THE GROWTH OF KNOWLEDGE:
Testing a Theory of Scientific Revolutions With a Formal Model

John D. Sterman
System Dynamics Group
Alfred P. Sloan School of Management
Massachusetts Institute of Technology

There has been a dramatic upheaval in the philosophy of science over the past two decades. The traditional notion that science is a logical, rational enterprise continually adding to the stockpile of knowledge has been challenged. A new school has arisen, emphasizing the discontinuous disruptions that punctuate the evolution of science. During such disruptions, or scientific revolutions, a tried and true theory is abandoned for an untested and often heretical alternative. The new theory destroys the old rather than building upon it, and though the successor may flourish for centuries, eventually another crisis develops and another revolution occurs. Yet the new view of science does not mean it is random. Indeed, the proponents of the new view, chiefly Thomas Kuhn in his Structure of Scientific Revolutions, argue that there is an internal structure to scientific activity guiding and shaping it.

This paper presents and tests an explicit theory of scientific revolutions. The theory is based on Kuhn's work and consists of hypotheses about the ordinary year-in, year-out conduct of science. But the theory goes beyond Kuhn and other analyses in one important respect: the theory is embodied in a system dynamics model.

Though Kuhn's work is not explicitly dynamic, he identifies a clear reference mode: the life cycle of a single paradigm. the life cycle consists of four phases: emergence, normal science, crisis, and revolution. During the

emergence phase, a new paradigm has only a few members, and consists largely of untested hypotheses and unorganized data. If successful, the paradigm attracts practitioners away from other schools as a wider array of the "puzzles" presented to the theory are solved. The confidence of the practitioners in the theory rises, and with rising confidence comes more efficient research, or normal science. During the period of normal science, confidence is high, the theory is successful, and large numbers of puzzles are solved. Membership continues to grow until the paradigm dominates the field.

Normal science may last for years or centuries, but gradually, as the theory is extended and empirical work is extended, anomalies, or unresolved challenges to the theory, begin to appear. Often anomalies can eventually be resolved into the theory, but as anomalies accumulate, practitioners slowly begin to lose confidence in the paradigm. The paradigm then enters the crisis phase, when anomalies are accumulating rapidly, and confidence is eroding. Under normal science, in a clash between theory and reality, the burden of proof is on reality. As anomalies develop, the burden of proof shifts back to the theory. Crisis persists until a theory emerges that can explain enough of the anomalies to attract members away from the dying pardigm. The transition is the scientific revolution.

The dynamic hypothesis to explain the life cycle of paradigms is synthesized from Kuhn's theory and the theory of metaphor. A scientific theory or paradigm is an extended metaphor. Like a new metaphor, a new theory illuminates previously obscured or unnoticed aspects of reality. It suggests new explanations, interpretations, and experiments. The elaboration and articulation of a paradigm corresponds to the exploration of the metaphor, just as a playwright or poet draws out the crosscurrents from an image. And just as a metaphor in language can cease to illuminate and grow stale, so the extended metaphors underlying scientific theories become more difficult to explore. The process of "metaphor depletion" leads to the emergence of anomalies. As the anomalies accumulate, practitioners' confidence in the paradigm erodes, leading them to search for new metaphors and ultimately to abandon the paradigm.

Though the model is quite simple (it has but six levels), it represents the recruitment and defection of members, the different research activities available to paradigm members, the determinants of practitioners' ontological and epistemological interpretations of their paradigm, and the ways in which these interpretations influence the conduct of their research. For example, rising confidence in and familiarity with a paradigm causes practitioners to see reality increasingly through paradigm-conditioned eyes, reducing their abilityto recognize anomalies.

The behavior of the model under the assumption of a strong paradigm shows a clear life cycle behavior (see figure). Initially, confidence is low, the paradigm has solved only a few puzzles, and there are but a few members. However, a rapid burst of progress and low anomalies rapidly raise confidence and rising confidence attracts new members. The rapid growth from emergence into normal science is due to a network of positive feedback loops: rising confidence reduces recognition of anomalies, further boosting confidence; high confidence attracts new members, boosting progress and confidence still further; rising confidence increases the efficiency of research, increasing progress and leading to even higher confidence. Through these positive loops the paradigm bootstraps itself from emergence into normal science.

During the period of normal science, solved puzzles increase rapidly. But anomalies gradually begin to accumulate as the underlying metaphor is applied farther from its initial realm of application. Eventually, anomalies

begin to depress confidence, and at that point, the positive loops that cause the paradigm's rapid growth reverse and cause an accelerating crisis and collapse: rising anomalies divert research from puzzle solving, reducing progress and confidence even more; declining membership as confidence falls further reduces progress; and most importantly, falling confidence leads to recognition of even more anomalies as the practitioners increasingly come to question their paradigm.

Tests of the model reveal other interesting results. First, the hypothesis for paradigm change advanced by Kuhn is dynamically consistent with the behavior he seeks to explain. Thus Kuhn's theory passes a test to which few other theories have even been subjected, providing support for his conception of science. Second, the results strongly suggest the observed life cycle of scientific revolutions arises within the structure of scientific research itself, from the ordinary day-to-day activities of scientists. It is not necessary to invoke the appearance of "great men" or the capriciousness of nature to explain scientific revolutions. Third, the behavior of the system over the life cycle is quite insensitive, over a broad range, to many of the initial conditions and internal policies for managing research activities and recruitment. Such insensitivity would be expected given the broad range of historical circumstance and diversity of content surrounding the various sciences that all experience the life cycle of scientific revolutions.

Finally, the model demonstrates the ability of the system dynamics method to test the behavioral plausibility of a structural theory, even when that theory is presented in an entirely qualitative manner, at a high level of abstraction, and in a context entirely divorced from explicit dynamic analysis.

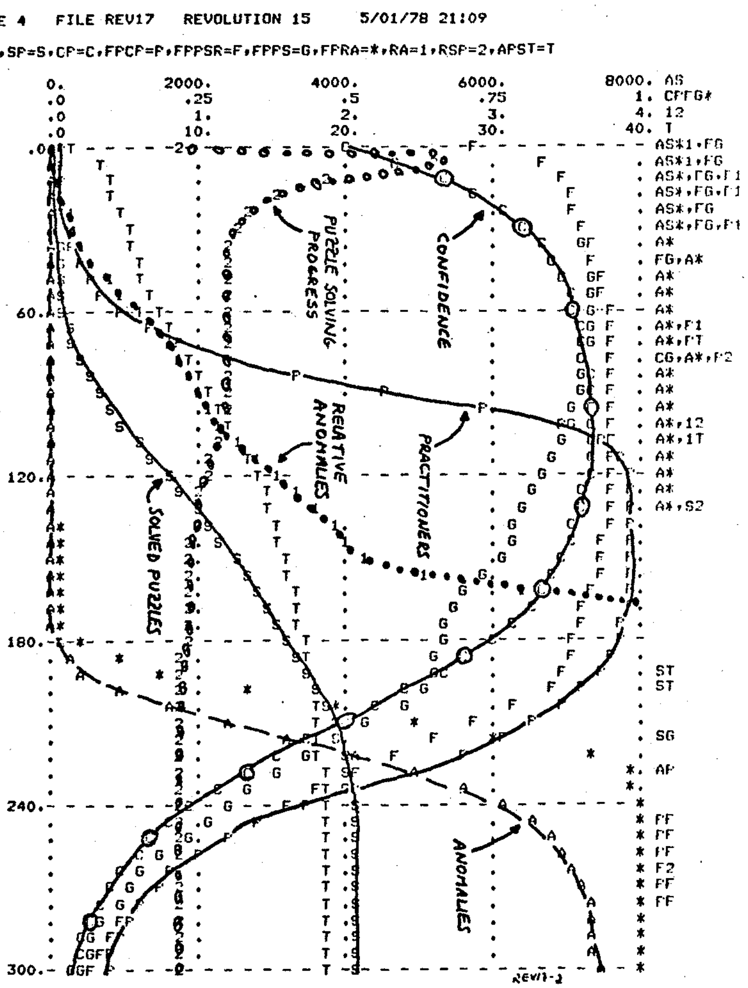

Figure IV-1. The Base Run.

THE GROWTH OF KNOWLEDGE:
Testing a Theory of Scientific Revolutions with a Formal Model

John D. Sterman
System Dynamics Group
Alfred P. Sloan School of Management
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

It is the fate of all knowledge to begin as heresy and end as superstition.

Huxley

1. Introduction

There has been a dramatic upheaval in our conception of science in recent years. The old notion that science is a logical, rational enterprise continually adding to the stockpile of knowledge has been challenged; many now recognize that the evolution of science is punctuated by violent disruptions. During such crises, or scientific revolutions, a tried and true theory is abandoned for an untested and often heretical alternative. The new theory destroys the old rather than building upon it, and though the successor may flourish for centuries eventually another crisis develops and another revolution occurs. Some even claim that science is completely anarchic, more a no-holds-barred brawl than a calm, reasoned investigation of reality. [1]

Yet the new view of science does not mean it is random. Indeed, the proponents of the new view, notably Thomas Kuhn in his Structure of Scientific Revolutions, argue that there is an internal structure to scientific activity guiding and shaping it. The purpose of the present work is to elucidate that structure by constructing an explicit theory of scientific revolutions. The theory draws heavily on the work of Kuhn and consists of hypotheses about the ordinary year-in, year-out conduct of science. But the theory goes beyond Kuhn in two important ways. First, the theory focusses on the dynamic processes involved, that is, it seeks to explain the recurring pattern of growth, crisis, and revolution in terms of the underlying structure of scientific activity. Second, the theory is embodied in a computer simulation model. The model uses the assumptions of the theory to play the roles of the actors and trace out the behavior of the system over time, thus, providing a test of the theory by checking whether the assumptions actually produce the lifecycle of scientific revolution.

2. Defining the Problem

The classic examples of scientific revolution are the Copernican and relativistic/quantum revolutions. It is widely recognized that these events signalled profound shifts in human understanding of nature. At the same time, revolutions were thought to be rare, marking the transition from superstition to science, from myth to method. But the history of science does not sit well with this view. Kuhn points out the dilemma facing historians of science:

> The more carefully they study, say, Aristotelian dynamics, phlogistic chemistry, or caloric thermo-dynamics, the more certain they feel that those once current views of nature were, as a whole, neither less scientific nor more the product of human idiosyncracy than those current today. If these out-of-date beliefs are to be called myths, then myths can be produced by the same sorts of methods and held for the same sorts of reasons that now lead to scientific knowledge. If, on the other hand, they are to be called science, then science has included bodies of belief quite incompatible with the ones we hold today. [2]

Kuhn resolves the dilemma by recognizing the scientific revolution as a basic feature of science. He argues that new theories tend to replace old ones rather than building upon them. New theories are usually incompatible with the old, built on different metaphysical foundations, relying on different observations, research methods, and criteria for validity. These and other characteristics of a scientific school define what Kuhn calls a paradigm. The concept of paradigms is central to Kuhn's analysis. It is rich in meaning and nuance, and several key senses need to be distinguished. [3]

> Paradigms are specific works that define a field. They are accepted examples of actual scientific practice--examples which include law, theory, application, and instrumentation together--[they] provide models from which spring particular coherent traditions of scientific research. [4]

Thus Newton's Principia and Keynes's General Theory are paradigms.

Paradigms define the nature of a particular science. They provide the tools, methods, and examples that guide practitioners in their research:

> Close historical investigation of a given specialty at a given time discloses a set of recurrent and quasi-standard illustrations of various theories in their conceptual, observational, and instrumental applications. These are the community's paradigms, revealed in its textbooks, lectures, and laboratory exercises. By studying them and by practicing with them, the members of the corresponding community learn their trade. [5]

The history of science is the history of the rise and fall of such communities. The great diversity of these cultures, from phlogistic chemistry to neoclassical economics suggests the forces behind the growth and decline are independent of the particular beliefs, people, and methods that make them up. Identifying a set of generic forces, the underlying structure common to all paradigms, is the purpose of this effort.

But what exactly is this common pattern of behavior? One element has already been identified: the growth and decline in membership. A new paradigm emerges with the work of a single, or at most several, persons. If successful, the paradigm attracts practitioners away from other schools. Growth ceases when nearly all the practitioners in a given field embrace the paradigm, when "physicist" comes to mean "Newtonian." Dominance may be long- or short-lived. Eventually, however, the paradigm starts to fail (in a special sense discussed below) and loses members to new theories. Often recruitment ceases as young scientists are drawn into a competing school and the paradigm simply dies away.

The character of scientific activity also changes over the life cycle of growth and decline. A new paradigm must fight to survive. It emerges in the context of older paradigms. New paradigms are largely untested--often the proper way to apply a theory is unclear at first even to its creators. The result is conflict over the fundamentals and disagreements over the "facts", slowing progress. When some paradigm gains the upper hand, however, the character of research changes rapidly. Kuhn recounts the history of electrical research to illustrate the process. [6] At first, there was no guiding paradigm and confusion reigned. No one could agree on which facts were important or even what the facts were. After years of conflict, Franklin proposed a theory that "could account...for very nearly all [the known electrical] effects and that therefore could and did provide a subsequent generation of "electricians" with a common paradigm for its research." [7] The result was dramatic:

> Freed from concern with any and all electrical phenomena, the united group of electricians could pursue selected phenomena in far more detail, designing much special equipment for the task and employing it more stubbornly and systematically than electricians had ever done before. Both fact collection and theory articulation became highly directed activities. [8]

Thus, conflict and confusion gave way to purposeful, efficient activity. Kuhn calls such activity normal science.

Normal science is puzzle solving. It is the extension and articulation of a paradigm. The dynamic feature of interest is the cumulative nature of puzzle solving within a given paradigm. Scientists try to build upon the foundation the paradigm work laid out and force nature into line with a paradigm, not elicit new and unusual phenomena from her. But new and unusual phenomena do arise. As normal science progresses, results are obtained that do not fit into the range of expectations determined by the paradigm. Kuhn terms such novelties anomalies. Anomalies are not simply disagreements between "fact" and "theory", for these occur constantly. Indeed, one of the tasks of normal science is to bring fact and theory into ever-closer agreement (and often this is accomplished more by adjusting the facts than by refining the theory). [9]

The null result of the Michelson-Morley experiment is a particularly famous example of anomaly. Other examples include the photoelectric effect and ultra-violet catastrophe of Newtonian physics; modern physics is plagued with "renormalization" and the seemingly never-ending growth in the ranks of the "elementary particles." Again, the dynamic feature of interest is that the progress of normal science, an activity whose aim is to suppress novelty, produces anomalies that begin to accumulate. When the level of anomalies reaches a certain point, the character of research changes again, entering a phase Kuhn calls crisis.

Crisis develops when scientists begin to lose faith in the paradigm. Under normal science, the paradigm has priority--in a clash between reality and expectation, the burden of proof is on reality. But when anomalies accumulate, that burden slowly shifts onto the paradigm. Having come to question their tools and methods, and unwilling to trust their paradigm-conditioned intuition, scientists find themselves adrift in an ocean of confusion. Research is diluted as practitioners increasingly abandon puzzle solving to take up the anomalies. Some even "desert science because of their inability to tolerate crisis." [10] New theories and ad-hoc patches begin to arise. The situation is much like that in electrical research before the emergence of Franklin's theory. Crisis persists until a theory emerges that can attract enough followers and explain enough of the anomalies to become the next paradigm. The transition is the scientific revolution.

Four stages in the life cycle of a paradigm have been described: emergence, normal science, crisis, and revolution. The basic pattern is fundamental, though there are individual differences in timing and severity. The entire process may take a few years or a few centuries; a new paradigm may appear rapidly or crisis may deepen for decades. Figure 1 is a generic representation of the basic pattern. It is the reference mode, the behavior that any theory of scientific revolutions must reproduce. To gain credibility, a theory must self-generate the reference mode without relying on

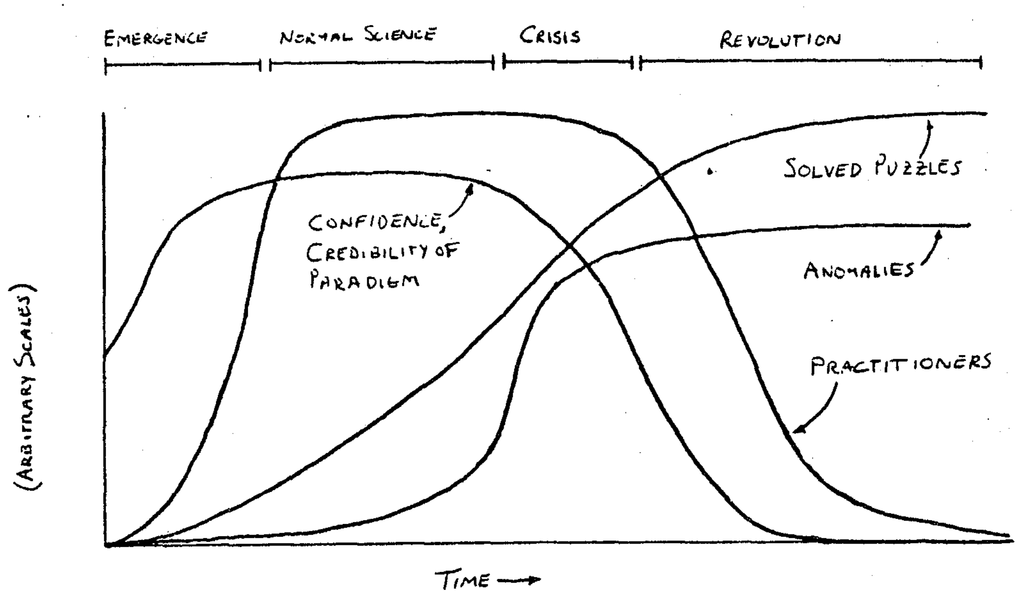

Figure 1

external driving forces such as the emergence, as if by magic, of a new and better theory. Further, the structure of the model must be a plausible representation of the way scientists actually work. All the variables should have real world counterparts. The triple requirements of reproducing the reference mode internally with a plausible structure are strong constraints. Satisfying them is the sine qua non of model validity. But they are by no means sufficient to prove the theory. Indeed, proof is the wrong dimension for evaluation. Rather, the goal is to illuminate the dynamics of scientific revolutions by making explicit the connections between the ordinary business of scientific research and the dramatic changes in our conception of the world that come out of it.

3. A Model of Scientific Revolutions

3.1 The Paradigm as Metaphor

The heart of the theory is the identification of the metaphysical and epistemological facets of paradigms with metaphors. In essence, the dynamic hypothesis is that paradigms are metaphors, and metaphors are limited representations of reality that crack when strained, producing anomaly and crisis. Four properties of metaphor in particular bear elaboration.

1. Metaphor is everywhere: I.A. Richards notes, "we cannot get through three sentences of ordinary fluid discourse without it." [11] Nelson Goodman echoes Richards by saying, "metaphor permeates all discourse, ordinary and special, and we should have a hard time finding a purely literal paragraph anywhere." [12] C.M. Turbayne goes farther by emphasizing that metaphor permeates all of our thought as well as our language. [13] Similarly, Kuhn stresses the priority of paradigms, speaks of analogies as the foundations of

paradigms, and suspects that "something like a paradigm is prerequisite to perception itself." [14]

2. Metaphor involves a "transfer of schema" from one area of experience to another: [15] "Metaphor" means "to carry across." Consider the metaphors "Richard is a lion", "the brain is a computer", or "capitalist economies are markets." The characteristics of lions, computers, and markets are transferred, via the metaphor, to Richard, the brain, and capitalist economies. The metaphors work because the characteristics of lions, computers, and markets are well known and carry a constellation of meanings, examples, connotations, and nuances with them that illuminate the subjects to which they are applied. Max Black calls this constellation a "system of associated commonplaces." [16] Both ancient and modern scientific theories are grounded in metaphor drawn from common experience: consider Heraclitus's "all is fire", the wave model of light, the corpuscular/atomic view, and the plum-pudding model of the atom.

3. Metaphors mask or filter reality: The image of the mask or filter appears constantly in discussions of metaphor, and is itself a crucial metaphor. Because a metaphor draws upon a system of associated commonplaces, certain relationships are highlighted while others are suppressed. Black's image is a piece of smoked glass. Looking at the night sky through such a glass blocks out some stars, thereby accentuating others. Stars that were not noticed before can also be seen. The filtering power of paradigms is central to Kuhn's theory as well: "In the absence of a paradigm...all of the facts that could possibly pertain to the development of a given science are likely to seem equally relevant." [17] It is interesting that the images used to describe the operation of metaphor are so similar. Black uses smoked glass,

Turbayne uses the emerald goggles of Oz, and Kuhn mentions glasses that turn the image of the world upside down:

> ...the scientist who embraces a new paradigm is like the man wearing inverting lenses. Confronting the same constellation of objects as before and knowing that he does so, he nevertheless finds them transformed through and through in many of their details. [18]

The filtering aspect of metaphor is often expressed by saying metaphors are models of reality. [19]

4. Metaphors define reality: In addition to organizing perception through the transfer of schema, metaphor creates the world, or at least a part of the world. Black notes:

> It would be more illuminating in some of these cases to say that the metaphor creates the similarity than to say that it formulates some similarity antecedently existing. [20]

Turbayne argues for the power of metaphor to shape the world with the history of Cartesian mechanism:

> ...enthralled by his own metaphor, [Descartes] mistook the mask for the face, and consequently bequeathed to posterity more than a world view. He bequeathed a world....Had he [chosen a different metaphor]...we should now be living in a different world. [21]

Kuhn attributes the same power to paradigms:

> ...the historian of science may be tempted to exclaim that when paradigms change, the world itself changes with them. Led by a new paradigm, scientists adopt new instruments and look in new places... [They] see new and different things when looking with familiar instruments in places they have looked before. Insofar as their only recourse to the world is through what they see and do, we may want to say that after a revolution scientists are responding to a different world. [22]

Parallels between metaphors and paradigms could be multiplied indefinitely but the point is clear. The term "metaphor" or "extended metaphor" can be substituted for the term "paradigm" without doing violence to the sense of either one.

Science, then, can be viewed as the elaboration and exploration of metaphors. More precisely, normal science is the elaboration of metaphor. The formulation of a theory corresponds to the initial transfer, and the exploration of the metaphor to what Kuhn calls the articulation of the paradigm.

Though a metaphor is usually inspired by a small number of similarities, other connections are soon noticed. When first formulated, not all the connections are apparent. The metaphor is unexplored. Indeed, the attraction and power of metaphor lies in its ability to suggest undreamt-of possibilities that open the door to elegant or useful visions of reality. The task of normal science is to search out these possibilities and build upon them just as a poet constructs an image and carefully draws out the crosscurrents.

When first formulated, the mask of a metaphor is particularly striking. Fresh metaphors jangle and startle; they are impudent and lively, setting the mind off in new directions and creating new insights into familiar problems. With usage comes familiarity, and familiarity breeds contempt; soon the edge wears off the metaphor and its masking nature is forgotten. Metaphors _qua_ metaphors are mortal, for "with repetition, a transferred application of a schema becomes routine, and no longer requires or makes any allusion to its base application. What was novel becomes commonplace, its past is forgotten, and metaphor fades to mere truth." [23] Similarly, newly proposed theories are often introduced explicitly as models or metaphors. Kuhn cites Copernican astronomy which was introduced as a convenient fiction, that is, it was convenient to treat the solar system _as if_ the earth travelled around the sun. Only later, as heliocentrism became accepted, did this fiction become scientific truth. Even today, more than seventy years after the introduction of relativity, the vast majority of people unequivocally believe in the truth of the Copernican system.

While metaphors that become commonplace live on as truth, not all metaphors are so lucky. Metaphors are inherently limited, and if pushed too hard will strain and crack. For example, suppose "man is a wolf" is proposed as a theory of human nature. Should it become the accepted paradigm, the task of normal science would be to extract useful insights from the metaphor. Statements such as "man is fierce and engaged in constant struggle", "men travel in packs", and "men have accepted leaders," might result. Such statements could be illuminating and even contribute to the design of governments, law, and technology. At some point, however, further application of the metaphor would begin to yield statements like "man has fur and big teeth", and "man has eyes, ears, and a heart." Such statements either blatantly clash with experience or are trite. They are anomalies. They arise from the fabric of the theory itself through the normal application of puzzle solving. The accumulation of anomalies undermines the utility and appeal of metaphors and can send them to the grave, disgraced as falsehood. [24]

3.2 The Structure of Puzzle-Solving

Capturing the complex and subtle processes of metaphor birth, exploration, and death in a model involves many simplifications. The basic structure of puzzle solving used in the model is shown in Figure 2. Three categories of puzzles are distinguished. Solved puzzles and anomalies are self-explanatory; the third category, puzzles under attack, consists of those puzzles that are formulated and actively being attacked, but which have not yet yielded or been recognized as anomalies.

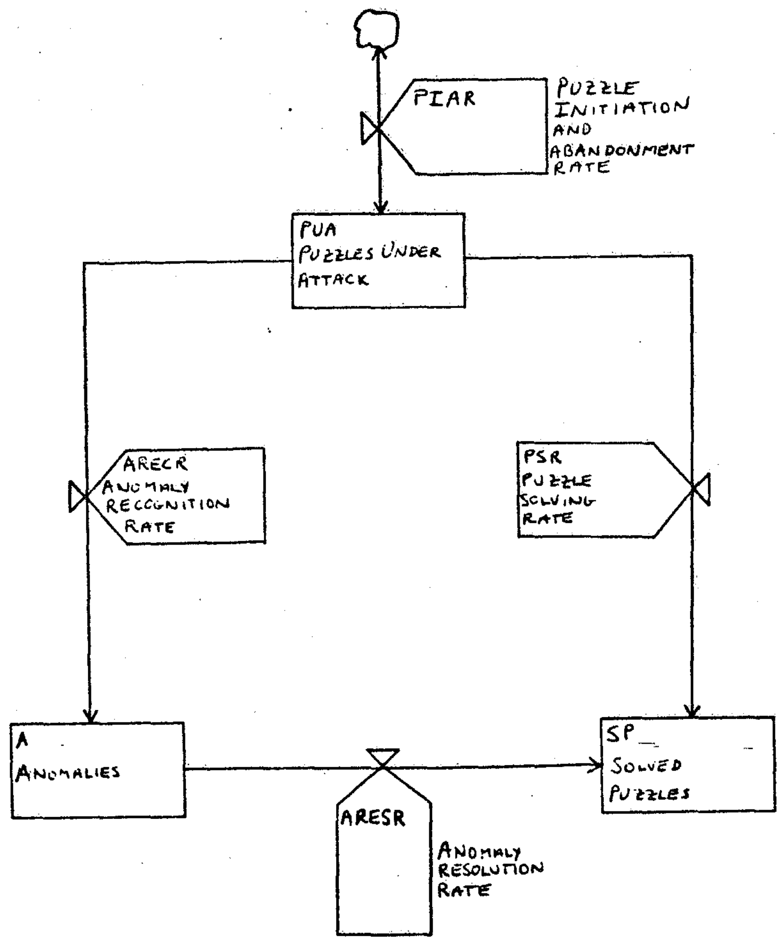

Figure 2

Four flows connect the different categories. The puzzle initiation and abandonment rate is the source of puzzles. Under normal conditions, puzzles are formulated and brought under attack as others are solved. Under conditions of collapse, however, there may be too many puzzles under attack for the number of practitioners remaining, and the abandonment of puzzles will dominate. If all goes well, a puzzle, once formulated and attacked, will be solved in fairly short order. Such puzzles flow into the class of solved puzzles via the puzzle-solving rate. If the puzzle is recalcitrant, however, it can become recognized as an anomaly. Anomalies can sometimes be resolved into solved puzzles. The shifting balance between these flows determines the behavior of the system, and thus the forces affecting them are crucial.

The determinants of the initiation and abandonment rate are straightforward. The number and average productivity of practitioners involved in puzzle solving defines a desired level of puzzles under attack. When the actual number differs from the desired level, research is initiated or abandoned to make up the difference. The number of puzzles solved each year depends on the number under study, the fraction of practitioners involved in sanctioned research and of those the number involved in puzzle solving, and the average difficulty of puzzles (Figure 3). Practitioners within a paradigm can be involved in different sorts of work. The majority will usually be involved in puzzle solving, while some will be working to resolve anomalies, and others, dissatisfied with the paradigm but unable or unwilling to defect to another, try to come up with alternatives, write philosophical essays, etc. Those involved in puzzle solving and anomaly resolution make up the fraction of practitioners in paradigm-sanctioned research. This fraction depends on the degree of confidence practitioners have in the paradigm. When confidence is high, almost everyone is involved in sanctioned research. If confidence drops, however, the number in sanctioned research drops as practitioners lose faith in the paradigm.

Confidence is defined on a scale from zero to one. A confidence level of one corresponds to absolute certainty in the truth of the paradigm. It implies that the practitioners have so much faith in the paradigm no experience could challenge it; no observation or result could convince them

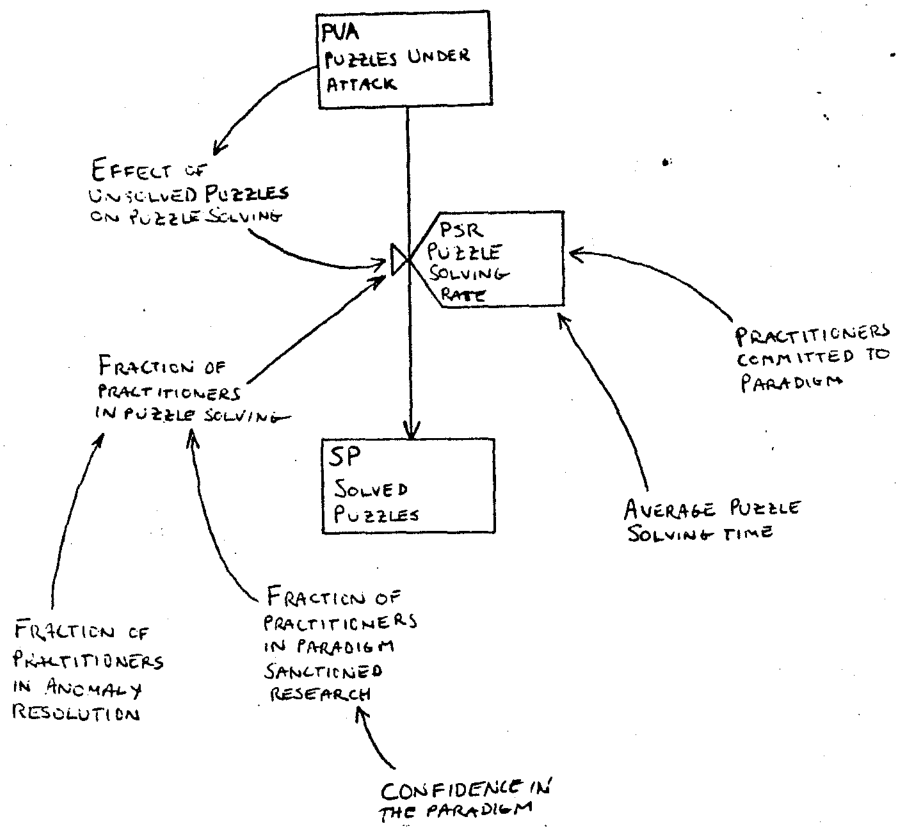

Figure 3

the paradigm was not true. Similarly, a confidence level of zero implies absolute certainty the paradigm is false. No experience could convince them it had anything to offer. The midway point is neutral, where practitioners are neither leaning toward the paradigm or away from it; it is the point of maximum doubt.

The most important determinant of the puzzle-solving rate, however, is the average difficulty of puzzles. It is assumed that, on average, puzzles become more difficult to solve as the number of solved puzzles grows. The increase in difficulty is reflected in an increase in the average time required to solve a puzzle. The purpose of this relationship is to capture the notion that a metaphor gets harder and harder to explore as it is elaborated and developed. It is the core of the dynamic hypothesis. There are several ways the "depletion" of metaphors could have been represented. One way would be to assume that a finite number of puzzles fall in the domain of the paradigm; when these are exhausted, further work produces nothing but anomalies. Masterman seems to advocate such a view:

> ...it is not only the case that a fully-extended paradigm, or theory, reaches a point where further extension of it produces diminishing returns. The situation is worse. The paradigm itself goes bad on you, if it is stretched too far, producing conceptual inconsistency, absurdity, misexpectation, disorder, complexity, and confusion, in exactly the same way as a crude analogy does, if pursued too far, say, in a poem.... The property of crudeness...[means] that a paradigm must be finite in extensibility. [25]

While Masterman's theory is sufficient to cause paradigms to collapse (in the same way that the extraction of all the oil from a well is certain to shut it down), it is not necessary. Indeed, the notion that paradigms are finite is a very strong assumption. Fortunately, it is sufficient to assume the puzzle-solving potential of a paradigm is infinite, but as normal science progresses, puzzles gradually become more difficult to solve. The growth in difficulty is gradual because practitioners attack the easy puzzles first, leaving the difficult ones for advances in technology or theory. In addition, successful paradigms get applied in realms quite far from their original field of application just as Newtonian mechanics, a theory formulated to deal with terrestial and celestial motion, came to be applied to subatomic phenomena. The farther from home the metaphor is applied, the more likely nature is to step outside the boundaries of the metaphor.

The forces determining anomaly resolution are similar to those affecting puzzle solving. The rate at which anomalies are incorporated into the theory depends on the number of practitioners in sanctioned research, the

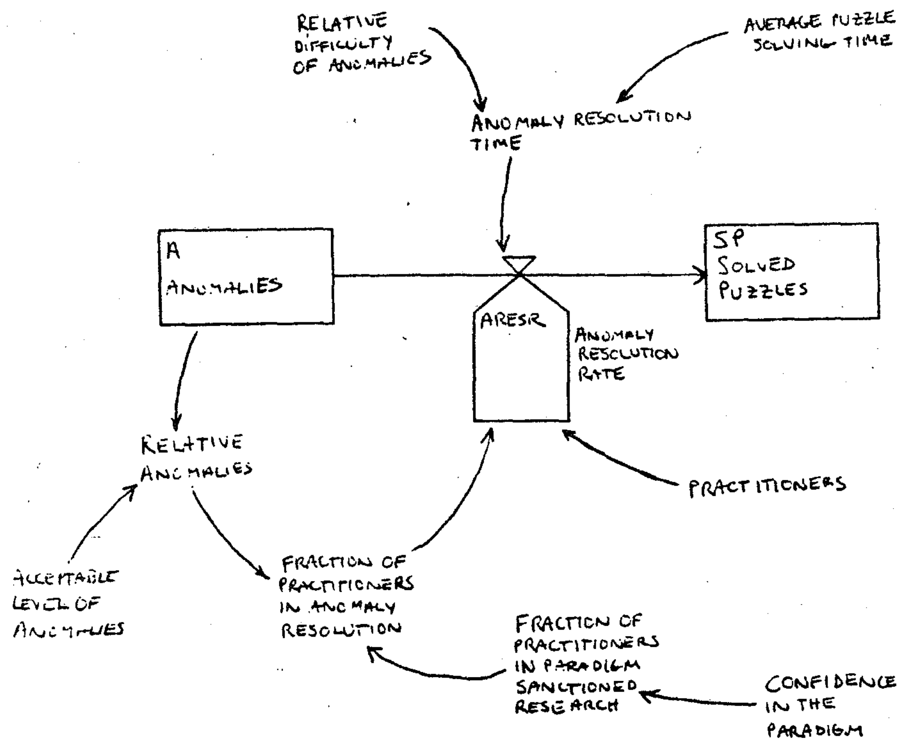

Figure 4

fraction of those involved in anomaly resolution, and the anomaly resolution time, as shown in Figure 4. Anomalies are assumed to be relatively more difficult than puzzles, and as the puzzle-solving time increases, the anomaly resolution time grows proportionately. The fraction of practitioners involved in anomaly resolution depends on the balance between the number of anomalies and the acceptable number. The acceptable level of anomalies is the number that can be tolerated without losing confidence in the paradigm. The acceptable level is not zero, for there are always some problems with any theory, and to lose confidence the first time an anomaly crops up would be to abandon a theory that may have a great deal to offer. If the number of anomalies increases, scientists are drawn into anomaly resolution; if they are successful the number of anomalies stabilizes or declines. Kuhn notes that practitioners are extremely reluctant to work on anomalies. Except for a few scientists who thrive on tension and confusion, the vast majority prefer the relative safety and professional rewards of puzzle solving.

Anomaly recognition is a tricky business. Anomalies are not simply experiments that run counter to expectation. There are always disagreements between fact and theory; it is the task of normal science to reconcile the two. Only when normal science repeatedly fails to resolve the differences or explain some novelty does a puzzle become recognized as an anomaly. Thus, in contrast to the theory that there are "crucial experiments" that provide potential falsifications of a theory, the view advocated here is that there is no fundamental difference between an ordinary unsolved puzzle and an anomaly except the length of time it has resisted solution. [26]

The longer a puzzle resists solution, the more likely it will be recognized as an anomaly. Thus, the fraction of puzzles recognized as anomalies depends on the balance between the average time required to recognize an anomaly and the average time required to solve a puzzle. When the recognition time is high relative to the average puzzle-solving time, very few anomalies will appear since few will remain unsolved for the length of time required to become an anomaly. When the recognition time is low relative to the average solving time, most of the puzzles under study will become anomalies before they are solved.

Anomaly recognition depends critically upon the practitioner's attitudes towards the paradigm. If the practitioners believe the paradigm is "true", they are more reluctant to recognize anomalies than if they are not as confident, in the same way that the freshly-minted metaphor is obviously only a metaphor but the old one is taken uncritically to be literally true. It is not simply a matter of knowing the paradigm is wrong and refusing to admit it. When practitioners are highly confident of a paradigm their perceptions are so conditioned by it that they cannot recognize or assimilate phenomena that violate it. [27]

3.3 The Role of Confidence

Confidence is the focal point of the model. It influences the way practitioners allocate their research effort, how they perceive anomalies, and determines recruitment and defections into and out of the paradigm. Confidence represents a constellation of attitudes and commitments. It reflects basic beliefs about reality by capturing the extent to which practitioners take the metaphor defining a paradigm as literal truth. Confidence responds to the progress of normal science and to the number of anomalies, and thus is a measure of the health and vigor of a paradigm. Figure 5 illustrates the forces acting on confidence. Confidence tends to decline when anomalies exceed their acceptable level and rises when they are below it. Confidence also declines when the rate of progress of normal science falls below a goal defined by the traditional number of solved puzzles plus a margin for the expected growth of solved puzzles.

The impact of progress and anomalies on confidence depends on the degree of confidence itself. If confidence is very high or very low, practitioners will be relatively unwilling to change their degree of confidence. At the extremes, where practitioners are absolutely certain that the paradigm is true (or false), confidence cannot change at all, by

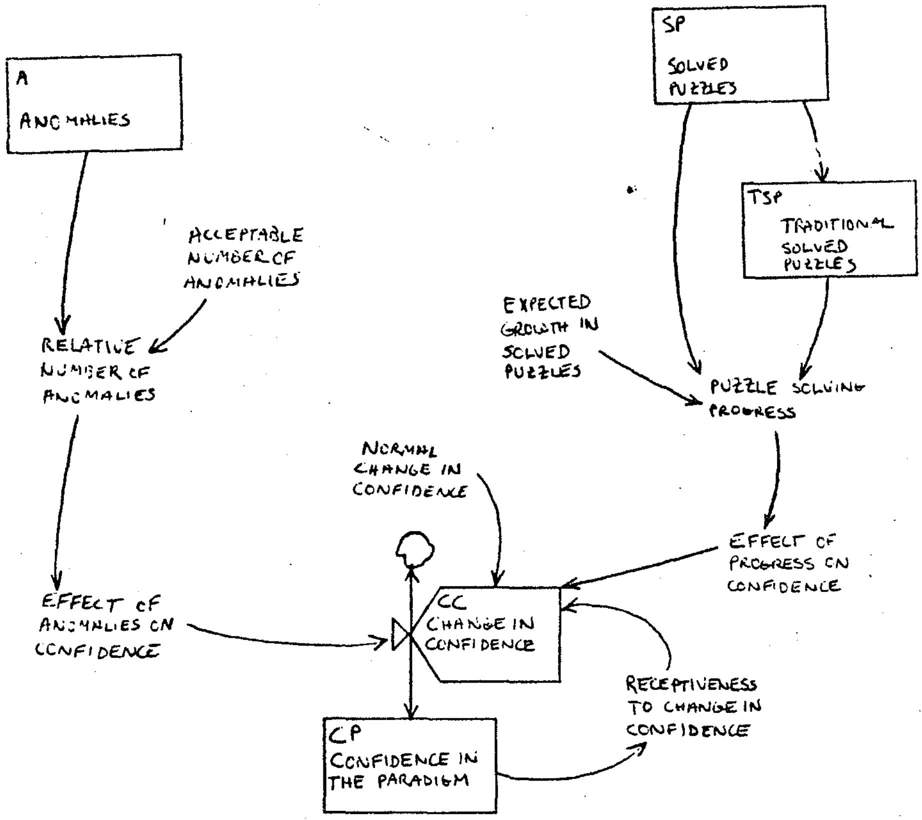

Figure 5

definition. In the midrange, where uncertainty and doubt dominate, confidence can change fairly rapidly since the practitioners have no strong reasons for holding or rejecting the paradigm.

3.4 The Paradigm as Community: Recruitment and Defection

At any given moment in the life of a paradigm there is a group of practitioners committed to it comprising some fraction of the membership in a given field. In reality, commitment is a grey area: there are degrees of training and familiarity, there are pure researchers and pure teachers, philosophers and technicians and all combinations in between. These distinctions are lumped together here. A practitioner is considered committed to the paradigm if the paradigm is the person's primary guide to professional reality. Membership grows through recruitment and shrinks through defections (including deaths); the determinants of these flows are therefore the determinants of the rise and fall of paradigms. In reality, many forces influence recruitment and defection such as demonstrated puzzle-solving ability, the presence of anomalies, the strength of alternative theories, state attitudes toward science, the availability of funding, etc. In the model, practitioners are assumed to respond to the confidence of those in the paradigm relative to the confidence of outsiders in alternative paradigms. Confidence represents the accumulated puzzle-solving ability and threat from anomalies; it is used to proxy funding and attitudes. Figure 6 shows these forces. When the paradigm is equally attractive as its competitors, that is, when the confidence levels are equal, recruitment and defections are in balance. If confidence falters, membership declines as defections exceed recruitment.

The confidence of practitioners belonging to alternate paradigms is assumed to be constant, corresponding to the assumption that there is always a competing paradigm available and that it has an unchanging degree of confidence. While clearly not true, this assumption is justified for several reasons. Competing theories do not arise at random. They tend to be born in the crisis phase of an existing paradigm, and are scarce during the period of normal science. They are part and parcel of the dynamic process. Thus, in the emergence phase of a new paradigm, recruitment would be easier than assumed here, since the old paradigm would be in crisis and confidence would be low. During normal science, alternatives would die away, and thus defections in the early phases of the crisis would be retarded, trapping the

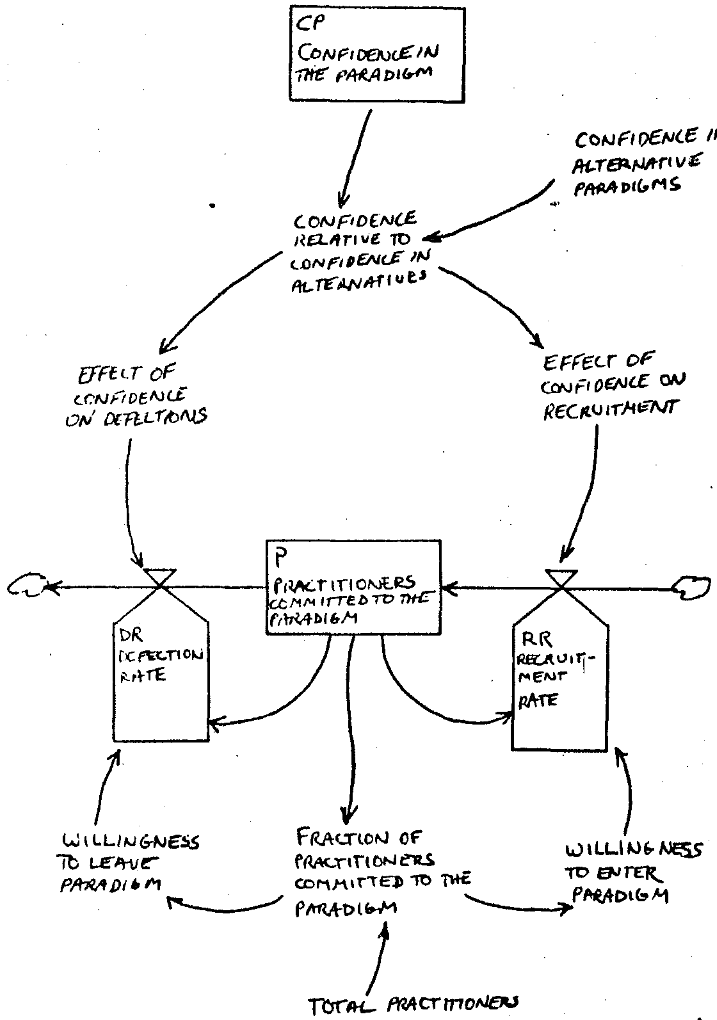

Figure 6

disgruntled practitioners in the dying paradigm. The ease of recruitment and the willingness of practitioners to defect depends on the fraction of the total number of practitioners committed to the paradigm. Recruitment gets more difficult as 100 per cent membership is approached since the most willing practitioners will be recruited first, leaving those who are either strongly committed to alternative paradigms or simply unable to make the shift. A symmetrical relationship affects defections, reducing the defection rate as the number of practitioners declines.

3.5 The Role of Data

There are no numerical data available on paradigms, no standard time series for puzzles, anomalies, or confidence. There is, however, a rich store of qualitative data, impressions, common sense, and historical episode and anecdote. These, along with theories of metaphor and science, form the basis for the model. But computers demand numbers, and the model does contain precisely quantified relationships. Obviously the numbers chosen are highly conjectural. Fortunately, the precise numbers do not seem to matter to the overall behavior of the model. Changing parameters does, as will be seen in the next section, change the timing of certain events and the particular values certain variables take on, but the fundamental behavior of the system is invariant to such changes. The role of data is to provide a vehicle for analysis by allowing the computer to test the consistency of the theory.

4. Testing the Theory

Figure 7 shows the reference run of the model. The simulation spans three hundred years, a rather long lifetime for a paradigm. The overall length of the lifecycle depends on the parameters of the model, particularly the inherent explanatory power of the paradigm. The reference run corresponds to a strong paradigm such as Newtonian physics.

The initial conditions represent a newly-emerged theory. There are only five adherents of the paradigm, they have solved but one puzzle, and there are no anomalies since the theory is so young and the metaphor virtually unexplored. The practitioners' initial degree of confidence is one-half, meaning they are in doubt about the validity of the paradigm, neither leaning toward it or away from it. Because of the low initial confidence, only three-quarters of the practitioners' time is spent in paradigm-sanctioned research. The paradigm is so new and shaky they still question the fundamentals just as the early quantum physicists spent years trying to understand the implications of the uncertainty principle.

In the first three decades, there is a dramatic increase in confidence. Confidence rises because anomalies are low relative to the acceptable level and because there is a large initial surge of progress in

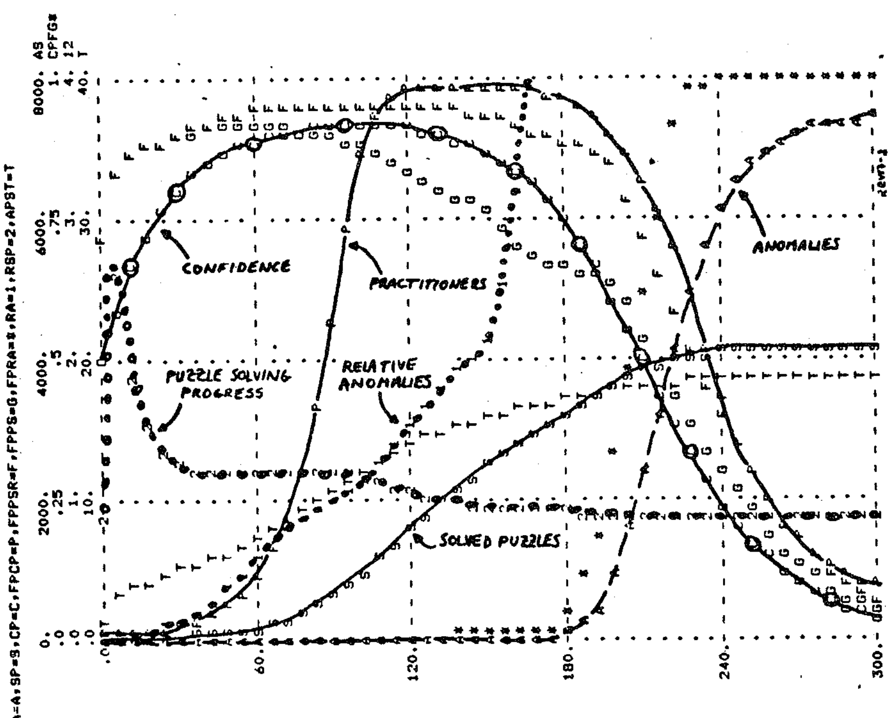

Figure 7

puzzle solving. Confidence is not simply responding to these pressures, however. Progress is high and anomalies are low <u>because</u> confidence is rising. Rising confidence creates the pressures that cause it to grow. Several feedback mechanisms are responsible for the self-reinforcing nature of confidence.

Rising confidence boosts progress. Initially, only three-quarters of the practitioners are involved in sanctioned research because confidence is so low. As confidence begins to rise due to upward pressure from the low number of anomalies, the doubts and confusion of the practitioners wane, and more are drawn into sanctioned research. Progress is increased, and with it upward pressure on confidence. An increase in confidence is a natural reaction to a new idea that can solve some problems, shows promise of solving many more, and has no serious strikes against it yet, even though it is largely unfamiliar. The effect of this positive feedback process, illustrated in Figure 8, is

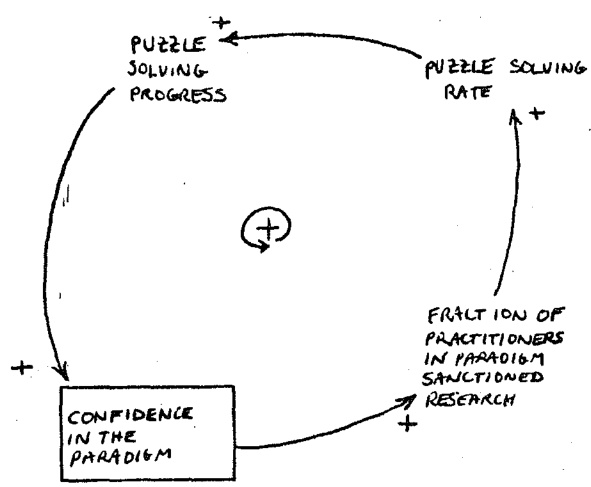

Figure 8

to create a burst of progress in the paradigm's early years, a flush of early success that spurs interest in the theory. The process saturates when the vast majority of practitioners are involved in sanctioned research.

Confidence continues to rise rapidly even after progress slows. Progress is still greater than expected, exerting some upward pressure, but most of the impetus causing confidence to grow comes from the low relative number of anomalies. Anomalies are held down by the growth in confidence itself, as shown in Figure 9. Rising confidence increases the anomaly recognition time as practitioners, increasingly familiar with and certain of the paradigm, start to see reality through paradigm-conditioned eyes. The result is to suppress the appearance of anomalies even though the average puzzle-solving time is increasing. In turn, lower anomalies boost confidence even more.

The two positive feedback loops rapidly raise confidence in the first years of the paradigm's life. In a short span, the practitioners have evolved from confusion and doubt to a high degree of faith in the truth of the theory. Their confidence allows them to focus their activities on puzzle solving and anomaly resolution. The paradigm has bootstrapped itself into normal science.

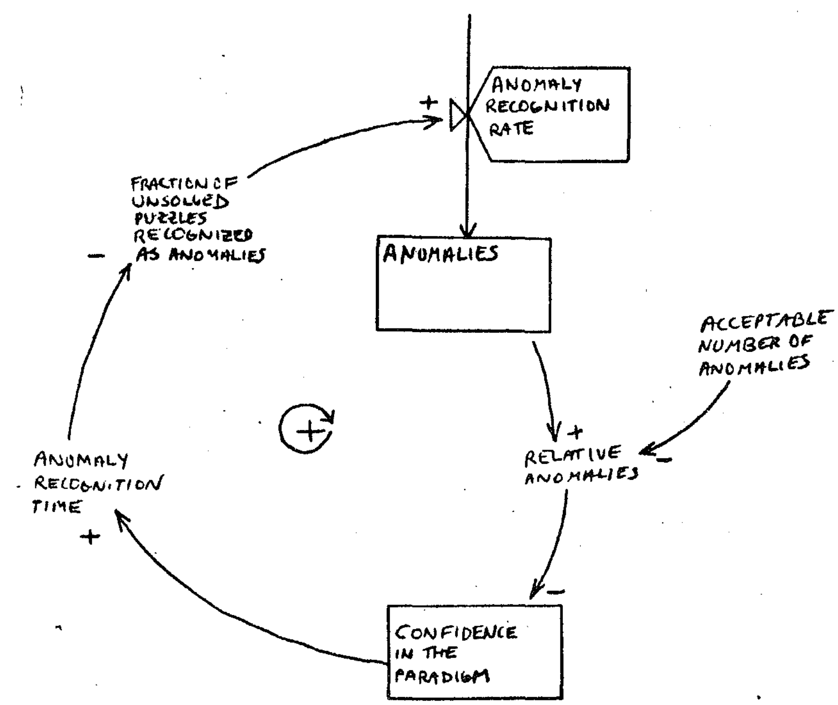

Figure 9

At the end of the emergence phase practitioners are growing, confidence is still rising, and solved puzzles are growing slightly. In the next eight decades, the paradigm grows from about ten percent of the total community to virtually one hundred percent. Again, there are self-reinforcing processes at work. Recruitment and defections are proportional to the number of practitioners committed to the paradigm. Thus, as the number of practitioners grows, the number of students they teach, articles they publish, conferences they attend, and societies they belong to increases, further increasing the number of practitioners. A bandwagon effect develops. At the same time, as the number of practitioners grows, the puzzle-solving rate increases, boosting progress. Higher progress raises confidence, and higher confidence increases the recruitment rate even more. The effect of this loop is to sustain progress at a level in excess of expectations during the growth phase of the paradigm.

Confidence starts to level off during the period of normal science for several reasons. Because doubt and confusion are not particularly pleasant states, practitioners responded quickly to the early successes of the paradigm. Further increases in confidence, however, corresponding to a shift from strong commitment to religious zeal, are slower in coming. More importantly, the pressures causing confidence to rise in the first place have diminished, particularly because the relative number of anomalies has risen. Anomalies rise despite a gradual shift of practitioners away from puzzle solving into anomaly resolution. The rise is due to the gradual increase in puzzle-solving time that has come with the growth of solved puzzles. Higher average puzzle-solving time implies that more puzzles are going to be recognized as anomalies before they are incorporated into the theory. In about the one hundredth year, the effect of anomalies on confidence becomes negative and then outweighs the positive effect of progress for the first time. Confidence then reaches its peak. Over the next half century there is a gradual erosion of confidence as the number of anomalies continues to rise.

By the 150th year, the paradigm has entered the crisis phase. During the crisis, the same self-reinforcing mechanisms that caused the paradigm to grow so rapidly in confidence and in membership now work in the opposite direction. As anomalies rise, the fraction of practitioners involved with puzzle solving declines, reducing the rate of progress and eroding confidence. As confidence declines, the fraction of practitioners in sanctioned research begins to falter, further reducing confidence. The process becomes a vicious circle in which lower confidence reduces progress, accelerating the decline in confidence and causing still more people to abandon puzzle solving. The increasing doubt of the practitioners soon lowers the anomaly recognition time. The practitioners begin to see that their theory is a limited representation of reality, just as the early practitioners did, and with this recognition comes increasing awareness of the holes and rough edges, the places where nature does not go along with the paradigm. Even more anomalies appear and confidence is challenged further. In addition, as confidence declines, practitioners lose the dogmatic rigidity they acquired in the period of normal science. As doubts mount, confidence becomes more volatile and responds faster and faster to rising anomalies and inadequate progress.

After one hundred and eighty years, crisis has deepened and accelerated. Anomalies are more than four times their acceptable level and rising at an increasing rate. Though confidence has fallen substantially, only a few practitioners have abandoned sanctioned research and still fewer have left the paradigm altogether. It is interesting that there is a lag of about sixty years (out of a three hundred year life cycle, or about one-fifth the lifetime of the paradigm) between the beginning of the decline in confidence and the beginning of the decline in membership. The model is likely to underestimate the length of this lag since it assumes alternative theories of neutral confidence are always available, while in reality new theories would have to evolve as the crisis developed.

In the next sixty years, crisis becomes revolution. Confidence falls from three-quarters to one-quarter, corresponding to a shift from fairly strong belief in the truth of the theory to an equally strong conviction it must be false. Puzzle solving nearly ceases as the number of anomalies grows to nearly twice the level of solved puzzles. The fraction of practitioners committed to the paradigm falls from near total dominance to less than half of the total field. At the end of three centuries the paradigm is essentially dead. Confidence is nearly zero--the paradigm is now viewed as error and superstition. Membership is approaching zero, still lagging behind confidence as a few extremely committed practitioners hang on despite overwhelming pressure to abandon the paradigm. Such practitioners, like astrologers, would no longer be viewed as scientists by the practitioners of the new paradigm.

In order to test the fundamental hypothesis that it is the gradual exhaustion of the metaphor that causes the revolution, a simulation was performed in which the puzzle-solving time was held constant. The resulting behavior, shown in Figure 10, is quite different. The first sixty years look very similar to the first run: confidence increases rapidly as there is an initial burst of progress and a low level of anomalies. The level of anomalies increases from zero to about three (note the different scales for anomalies and solved puzzles), but soon level off well below the acceptable level. The number of practitioners rises to one hundred percent of the field, and there is steady growth in solved puzzles. Because anomalies do not increase, confidence continues to rise indefinitely. Without the gradual increase of puzzle difficulty, normal science continues forever while the paradigm comes to be regarded as absolute truth.

In a second test, the impact of confidence on the anomaly recognition time was eliminated. Now the positive feedback that suppressed anomalies in the two previous runs is severed; rising familiarity and confidence no longer condition practitioners to 'see' what the paradigm suggests they should see.

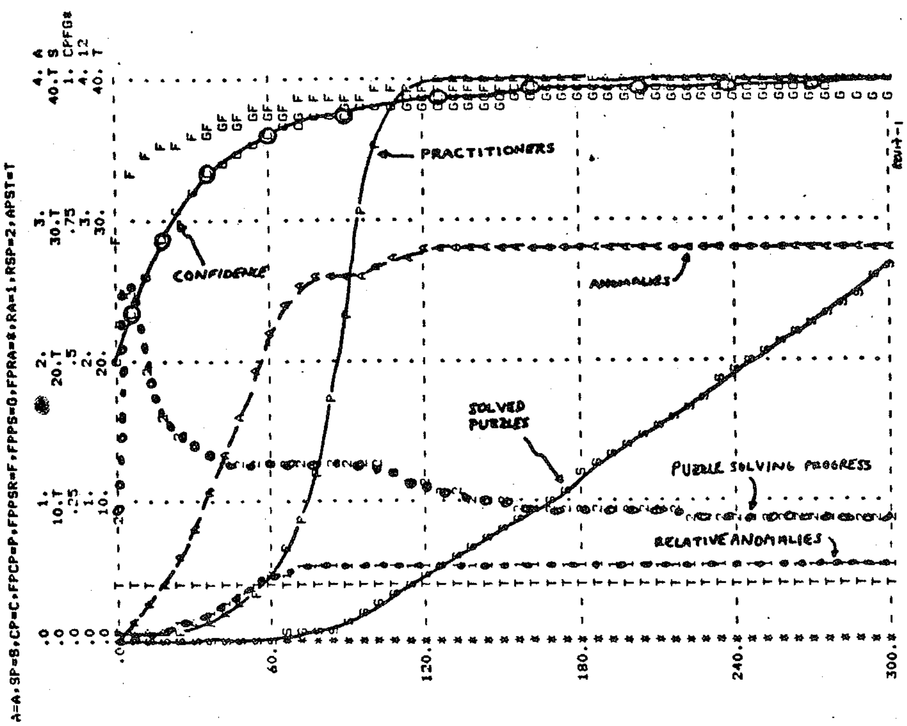

Figure 10

The result, Figure 11, is striking in that the overall behavior is extremely similar to that of the base run. The major difference is one of timing--the lifecycle is shorter. Confidence does not increase quite as high as before, and membership never reaches one hundred percent. The test suggests that the degree to which a paradigm conditions a practitioner's perception of reality is not essential in causing revolutions. It does seem to be important, however, in determining the effectiveness of research. Without the effect of confidence on anomaly recognition time, the number of anomalies increases much more rapidly than in the base run, causing confidence to peak and decline just as normal science gets underway. Interestingly, the growth and decline of practitioners still lags confidence by about sixty years. Though membership never reaches one hundred percent, it peaks just when confidence is falling the fastest.

The paradigm has died after two centuries, two-thirds the time it took in the base run. Thus, it could be argued that the overall rate of progress of science could be enhanced if practitioners did not tend to become rigid in their expectations and perceptions as science progresses, certainly an appealing hypothesis. However, after two-thirds as much time, the paradigm has produced only three-eighths as many solved puzzles, a reduction in productivity of nearly 50%. The explanation for the reduced productivity lies in the emergence and crisis phases of the life cycle. In both Figure IV-6 and the base run, emergence and crisis phases require approximately the same length of time. The major difference between the two simulations is the drastic reduction in the period of normal science when confidence is divorced from anomaly recognition.

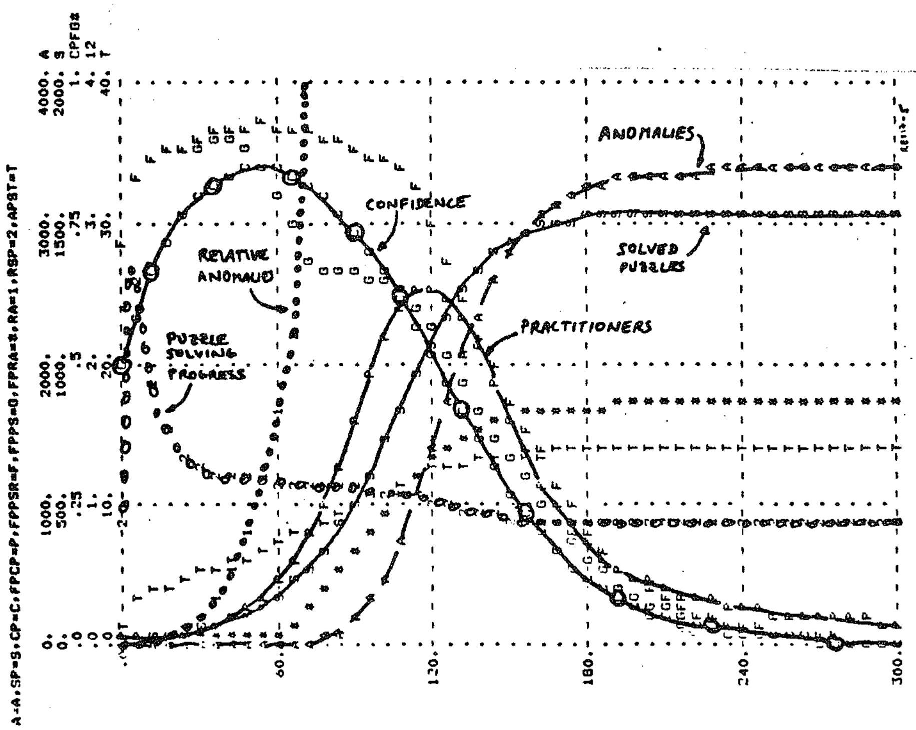

Figure 11

When anomaly recognition time remains low, practitioners are more open to novelty and less bound by traditional ways of seeing, so new theories arise more readily. At the same time, the old theories, because practitioners never achieve the narrow focus normal science depends on, do not probe as deeply into nature as they might. When rising confidence suppresses the appearance of anomalies, radically different metaphors are more likely to appear since the anomalies that are recognized strike into the fundamental tenets of the paradigm. The behavior of the model lends support to Kuhn's statement that "resistance to change has a use":

> By ensuring that the paradigm will not be too easily surrendered, resistance guarantees that scientists will not be lightly distracted and that the anomalies that lead to paradigm change will penetrate existing knowledge to the core. [28]

In another test, the metaphor was assumed to be fully four times as strong as in the base run, but the life of the paradigm was prolonged only a century, or by 33%. Further, the number of puzzles solved was only a little more than double the number solved in the base run. Because the difficulty of puzzles increases more gradually with solved puzzles, there is less resistance to puzzle solving and the rate at which puzzles are solved is higher. Thus, the puzzle-solving time rises nearly as quickly as before, causing anomalies to appear just a few years after they accumulate in the base run.

To test the sensitivity of these results to the particular numerical assumptions of the model, numerous simulations were performed in which virtually all parameters were varied, typically by factors of 2 or more. None of the tests changed the behavior of the system. A few changed the timing, especially the duration of normal science. Most, however, had but little effect on anything. In all these cases, the changes in timing and puzzles solved were minimal because of compensation from within the system.

5. Conclusions

The theory presented here, as tested by its model representation, shows that the observed life cycle of scientific theories arises within the structure of science, from the ordinary day-to-day activities of scientists. It is a systematic phenomenon with identifiable causes. These causes are feedback processes. It is not necessary to invoke competition between theories or "great men" hypotheses to account for scientific revolutions.

The model supports the hypothesis that the cause of revolutions is the gradual exhaustion of the root metaphor defining a paradigm. It is not necessary to assume metaphors are finite, only the weaker assumption that they gradually become more difficult to apply as they are used farther from their realm of formulation.

Sensitivity tests reveal the behavior of the system to be dominated by its internal structure. Large variations in initial conditions and parameters produce a remarkably small range of variation in behavior. The insensitivity is due to compensating feedback mechanisms deeply embedded in the system, suggesting historical circumstances, personalities, diverse content, and sheer luck play a much smaller role in the broad evolution of science than is commonly thought.

On another dimension, since the theory is largely a representation of Kuhn's vision of science, it shows Kuhn's theory to be dynamically consistent, that is, it shows the behavior he sets out to explain and understand can actually be produced by the forces he postulates. And in a larger sense, the results show how formal models can be used to test theories even when those theories are stated in entirely qualitative terms, at a high level of abstraction, and in a context entirely divorced from explicit dynamic analysis.

Notes

1. Paul Feyerabend is the leading advocate of anarchism in science. See his Against Method, London: NLB, 1975.

2. Thomas A. Kuhn, The Structure of Scientific Revolutions, Chicago: University of Chicago Press, 1970, second edition, p. 2.

3. Masterman notes 21 distinct senses of the term "paradigm" in Kuhn, and divides these into three main categories, metaphysical, sociological, and artifact paradigms. Though similar to the categories used here, there is a difference in emphasis. See Margaret Masterman, "The Nature of Paradigms", Lakatos, Imre, and Alan Musgrave, (eds.) Criticism and the Growth of Knowledge, Cambridge: Cambridge University Press, 1970.

4. Kuhn, p. 10.

5. ibid., p. 43.

6. Kuhn uses the example of the electricians to illustrate a science before the emergence of its first clear paradigm. The lesson, however, is the same since during a crisis, the science reverts to a state much like the pre-paradigm state, as Kuhn himself notes.

7. ibid., pp. 15.

8. ibid., p. 18.

9. On the mutability of facts, see Kuhn's discussion of the law of fixed proportions in chemistry, pp. 134-135.

10. ibid., p. 78-79.

11. Richards, I.A. The Philosophy of Rhetoric. New York, 1936, p. 92.

12. Goodman, Nelson. Languages of Art. Indianapolis: Bobbs-Merrill, 1968, p. 80.

13. Turbayne, Colin. The Myth of Metaphor. Columbia, South Carolina: University of South Carolina Press, 1970, (revised edition).

14. Kuhn, p. 113. For the role of analogies in paradigms, see pp. 14-15.

15. "Transfer of schema" is Goodman's term (pp. 71-80 of [12]). Gilbert Ryle calls the transfer "a calculated category mistake"; Turbayne calls it "sort crossing."

16. Black, Max. "Metaphor," Proceedings of the Aristotelian Society. LV, 1954-1955, pp. 273-294.

17. Kuhn, p. 15.

18. ibid., p. 122.

19. A complex philosophical problem is raised here: the problem of "the literal" or "the true" from which these models are abstracted or the falsity of a metaphor established. The solution to this problem is left to the reader, who is encouraged to adapt the theory of metaphor to any brand of epistemology desired. One solution consistent with the theory of metaphor and paradigms is that "reality" *per se* cannot be directly perceived, that is, that the sky must always be seen through some sort of smoked glass. Though emerald goggles might be traded for inverting lenses, as in a paradigm change, it is not possible to gaze with the naked eye on the stuff metaphors are made of. Essentially, the view is that of Lao Tze: "The Tao that can be spoken is not the true Tao."

20. Black, pp. 284-285.

21. Turbayne, p. 69.

22. Kuhn, p. 111.

23. Goodman, p. 80.

24. The notion that anomalies arise from dying metaphors sheds an interesting light on a curious aspect of science. When practitioners become disenchanted with a paradigm, its fundamental tenets start to look like tautologies. Poincare notes during the crisis of Newtonian mechanics that "F=ma" is not a law of nature but merely a definition, since any one of the variables can only be defined in terms of the other two. Similarly, biologists upset with evolutionary theory say the doctrine of differential reproduction states nothing more than "those who have the most offspring have the most offspring." The same is true of metaphors that have been brutally killed through overuse. Consider Gertrude Stein's famous quip, once called "an epitaph for a dead metaphor": "A rose is a rose is a rose."

25. Masterman, pp. 83, 80.

26. The crucial experiment or falsificationist view is associated with Sir Karl Popper. See his Conjectures and Refutations. Also see Lakatos, Proofs and Refutations.

27. As an example of paradigm-induced "blindness", Kuhn (pp. 62-65) discusses the Bruner-Postman experiment in which subjects shown anomalous playing cards (e.g., a red six of spades) were unable to identify the card as an anomaly. Even with long exposure times, most people unhesitatingly assigned the card to one of the categories prepared by previous experience, i.e., they said "six of spades" or "six of hearts."

28. Kuhn, p. 65.