as stimulating, but as Peierls additionally noted, "there was not much else to do" (Szasz 35).

Another social concern at Los Alamos was the ever-present cultural difference between the Europeans and the Americans. The members of the British Mission enjoyed the arrangements at Los Alamos because no one made distinctions by rank or class, unlike what they were used to back in Europe. One example given by a member at Los Alamos was when Robert and Kitty Oppenheimer joined a long line outside of a popular movie, rather than calling rank. As Peierls wrote, "it is an enormous pleasure... to be at a place... where work is guided by the necessity to

get the best answer in the shortest possible time rather than by questions of formal organization and prestige" (Szasz 41). Yet due to their distinct cultural differences, the British were still viewed by the Americans as guests at Los Alamos, more than permanent residents to the establishment.

Despite the close interaction between the British and American scientists, tension still existed. In 1942, the United States government asked Oppenheimer to join the development of the atomic bomb. That summer, he and a group of theoretical scientists (mostly from California) did an intensive study on the properties of nuclear explosions, and by early 1943 Oppenheimer was director at Los Alamos and beginning to set up the labs. But the British Mission did not begin to arrive until December 1943 due to political tensions between the two countries up until that time.

Before war broke out in Europe, the United States distrusted Britain and its intense interest in atomic weapons, even though England admitted that it needed such weapons in preparation for postwar existence in Europe, if it was to retain any power. In August 1943, England and the United States signed the Quebec Agreement, which established the nuclear relationship between the two countries for throughout the war. In other words, the United States and England agreed to never use the weapon either against each other or against a third party without mutual consent, as well as to not convey information on the weapon to a third party without such consent. In addition, the Quebec Agreement stated that the two countries should combine their efforts on the development of the atomic bomb, so the British scientists were sent to the United States,

for safety and security reasons. Despite such official agreements, however, distrust occasionally ran through the members of Los Alamos.

Despite the tight security and army checks, information at Los Alamos did manage to escape. Perhaps the most infamous spy case at Los Alamos was that of Klaus Fuchs, a Soviet agent who was responsible for the greatest leakage of information from Los Alamos. Due to recent declassification of FBI documents, more information is available on his role at Los Alamos than any other of the "Atomic spies" (Szasz 82).

Fuchs was from a religious family in Germany. His father, who was an active Quaker, was sentenced to a concentration camp, while his sister committed suicide during the war due to Nazi political harassment. While studying at the university in Germany before the war, Fuchs became active in communist circles. In 1933, Fuchs moved to Britain when the Nazis took over Germany, and became a naturalized citizen of England in 1942. While in England, Fuchs made little attempt to hide his affiliations with the communists. During his work on the British atomic bomb project, Fuchs got in contact with the English Communist Party, and through a courier he informed the Soviets on his own work because "he felt that the Allied Powers had deliberately planned to allow Germany and Russia to bleed each other to death" (Szasz 87-8). Fuchs was transferred by the British to continue his work in New York, and he moved out to Los Alamos in August 1944.

Fuchs was not exposed as a spy until four years after he left Los Alamos. "In 1948, the FBI discovered that somebody had

supplied the Soviet Consulate with a top-secret scientific report... as well as other Manhattan Project data. Soon they became convinced that the KGB had had an agent within the British Mission" (Szasz 82). The Soviet Union had already been working on the development of an atomic weapon also, and in 1949 the Soviets detonated their first atomic bomb, beating England as they become the second world nuclear power. It is estimated that the information Fuchs supplied to the Soviet Union sped up their discovery by possibly one and a half to two years.

The question that arose was, of course, how did he manage to transfer so much information. All of the members at Los Alamos were surprised, or even "dumbfounded," that Fuchs had been involved in such a vast transfer of information. Feynman, who lived near Fuchs in the dorm at Los Alamos and with whom Fuchs discussed politics often, admitted that he had had no idea of what Fuchs had been doing. Ironically, the two men had a discussion one time about who was more likely to be a spy, and they both agreed that Feynman led a more suspicious life, since he made frequent trips to Albuquerque in Fuchs' car to visit his wife who was dying of tuberculosis. But in looking back on their years at Los Alamos, no one admitted to really knowing much about Fuchs at all, since by nature he was quiet and reserved. One may even question those who have claimed that they thought Fuchs was suspicious, for their reasons are usually based on their judgments of his personality rather than any actions or reasonably incriminating information.

Although Fuchs' name is so highly tied with the espionage, he was one of the top scientists at Los Alamos. Most of his

scientific contributions were in the Theoretical Division, where he made valuable contributions on blast-related issues, which played a "vital role in perfecting the implosion process... [since it was] the most difficult aspect of weapons production" (Szasz 92). The head of the Theoretical Division claimed that "[Fuchs] was one of the most valuable men in my division... one of the best theoretical physicists we had" (Szasz 89). Hence, because his work was so impressive, Fuchs gained access to a lot of the information on the Manhattan Project.

How does this example tie into Kitcher's model of personal incentives? One possible explanation for Fuchs' outstanding work may have been to achieve such trust and recognition within Los Alamos. The community of scientists encouraged its members to produce ideas and to solve problems related to creating the bomb. As one who obviously made contributions to the group, Fuchs gained access to more information. Oppenheimer wanted an open community of scientists in which knowledge was shared with all members, especially those members who could use the knowledge available and expand on it. It is possible that Fuchs was driven to make such contributions at Los Alamos knowing that his "reward" would be further access to information, which in turn he could send to the Soviets. In fact, it may be hard to take such political incentives *out* of one's understanding of Fuchs and his accomplishments.

But looking back on the situation, General Groves was annoyed that the information had been so open and available for all of the scientists involved. "[Fuchs] told the Russians exactly how to

assemble the bomb; how to calculate the yield of the bomb and the neutron diffusion, and he certainly told them about the critical mass... which is the most important for knowing the size of the reactor you must build" (Szasz 92). But despite the attack on the openness of the scientific community, Oppenheimer felt that Fuchs could have done enough damage from within his division alone, since it was responsible for many of the contributions made to the construction of the atomic bomb.

In 1959 Fuchs was released from a British prison and moved to East Germany to accept a position at their Central Institute for Nuclear Physics. Fuchs publicly admitted that he feels no regret for his actions at Los Alamos or before, since he felt that the governments were wrong in not including the Soviet Union in their efforts. In 1977, Fuchs called the development of the American neutron bomb "an abominable example in the misuse of science" (Szasz 96).

But Fuchs was not alone in sensing a misuse of science by the political heads-of-state. After the creation of the atomic bomb and the ending of the war, several British scientists remained at Los Alamos for a few more months, where they "watched the American political battles over postwar atomic control with interest" (Szasz 47). Notwithstanding the previous agreements between the United States and England, the United States still distrusted handing over the necessary information to Britain concerning the construction of the atomic bomb.

In 1946, two atomic tests were conducted in the South Pacific on or near Bikini Island, called "Operation Crossroads." The

first detonation, "Able," was postponed several times by President Harry Truman in order to coordinate the event with the schedules of various congressmen, and was possibly the most photographed event of the year. The testing done at Bikini Island has been described as a "raw display of power... [which] served *less to advance bomb technology* than to demonstrate its results to the world" (my emphasis, Szasz 47). In addition, the open society of scientific information that had been present during the years of the war came to a sudden halt when in 1947 the United States passed the McMahon Act, which denied any foreigner access to restricted information, *including their own reports*. From that point on General Groves denied the foreign scientists access to restricted data, regardless of the scientist's previous status or contributions at Los Alamos.

In hindsight, Laura Fermi claimed that the scientists at Los Alamos were "so involved with their work and under such pressure of time that they gave little thought to what later became known as the "social implications of the bomb"" (Badash 98). Even though the scientists on the British Mission enjoyed the scientific aspects of working at Los Alamos, once it was over with, many left with a sense of ambivalence. As members went back to their country, several refused to continue in the same work as they had done at Los Alamos.

For example, D.J. Littler declined a position to work on Britain's atomic weapons and went into energy production. James Tuck took a job developing commercial power from fusion rather than working on the creation of a hydrogen bomb. In addition, Tuck

suffered from reoccurring nightmares of running away from a mushroom cloud. Bohr regretted the creation of the atomic bomb in private, and Rotblat in public. In 1986, Peierls summed up many of the scientists' thoughts when he remarked: "I'm afraid, forty years ago, we over-estimated the capacity of those in power to understand the implications of what we had created" (Szasz 106).

The scientists at Los Alamos broke ground not only in atomic physics, but in international relations as well. The creation of the atomic bomb in the United States depended on the joint effort of the European (the British Mission) and the American scientists. Without such collaboration, the atomic bomb would not have been completed by either country in such record time. As Oppenheimer claimed, the development of atomic energy "lay in a field international by tradition and untouched by preexisting national patterns of control... [the problem was] to see whether in this area of international barriers might not be broken down and patterns of candor and cooperation established" (8). Laura Fermi, looking back on the years at Los Alamos, points out that several factors were necessary for success in producing the atomic bomb:

> An *enormous* concentration of brain power in one country; the very close collaboration of European and American scientists with their different skills and intellectual traits; their unity in their will to defeat the dictators; the formidable American industrial and financial power... (Badash 101).

C. *Criticisms: the goals and direction of a scientific community*

How does this historical model of the Manhattan Project at Los Alamos fit Kitcher's theory, and what does it show? Los Alamos is an ideal model for studying group rationality. It is an enclosed community of scientists working towards a common goal, the development of the atomic bomb, with a distribution of labor among the five divisions: Theory, Experimental Physics, Chemistry, Ordnance and Engineering, and Administration. There was a range from newly-finished college students to world-famous physicists from different nationalities, theoretical experience, and background. As one of the scientists later said in a lecture about his experiences from working on the Manhattan Project, "Los Alamos was three deep in experts and tensions were so high that if the atom bomb had not functioned [when tested at Alamogordo], there would have been a much greater explosion between the scientists" (Badash 73)!

But not only did the community share a common goal, the community achieved it. Indeed, even the scientists themselves have admitted how ideal the situation was for progress: the government provided anything they needed for their work. Communication was also facilitated, since their colleagues were literally down the hall, rather than across several states or across an ocean. So what was key to the Los Alamos success?

Although motivated by various factors, all of the scientists at Los Alamos wanted their "team" to discover the atomic bomb before the Germans: an elevated sense of *esprit de corps*. In fact, it was this political pressure that brought them together in the first place. In addition, the Manhattan Project at Los Alamos is now

recognized as the fountainhead of the atomic bomb. There was an effort to achieve ranking not only as a crucial scientific group in the field of atomic physics, but also of the individuals within the group as well.

The scientists were motivated by strong personal and political opinions (to stop the Germans) as well as a drive for fame. There was a sense of responsibility to their governments - they were no longer working for simply the development of a theory or the advancement of science alone. Founding an isolated community at Los Alamos was a way of maximizing their chances of inventing the atomic bomb, and the open society they created within its boundaries allowed for the free flow of information which sped up the process of discovery.

So we obviously have a community of scientists motivated by various factors other than purely epistemic reasons of truth or knowledge. Kitcher would agree that such a situation augmented the advancement of science, for the scientists had to maximize their chances for personal fame and the personal and political desire for what they believed could end the war or put a stop to Hitler. But should this be the way scientists and their community *ought* to behave?

If the aim of science is to develop science as quickly as possible, such as in the wartime effort of World War II, then the situation at Los Alamos is ideal. Kitcher proposes looking at the social structure of science, for that is where change is needed to advance science further and faster. But the social structure has already changed, and especially the role science plays in the society

as a whole. Science provides the world with knowledge and technology, two things that define society today, and certainly the aim of science is to achieve progress in these fields as best as possible. Science also supplies society with a sense of rationality. Fields such as political science and sociology look to the hard sciences for a scientific method: rational, orderly examination and research of phenomena.

Moreover, the sciences influence the field of ethics and justice. Perhaps just as John Rawls' A Theory of Justice played a key role in influencing Sarkar to write A Theory of Method for the philosophy of science, so too the philosophy of science may offer something back to ethics in the study of group rationality. Societies have always looked to science as a way of defining what is "rational," and have integrated theories of science into everyday life to justify their actions. As a member who had been a part of the community of Los Alamos, J. Robert Oppenheimer once claimed, "*science* rests on, intersects with, alters, affects almost all of man's *ethical life*" (my emphasis, Oppenheimer 104).

Kitcher's model of a sullied community may be a sufficient condition for the advancement of science, but he does not show that *using* such personal incentives and motives is a necessary condition. Such personal biases and motives (good and bad) are present in individuals and cannot necessarily be removed, nor should they be removed. To an extent these drives may help (or hinder) the scientists in their work. But it is not the role of the scientific community to manipulate such psychological mechanisms of individuals to achieve its goals. The scientific communities of today

must include another element: a responsibility to society, or an ethical position.

Drawing a parallel to ethics for a moment, what kind of a social structure does Kitcher suggest? A society in which stratification or division of labor is achieved by appealing to its members' sullied motives. Science strives for the advancement of science: newer and better theories. What does a society strive for? Another paper could be written on that question alone. In some ideological views of social justice, a society aims for fair treatment of all citizens, and to be considered a just society or a just state (the advancement of society). Is it not ironic, if we take such a view, to have a *sullied* community aiming for such *just* ends? The argument has been made, for example, recalling that Adam Smith proposed that the pursuit of self interest by individuals leads to the greatest good, but again, is it the role of the community to manipulate such interests?

The scientists at Los Alamos were driven to discover the atomic bomb in a race against the Germans. Had Hitler discovered the bomb first, history would be considerably different today. But most of the scientists had an idealistic approach to working on the bomb, such as the British scientist who believed that such a weapon would make the world "come to its senses."

In a lecture in 1962 on science and culture, Oppenheimer claimed that

> One cannot doubt that in the sciences the direction of growth is progress.... I do not mean that moral progress is impossible, but it is not, in any sense, automatic. Moral

> regress, as we have seen in our day, is just as possible. Scientific regress is not compatible with the continued practice of science (125).

At another time, he said that scientists "typically rejoice in any success that anyone else has had, even though they may wish that they had had the success themselves. This is a hallmark of science. If people are jealous of each other more than they rejoice... then somehow this has nothing to do with science" (106). Such a community as Kitcher proposes, one where there is such a high level of competition for personal gain, would lead to a community of jealousy, rather than one that celebrates the achievements of the group as a whole. A scientific community that can look at what it has done without a sense of regret has accomplished something.

Given a wartime effort, such a situation as Los Alamos or as Kitcher proposes succeeds in achieving its ends: given the personal motives (varied as they are) the scientists were driven to excel and develop their work in producing a bomb. But should we then maintain that this is how a scientific community must always be fashioned? Would such a community be fruitful in peacetime and the normal efforts of science?

I propose that in a peacetime effort, scientific communities should take personal incentives into account, for the psychological element is powerful. But instead of trying to create a community that manipulates such drives, perhaps we should adapt the sullied community. For more than money or profit (the "prize"), scientists are driven for recognition of their work. And already

successful scientists are recognized in their fields for their achievements.

An ideal scientific community, in my view, is one where the scientific community is not segregated from the rest of society in an ivory tower. Rather, I propose that it is impossible to develop and open society of scientists that can fully prosper without placing it within a social community that takes part in the scientists efforts. Granted, science has developed to a skill beyond layman's understanding. But it is still possible to place scientists in a community where they are socially recognized for their efforts, in a way similar to famous politicians, for a scientist's efforts are no less admirable or important than the achievements made in law.

Finally, I suggest that a situation similar to Los Alamos can be recreated in a larger scientific community. Research institutes and universities provide for close interaction between scientists, and I agree that such intense situations develop science and advance its efforts rapidly. Perhaps in an utopia, the scientific community could consists of groups of scientists or intellectuals who live in close, intense communities, with an open flow of discussion between science and society that promotes understanding and recognition of the scientific accomplishments.

Works Cited

Badash, Lawrence, Joseph O. Hirschfelder, and Herbert P. Broida, eds. Reminiscences of Los Alamos: 1943-1945. Boston: Reidel, 1980.

Kitcher, Philip. "The Division of Cognitive Labor." The Journal of Philosophy LXXXVII (Jan.-Dec. 1990): 1-23.

Oppenheimer, J. Robert. Uncommon Sense. N. Metropolis, Gian-Carlo Rota, and David Sharp, eds. Boston: Birkhauser, 1984.

Sarkar, Husain. A Theory of Method. Berkeley: U of California P, 1983.

Smith, Alice Kimball. A Peril and a Hope: The Scientists' Movement in America: 1945-47. Chicago: U of Chicago P, 1965.

Szasz, Ferenc Morton. British Scientists and the Manhattan Project: The Los Alamos Years. New York: St. Martin's, 1992.