Laboratory of Economics and Management
Sant'Anna School of Advanced Studies

Piazza Martiri della Libertà, 33 - 56127 PISA (Italy)
Tel. +39-050-883-343 Fax +39-050-883-344
Email: lem@sssup.it Web Page: http://www.lem.sssup.it/

LEM
Working Paper Series

Bounded Rationality, Cognitive Maps, and Trial and Error Learning

Richard R. NELSON

Columbia University, USA

2005/28 December 2005

BOUNDED RATIONALITY, COGNITIVE MAPS, AND TRIAL AND ERROR LEARNING

Richard R. Nelson

Columbia University

Bounded Rationality, Cognitive Maps, and Trial and Error Learning*

I. Introduction

The term "bounded rationality" is meant to connote the reasoning capabilities of an actor who, on the one hand, has a goal to achieve, and goes after his objective with an at least partially formed theory about how to achieve it (this is the "rationality" part of the concept), and on the other hand that the theory is somewhat crude, likely will be revised at least somewhat in the course of the effort, and that success is far from assured (this is the meaning of the "bounded" qualification to rationality). Both aspects of the concept seem necessary to capture what we know about human and organizational problem solving, in a variety of different arenas.

The bounded rationality concept is employed, sometimes with that name and sometimes without, in many places in social science research. It shows up in several strands of cognitive science (explicitly of course in the writings of Herbert Simon and colleagues, e.g. Newell and Simon, 1972). It is central in recent writings about the role of business strategies in guiding what firms do, and sometimes leading them into dead ends (two apt examples are Leonard-Barton, 1992, and Tripsis and Gavetti, 2000). Many recent analyses of how engineers solve problems and design things fit this mold (Vincenti, 1990, is wonderful on this). And the assumption of bounded rationality of course lies at the heart of the evolutionary theory of economic change that I developed with Sidney Winter (Nelson and Winter, 1982)

* The author is grateful to Giovanni Gavetti, Annetine Gelijns, Paul Nightingale, and Sidney Winter for helpful comments on an early version of this paper, and to the Sloan Foundation and the Merck Foundation for research support

My focus in this essay as on individual and organizational problem solving, efforts to find a satisfactory or a better way of doing something in a context where present practice is not deemed fully satisfactory. While I believe that the perspective I develop can provide quite general insight relevant to analysis of problem oriented search under bounded rationality, the modeling is tailored to suit the empirical arena where I have done most of my own work, efforts to advance the performance of a technology, and in particular my current focus of that research, medical innovation. In the analysis that follows, search for a satisfactory or a better way of doing something is oriented, and limited, by a theory held by the actor. Empirical exploration of alternatives at any time is treated as like going down a path, which current theory suggests is promising, finding out where it in fact leads, and then perhaps trying another path. As a result of what is learned in exploration, theory may be revised. In turn this may lead to choice of different kinds of paths. What our explorer is able to achieve ultimately depends on the efficacy of this interactive learning process, and on luck.

The model puts a spotlight on two intertwined variables. One is the ability to see and recognize key path characteristics, and thus to be able to discriminate among different paths, and potentially to control the path one is on. The other is the ability to develop a theory, a map to guide exploration, based on what one has learned. The former capability is highly relevant to the latter. Indeed, in fields like medicine the basis for a significant improvement in practice often is laid by enhanced ability to make fine grained discriminations, or a more effective way to control and fine tune action taking, and the improvement in understanding (theory) that this enables..

Note that this characterization of problem solving sees the actor as both theorizing, and engaging in trial and error learning. Gavetti and Levinthal, 2000, and Gavetti, 2005, similarly stress that search tends to involve both knowledge which guides search, and fumbling.

Recognition that both are involved leads one to see concepts that are widely invoked in the literature on problem solving in a somewhat different light. Thus within this framework the "local" of "local search" needs to be understood as theoretically or cognitively local. Similarly, whether a topography is smooth or rugged depends on our problem solver's cognitive mapping. The "natural trajectories" idea in Nelson and Winter (1982, chap 11), and in Dosi, 1982, stresses the importance of theory in identifying paths that tend to lead in the direction one wants to go. The analysis here builds on that conceptual foundation.

Another conceptually interesting, and very real, phenomenon that becomes clear under the formulation I propose is that theories need not be causally valid to be useful in guiding action taking. On the other hand, the practical value of a theory that is not connected with the actual causal forces at work may be very context specific, and holding to that theory may lead to disastrous choices if the context changes. Theories that have the causal factors basically right are less likely to be vulnerable to such problems.

While I believe my formulation is a particularly apt characterization of the cumulative efforts to improve technology in areas where, as in medical practice, over the years there has been significant progress, I also believe that it sheds light on why it seems so difficult to improve practice cumulatively and significantly in a number of other arenas, for example in education or, I would argue, business management. More generally, the formulation helps to illuminate some of the key reasons why progress has been so much more rapid in certain fields of activity than others.

I proceed as follows. In Section II I lay out the general analytic formulation. In Section III I narrow the focus and explore a highly stylized model, in order to develop concrete if abstract examples of some of the general analytic arguments presented earlier.

Then, in Section IV and V I broaden the analysis, adopt a less abstract style, and reflect on the factors that make problem solving in a field hard, or relatively easy. The examples I use will be drawn mostly from the history of technology, and most of these will be concerned with the advance of medical know-how. However, I also will draw contrasts with search and problem solving in the fields of business practice and in education.

II. The General Model

Consider the following general problem. Our actor faces a goal oriented task that must be done periodically. Each time he performs the task his objective is to get as close as possible to his goal. He faces a set of S paths of which he is aware and believes are open to him, that vary in how close they get. How close a path gets to the goal is measured by a figure of merit, M. The actor wants to pick a path with as high a value of M as is possible. His learning task is to come to identify paths with high expected M.

The task environment may be a constant, at least for a considerable period of time, or it may be variable. If it is variable, what paths are best may be a function of the environmental state. The actor's learning task clearly is more complicated under these latter conditions than when the task environment is a constant. He either needs to learn how to identify different contexts, as well as a set of context specific guides to action, or find a broad guide to action that works reasonably well in all or most contexts he will face. To facilitate coherent development of the main argument, for the moment, I hold off consideration of discrimination among contexts, and assume that our actor proceeds unaware of or unable to distinguish different contexts; later I will open up the analysis in this regard. Under these restrictive assumptions our actors learning task, then, is to find a way of choosing paths of action, among those that he believes open to him, that work well on average, given the possibly broad range of contexts in which he may have to operate.

Each possible path is marked by the values taken on by N characteristics or attributes. Some of these characteristics may simply be present or not on a path; others may have a number of possible values. Our actor is able to see at least some of these attributes, before committing to a particular path, but not necessarily all of them. At any time he has a set of beliefs about what visible attribute values mark good paths (and may have some conjectures as to desirable values of attributes he presently is unable to see, but in the future might learn how to). Our actor can not in general control the exact path he takes, but only that that path have a set of attributes. If there is more than one path that has the attribute values our actor believes mark the best paths, which path actually gets taken is determined randomly (equal probability) across the set of paths with those attribute values.

Thus what happens as a result of the actors choice is determined partly by his deliberate choice, and partly by chance. It is as if a choice by the actor involved picking a particular urn, defined by the visible attributes chosen and under control, and containing all paths with those markings, from which random draws are taken, with replacement (for a more extensive development of the "urns" model, see Nelson, 1982). If our actor keeps choosing a course of action with a particular set of controlled attributes, he will achieve a mean payoff. But there may be considerable variance. Part of that variance may be due to the fact that the paths actually taken vary. Part may be due to the fact that there are different context conditions that the actor does not notice.

The "rationality" part of the bounded rationality characterization of reasoning in problem solving connotes, correctly I believe, that individual and organizational actors generally have some reason for what they do, although in a familiar context behavior may be largely automatic. I will interpret the actor's choice of paths with a particular set of attribute values, a particular urn, as being the result of a "theory" held by the actor. I do not want to be too constrained here in the interpretation of what a "theory" means. A theory about good paths may have backing in scientific understanding, or it may not. It may simply reflect the actor's experience. The theory may be deep and provide an explanation for why particular kinds of paths are better than others, or it may simply be about the nature of good paths. The empirical evidence supporting the theory may be strong, or weak. The theory may be no more than a hunch. In any case, the important thing about our actor's theory is that it identifies the characteristics of paths that he thinks lead to good outcomes.

As I indicated earlier, the analysis here is strongly motivated by my developing understanding of innovation in medicine. From this orientation, the particular paths can be interpreted as different treatments for a disease, or at least a set of symptoms. Given those symptoms, our actor-physician's choice problem is to come to find paths-treatments that are likely to have a high figure of merit, given those symptoms.

As the formulation is set up, there clearly is a lot of room for learning by doing. That learning by doing may simply come as a byproduct of doing. However, there also generally are opportunities to pick a path to follow in part at least because doing so will help test or fine tune prevailing theory, or because such action will enable the exploration of a somewhat different theory. At this stage of the analysis, I do not want to get into the issue of possible trade-offs at any time between choosing a path which, given what the actor then believes, promises the highest expected figure of merit, and choosing a path that is likely to increase knowledge, but at an opportunity cost in terms of expected merit this time. Later in this essay I will distinguish between search that proceeds off-line, as in a separate R and D activity, and on-line learning by doing, and their often important interactions. But in this and the following section I will repress these complications and simply assume that search goes on as long as our actor believes that there is something important yet to learn.

If there are any costs to searching, and our actor is confident that his prevailing theory is about as good as can be achieved, then there is no point in experimenting. However, if he has reason to suspect that there is something not quite right, or incomplete, about his current understanding, and can see ways that exploration might improve his understanding, then there is incentive to do such experimenting.

Search to improve theory can be oriented in a variety of different ways. For example, in the actor's mind the objective of search may be to develop a theory that discriminates among different path markers in a more fine grained way. Or the relevance of markers not considered under prevailing theory may be explored. Or, and here I loosen my earlier assumption, attention might be given to trying to distinguish among different context conditions.

Given a broad orientation, there are a variety of possible search strategies. I suggest that the concept of local or neighborhood search can be interpreted as a particular kind of search strategy. The concept is based on the notion of a starting point, under our assumptions here associated with a particular path or set of them that have been taken as a result of a broad theory held by our actor, and then exploration as to whether some marginal change in orientation might increase expected M. My use of the term "path" to denote a course of action suggests that geography matters, and a theory about the characteristics of good paths might be that they tend to start in a particular geographical area. Under that theory, geographical proximity is a natural way to think of closeness or neighborhood, and local search might involve experimentally shifting search to an area that is physically proximate to the one that had been being employed as a starting point.

But a theory about the characteristics of good paths may have nothing to do with the geographical location of the start of the path. The broad theory may be that paths that go through dense woods tend to be good ones. and that starting in a grove of trees is a good indicator that the path will spend a lot of its course in the woods. Under that theory, our actor might think of attending the number of trees in the grove at the start of a path. Assume he gets close to his goal after taking a path with say X trees in its starting grove, and recognizes that number. He might then be interested in exploring whether paths with a few more or a few less trees in the grove at the start might yield even higher expected M. Those paths might not be nearby geographically, but would be "nearby" in their markers. Local or neighborhood search seems generally to be regarded as theory free. However, I am proposing here that notions of nearness, and similarity, are very much influenced by the background theory the actor has.

Consider, for example, the concept of similarity of chemical elements. If a chemist is looking for an element that has many of the same reactive characteristics as the one he presently is working with, but which may have fewer drawbacks, he will be drawn to other elements on the same column of the periodic table, a similarity defined in terms of the number of electrons in the outer shell. A physician prescribing for angina, who has a patient that has reacted badly to a particular beta-blocker, likely will try another beta-blocker. In both of these cases, what is similar, what is a local change, is highly "theory" dependent.

Often the concept of neighborhood search is linked to the concept of hill climbing. The notion of hill climbing assumes a space of alternatives that can be connected in a chain, so that it is feasible for the actor to move from some initial point "A", to some point in its neighborhood, say “B”, where the payoff may be higher, and if it turns out to be higher, then move from point B’ to point “C”, which is in the same general direction from B as B was from A, etc. Like the concept of neighborhood search, bill climbing generally seems to be regarded as proceeding in a natural physical space, and as a theory free activity. But, following the line of argument above, I propose that, in problem solving, hill climbing generally should not be conceptualized as proceeding in physical space. Thus our actor may pose the question in terms of whether the number of trees at the start of a path matters, and move from two trees in the grove paths to three tree paths to four tree paths, etc.

It may turn out that, while merit changes erratically as one moves from one path to another that is nearby physically, average merit is smoothly related to the number of trees that mark the start of a path. What is a rugged topography and what is not, just as what is close and what is not, depends on the theoretical lenses through which one views the terrain. (Levinthal, 1997, and Gavetti and Levinthal, 2000, have interesting models of seach on rugged landscapes).

To return to our example in chemistry, the search for a better element to stick into a molecule likely would proceed over extremely rugged terrain, in the sense that M would bounce around a lot, if our molecule designer tried one element, then the one in the periodic table heavier than it, etc. But the terrain might be quite smooth if the first larger element considered seriously was one with the same structure of the outer electron shell, then a still heavier element with that same configuration, etc. More generally, it is clear that in their exploration (either actual physical trial or by modeling) of designs for a device, engineers try hill climbing regarding parameters of the design their theoretical understanding suggests are important.

The perspective I am developing here flags attention to the fact that there may be several different theories that are roughly consistent with the empirical evidence regarding the characteristics of good and bad paths. Of these theories, some may work for quite spurious reasons, and be quite context specific.

Thus our actor's theory that paths that go through the woods are on average better than paths that go over fields may be supported by considerable experience. But some of the paths through the woods go by lakes and others do not. And in fact it may be passage by lakes that gives a path its good properties. The reason that paths through the woods pay off more on average than paths through meadows may simply be that a much larger percentage of the former than the latter go by lakes. If this is learned, our actor might well be able to do better with this new and more accurate theory, at least if there are visible indicators that are more reliable than a grove of trees at the start in telling whether a path will go by a lake.

An interesting real example is the growing perception, during the 18th and early 19th centuries, among physicians concerned with epidemics, that malodorous air, miasma, was an important source of contagious disease, and that cleaning up the urban environment helped to prevent epidemics. This "theory", which clearly contributed to good public policies, existed well before there was general acceptance of the germ theory of disease (see e.g. Porter, 1997, chap. 13). However, when the germ theory of infections disease came to be accepted and understood, and discriminations made among different diseases and their carriers, policies could become much more focused on preventing the presence of the carriers of various kinds of diseases.

I now want to open up an issue that has been repressed in the preceding discussion; the achievement of a high average M by our actor may require ability to recognize different contexts and pick paths that are appropriate to each. A "by a lake" theory of the key characteristic of good paths may work well in a context where there are lots of lakes, and no rivers, but not be helpful in a context where there are no lakes. Conversely, if rivers have the same favorable properties as lakes, a “good paths go by a river” theory may be fine in a context where there are no lakes and several rivers. If our actor always is in one context or the other, he only needs to have one of these theories, the right one for the context. But if he sometimes is in one context and sometimes in the other, he needs to be able to recognize the difference and respond to it correctly.

Alternatively, he might do fine if he had a general theory that what counts about paths is whether or not they go by a body of water, and paths had visible markers that reliably signaled that. Coming to understand the general theory sounds both more parsimonious and deeper. However, if there is no reliable general marker, knowing that good paths go by some kind of body of water may tell our actor little about how to identify paths that go by lakes if he is in lake country, or by rivers if he is in river country. At the same time, there might well be reliable markers for whether a path goes by a lake in lake country, and others for whether a path goes by a river in river country, and ways that our actor can tell where he is. In this case, holding the correct general theory may add nothing to our actor’s ability to choose good paths. I would like to argue that in many arenas of problem solving, even many that are illuminated by a formal science, general theory often points only very broadly towards good practice. The devil is in the details, and understanding of the details is won only through detailed empirical exploration guided only broadly by that general theory.

The germ theory of (many) diseases clearly was an important breakthrough, and greatly aided efforts to find ways to deal with human disease. But there are many different varieties of infectious disease, and public health measures that are effective for one may not do much to stop another. Cholera bacteria are carried in contaminated water, and effective campaigns to stop cholera (and typhoid) focus on keeping the water supply clean. Malaria is carried by mosquitoes, and effective public health programs are focused on this vector.

These last examples highlight that in trying to make progress in an area a lot depends on what one can see, more generally on one's ability to discriminate among pathways. In turn, these capabilities depend on the instruments and techniques one can use to see. The development of the germ theory of disease, and the evidence that the medical community found convincing, depended on instruments like the microscope, and a significant number of carefully controlled experiments and demonstrations. This is equally true regarding the identification of the particular microorganisms responsible for particular diseases, and exploration of their vulnerabilities.

I want to highlight that in the examples here, enhanced ability to see and discriminate contributed to stronger ability to control what one actually was doing, the path one actually was on. But ability to control can be considered as a capability in its own right. From this perspective, variation in the paths actually taken by our actor may simply reflect meanderings that, while he knows they should be avoided, he cannot always control. Greater ability to see differences may have little value in problem solving if one cannot control one's actions sufficiently to take advantage of enhanced ability to discriminate. However, in what follows I will assume that ability to discriminate, and ability to control, go together.

In my discussion thus far I hardly have mentioned the discovery or development of new methods, new pathways, like the use of new antibiotics and surgical instruments that have been so important in the advance of the efficacy of medical practice, and in other areas of human activity where progress has been sustained and significant. Technological advance of this sort is an enormous topic in its own right and can not be dealt with generally as part of a paper focused on the importance in problem solving of ability to see and discriminate, and control the actions taken. However, I will argue later that recognition of exactly these features of problem solving can go a long way towards illuminating both the sources and the roles of new techniques.

In the following section I will repress the role of the development of new techniques in problem solving. However, these matters will be dealt with extensively in the concluding two sections.

III. A Specific Example

I present now a highly abstract version of the general analysis sketched above, for the purpose of making clearer some of its implications. We observe this particular example from a position of full knowledge of the problem facing our actor, so we can appreciate his difficulties with the problem, see what can lure him to false conjectures, and the ways he might be able to develop a good theory and course of action. In this particular case it is assumed that the context facing our actor is a constant, or at least is a constant over the period of time we consider in the formal analysis. On the other hand, one of the questions we explore is the relevance of what the actor has learned in this context if the context changes.

As displayed in Figure I, there are 36 actual different paths of action. No new ones are introduced or discovered over the course of the analysis; there is no technological innovation possible in that sense. The merit of each available pathway is shown in the upper left hand corner, which we can see but the actor can't until he has gone down that particular path. In this particular model I assume that once a path is traversed, the actor has no trouble assigning a merit rating. His problem is in knowing what particular path he actually took. He only knows certain characteristics or signs of that path. Thus what he can see influences his ability to control what he actually does. For this reason, among others, just what he is able to see matters a lot.

I, the author, have laid out the alternative paths on an orderly grid which you, the reader, can see. However, our actor has no map and has no compass, and location in Euclidian space is something he cannot fathom. On the other hand, he can see, if he pays attention, that paths, or their starting points, differ regarding whether they have an X mark in the middle, or a rail line like mark running across the bottom, or neither of these. There are other features, or finer features, that differentiate the alternatives, and these are displayed in Figures II and III. But at the start of the problem solving venture in question, our actor cannot make out these differences.

The reader may verify that the average merit from picking a path at random is 1. If our actor started by randomly picking paths, it might take him some time before he pinned this down, because there is considerable variance. However, perhaps the variation in achieved merit in itself would start him thinking about whether certain identifiable features might matter, for example an X mark or a rail line mark. If he explored one or another of these possibilities, he would find out that those variables did matter. Paths marked by an X have an average merit of 15/6; paths with rail lines have on average a merit average of 21/8. Either of these is a useful theory to guide choice, in this particular context at least. And in this context, our actor might find it difficult to decide which of these theories is a better guide to action, because there is not much difference between the average merits of the paths pointed to by the two theories.

This is interesting because, as will become evident shortly, the "marked with an X" theory of high merit paths is completely spurious causally, and works in this context only by chance. It likely would be of no use at all if the context changed. In contrast, the "marked with a rail line " theory does have a relationship with the underlying causal mechanisms.

I note that, in the spirit of the discussion of the prior section, if our actor can only tell whether a path has a rail line like marker, or an X, he has limited control over the path he actually takes at any time. A consequence is significant variation in the values of M achieved from taking paths that, for all our actor knows, are the same.

Figure II displays what our actor can see after he gets eyeglasses that enable him to see more clearly. One thing he now can see is that, while he missed it without glasses, some of the paths marked by an X also have an x, while others do not. Being able to make this discrimination clearly is helpful, at least in this context. Paths marked by an X but without an x have an average payoff of 11/3. If he can find this out by experimenting, and sticks to paths so marked, this beats the average he can get if he guides his choices by whether a path has a rail line mark or not. However, while a theory that good paths are marked by a simple X will help guide him in this particular context, the theory is spurious regarding true causality.

Good glasses also enable our actor to see more finely into the markers on paths that he used to think were characterized by rail line markers. He now can see that these really are a set of connected crosses. He also can see that the set so marked seems to differ in terms of the number of crosses or cross bars. What he can make of that knowledge depends on how finely he can discriminate. If he can actually count the number of crosses or cross bars, and experimented to find out the relationship between average merit and number of crosses, he might be able to home in on paths with seven crosses, which have an average merit of 13/4, not quite as high as the paths marked by a simple X, but pretty close.

However, it would not be easy to reach that understanding. There is a large number of paths marked by collections of crosses, and several for each number. Attempts at hill climbing through finding the average merit for paths with a given number of crosses and then moving up or down one wouldn't work very well. First of all, there still is more than one path associated with any marker our actor can identify, and therefore different values of M in going down paths of that class. At this level of discrimination, it is average M that counts, but getting the averages close to right might take a number of trials. Second, the relationship between average merit and number of crosses is far from smooth. The topography is rugged, when looked at from the perspective lent by a theory that focused on the number of crosses..

Now equip our actor with an even better set of glasses. He now can see, as shown in figure 3, that some of what he thought were crosses on different paths in fact are asterisks. He could then explore the merits of the paths that had some asterisks. If he had any luck, he would quickly come to a theory that asterisks marked good paths. The average merit of paths with asterisks in 21/6. If he reached this conclusion, he might then consider whether the number of asterisks mattered. This would be an easier task than exploring whether the number of undifferentiated cross like things mattered, first of all because the six paths with asterisks each have a different number of them; no intra class variation to make things complicated. That is, finer discrimination gives the actor better control of what he actually does. And second, here hill climbing works. The topography is smooth. Merit goes up linearly with the number of asterisks.

In fact, this is the only structural systematic relationship between attributes and merit in the set of paths. I know that because I built the model. There are six paths with asterisks. In each of these merit equals the number of asterisks. The merit for the other thirty paths is either zero or one, half each. Any relationships between attributes other than the number of asterisks and merit are due to chance, or to the fact that certain attributes are partially tied to the number of asterisks. The particular context characterized in Figure III was achieved by scattering the paths so characterized within the grid more or less at random.

It is useful to reflect a bit on the robustness of the different theories that, in our account, the actor used to guide his choice of paths, all theories that, in this particular context, helped him to achieve an average level of merit that was better than chance. The theories relating to the presence of an X, which our actor could see without glasses, or the theory about the high average merit of paths marked by an X but without an x, which he could see with glasses, are as I stated earlier completely spurious causally. They are helpful in this particular environment. But lay out the random elements of the grid in another way, and guidance by such theories is likely to be worthless, or worse.

The "marked by a rail line", or connected cross like things, theory will provide some guidance in another environment, because all paths that are marked by asterisks also would be seen as being marked by a rail line or connected crosses, if the actor had no glasses or weak ones. And in most contexts, having a large number of cross like things likely would be a guide to good paths, because paths with a large number of asterisks are in fact good paths.

With the vision of hindsight, you (the reader) know that the true guiding theory is simple. Merit equals the number of asterisks for those paths so marked, otherwise zero or one with equal probability. If our actor arrives at that theory, he is in good shape even if the context changed. However, note that within the context we have just considered, it took the development of ability for fine discrimination for our actor to have any possibility at all to arrive at the correct general theory; without that capability while our actor might develop a theory that would enable him to do well in the context in which he is operating, there is no way that he can achieve the correct theory of what determines merit. And even if he had high quality glasses, it might be very difficult for him to find the correct general theory. Among other things, in this context there are other theories that provide almost as good guidance.

But if the context changed, how well our actor would do would depend very much on the theory he had developed in the earlier context.

IV. On Factors that Enable Progress in Problem Solving

In this section I lower the level of abstraction and reflect on what enables difficult problems ultimately to be solved reasonably well. My focus here is in persistent problems relevant to a significant group of individuals or organizations, on which progress is made not all at once, but rather from a series of partial solutions that enable the goal to be approached more and more closely. This aptly characterizes the advance of most technologies.

I want to start by proposing that the ability to see fine grained differences, and to control the path taken based on fine grained discriminations, a key variable influencing the power of search highlighted in Section III, shows up as important in many empirical studies. While the point made there would seem obvious, almost trivial, there is much more than tautology here. Ability to see, discriminate, and reliably select and control that the actions have particular characteristics, is something that cannot be taken for granted. In many cases creating the capability to discriminate and control has been the key to becoming able to solve the problem.

Consider, as a famous example, the inability of navigators of ships at sea to assess accurately their longitude, prior to the development of the chronometer. As a result, those responsible for choosing the ship's path often could not judge accurately or control the longitudinal dimension of the path they were taking, with occasional disastrous consequences. The development of the chronometer enabled navigators to have much finer control over their ships' actual path, given the course they aimed to take. (For the story see Sobel, 1996). In effect the "subset of paths" on which they might be at any time was substantially reduced.

This was a case where, while improved ability to see and control greatly improved ability to pick and hold to good paths, these developments had no significant affect on broad theory. The importance of being able to measure longitude had been understood for a long time; only the ability was lacking. However, in many cases such improved abilities have led to new theories, or at least permitted them, which in turn led to the discovery of better practice.

Thus prior to the development and acceptance of the microbial theory of infectious diseases, largely as a result of the research and publicity done by Pasteur and Koch, physicians had little clue regarding how to deal with such illnesses. Efforts to cure individual patients usually were worthless or worse. With the acceptance of microbe theory, this changed. But microbial theory could not have been discovered or demonstrated (although it had been conjectured) without instruments and techniques that enabled microbes actually to be seen. In turn that theory, and the observational capabilities that enabled it, set problem solving in the direction of identifying the particular microbes involved in various diseases, again a path requiring fine grained perception and control. Once the particular micro-organism was identified, problem solving could turn towards trying to develop an effective vaccine. While a persuasive theory of why vaccines often worked only came later, experience gave some strong clues as to how to develop a vaccine, once the offending microbe had been identified. And while risky, tests of the efficacy of a new vaccine generally gave sharp results. The germ theory of disease also provided useful focus for efforts to find or create antibiotics, but significant progress in this area took a long time to achieve.

There was an interesting difference between the ability to treat individual patients and ability to devise public health measures for infectious diseases, prior to germ theory. The dominant theory that guided treatment of individual patients by many physicians was based on the notion that illness was the result of the balance of body fluids getting away from what it should be. While for some physicians this theory led to advice to patients regarding a healthy life style, it also often led to treatments for illness like bleeding, that usually did more harm than good. In contrast, as I noted earlier there was a relatively widely held theory about the health problems caused by "bad airs" that led to various campaigns to clean up areas that were malodorous. In some cases this led to the elimination of various public health hazards, and generally led to actions that did more good than harm. This is a good example of a an invalid theory that in fact enabled some effective problem solving, because the factors it identified as relevant were often correlated with the real causes, and dealing with the correlated variables often dealt with the actual culprits. Of course with germ theory, public health policy would rest on a new and more effective foundation.

I have been arguing that one of the important advantages of being able to identify, specify, and control closely the paths of action that are taken is that one then has the possibility of getting useful feedback regarding efficacy. But while ability to fine tune action may be necessary for sharp feedback, it is not sufficient. In the model in Section III our actor had no difficulty assessing merit when he got to the end of a path. However, there are a wide variety of problem areas where, while actions can be tightly specified and controlled, feedback is not sharp or rapid. This is often the case in medicine.

Areas of medical practice differ significantly in this respect. In general, one can learn relatively quickly if a new antibiotic works or not, although it may take more time to learn about possible side effects. However, with treatments that are basically intended to prevent premature deaths, which is usually the case with treatments for early breast or prostate cancers, it can take a long time before one has statistically significant data on the ages of death of a treated group, and a control group. This very slow feedback regarding efficacy clearly adds to the difficulty of making significant progress in these areas.

Another striking feature of progress in medicine, and the evolution of technologies more generally, that is repressed in the model, is that achieving a goal, or a continuing improvement in performance, almost always is the result of work done by many different parties. There are exceptions. Jenner's bet on the efficacy of using materials from cow pox as the stuff for vaccination for small pox was based on his access to folk wisdom, but the idea of the vaccine was his. Similarly. Pasteur's vaccine for rabies was largely his doing. But in many cases the development of a successful means to prevent or cure a disease has been the result of the work of many different people and organizations.

The penicillin case is a good example. Here the original discovery of the antibiotic properties of the mold was done by one person, Fleming. The experimental work on using the substance as a prevention or treatment for infection was done by someone else, Florey and colleagues. And the development of an effective production method was done by still other people. There are many other similar examples in the history of medicine. In virtually all detailed studies of particular important inventions I know about, a history of earlier work that laid the basis for the invention in question is part of the story.

Again, I want to propose that a necessary, if not sufficient, condition for the ability of today's problem solvers to build on the work of yesterdays' is the ability to closely specify and control the paths being worked on. You can't build on someone else's work unless you know the nature of that work in a certain amount of detail, and have the ability to pick up where that prior work ended.

In Sections II and III I ducked the question of just how exploration for better ways of doing things proceeded, while noting that it could proceed on-line or, in some cases, off-line. Ability to experiment, to learn, off-line obviously has major advantages. Doing experimentation on-line involves opportunity costs, if a path other than in the set currently judged most promising is taken in order to explore its merit. If one tries to keep opportunity cost down by not diverging much from known best practice, the range of exploration is seriously constrained, and what one learns has to come largely as result of "learning by doing" where ability to pin down cause and effect relationships may be quite limited. In most of the modern technologies where advance has been rapid and sustained, much of the exploration and fine tuning of new paths is done off-line, in R and D that is separated from actual practice.

The fact that today new medical practices, including the use of new pharmaceuticals, and devices, are largely developed off-line is so familiar that it is not recognized as remarkable. Until recently, all advances in medical practice were achieved in exploration on-line. Jenner's discovery of vaccination is a good example. While the discovery that the availability of citrus fruits could prevent scurvy involved a relatively controlled experiment in which different sailors were provided with different supplements to their standard diet, that experiment was very much on-line. As this example illustrates, it always was possible to experiment on a small number of people, with broader practice for the population as a whole learning from advertised experimental results. However this is very different from doing experimentation off-line, as it were, with application to even a small number of patients dependent on "laboratory" results.

Indeed, prior to the mid-nineteenth century there was hardly any "laboratory" work that led to significant advances in medical practice. A combination of the development of the germ theory of disease, identification of cells as the basic building blocks of living creatures, and significant advances in understanding of biochemistry, changed all that. Nowadays it is expected, indeed required by law, that a new treatment, a pharmaceutical for example, be sufficiently well tested off-line, in the laboratory, or using nonhuman subjects, before actually trying it on humans, that the prospects that the new treatment will be effective are relatively good and the chances of its causing harm relatively small.

Let me flag what by now should be obvious to the reader. The efficacy of off-line R and D depends on the ability to specify closely, to control, what is achieved in R and D. And it must be possible to take what has been learned off-line, and use it on-line with much the same results as was achieved in R and D. As Paul Nightingale (2004) has argued, much of what is achieved in R and D is achieved under tightly controlled conditions. Thus the transfer to practice often involves developing shields so that the relevant aspects of practice resemble the controlled conditions of the R and D setting. Reflect on how much shielding from possible adverse environmental effects goes into the design of high tech products, like a micro-processor . Pharmaceuticals are protected by a outer shell, and conditions of use closely specified.

It also is important to recognize that clinical trials still are a necessary part of the process. And many new treatments fail at this stage. In the language of Sections II and III, what has happened is that the characteristics, the markers, that suggest a path is worthwhile going down now take the form of results in a laboratory, or in an animal model. But there still is the need to go down that actual path and see where it actually ends up.

The cases discussed in this section and earlier make it clear that effective problem solving often involves the enabling or creation of ways of doing things that were not feasible before, or not even envisaged as possibilities. Indeed, such technological innovation has been the basic driving force in those areas of human activity where advance had been sustained and cumulatively great. In some cases the key to advance is new ways of observing or assessing: the better glasses of Section III, the chronometer, or modern medical diagnostic equipment which, among other things, enable much finer discrimination among circumstances and pathways. In other cases the advances involve new paths towards the objective: new vaccines, antibiotics, new surgical procedures and the apparatus that makes them possible, like heart-lung machines. To encompass these kinds of developments, the model of Section III would have to be enriched to include the creation of new paths, or the unblocking of ones which were known, in a sense, but not regarded as feasible.

I have dealt extensively with innovation, in the sense of the development of significant new ways of doing things, in a number of other publications. Here, in the context of the perspective I am developing that stresses the key roles of observation and precision of action taking in problem solving, I want to limit my discussion to those aspects, elaborating somewhat on some themes I introduced at the end of Section II..

I want to argue that when new effective paths are opened by innovation, an important reason why they are effective often is that they give tight control. I noted earlier that going effectively from what works in the highly controlled setting of off-line R and D to actual practice will often require the shielding of a new artifact from the range of contingencies that might interfere with its efficient operation. But more generally, controlled reliable operating characteristics is an important value in its own right, as I have been arguing throughout this essay. This certainly is so regarding new medical practices that are widely accepted as improvements over prior art. Thus in the development of new drugs and vaccines a considerable amount of attention is given to trying to optimize properties, and an important concomitant to fine tuning design is that the production processes need to be carefully controlled to assure that what is produced is what is intended to be produced. If there were considerable variation in the nature of a drug with a given name, on average treatment using it would be much less effective than treatment using a controlled optimized version. Tests of efficacy would yield a lower mean

M, and considerable variance. Similarly regarding new surgical procedures and the devices used in these, both those involved in the surgery and those implanted, like a pacemaker. While tight control certainly is not a sufficient condition for an innovation to be a significant advance, control certainly is a contributing factor, and in many cases a necessary one.

I want to conclude this section by pointing to two areas of practice where progress has been very slow, and a central reason (in my view) is that the operating practices are very difficult to pin down in any detail. (For a more extensive discussion see Nelson, 2003.) One is education. Consider as an example the "phonics" method of teaching reading. The broad outlines of the phonics method are relatively clear. However, few schools or teachers use it in a pure form, and more generally the details of how reading is taught under its broad rubric differ from school to school, teacher to teacher, and even student to student. For this reason, among others, it is very hard to evaluate the efficacy of using phonics, as contrasted with other philosophies of teaching reading that march under a different flag. This problem regarding educational techniques has been well recognized, and it has been proposed to establish tight standards for particular techniques. But this is easier said than done, partly because it is recognized that the imposition of tight standards might in many cases eliminate the possibility of context specific variation that is needed for efficacy in particular cases. But to the extent that broad flexibility is needed, what is learned though experiment and experience is going to be coarse grained.

I would like to propose that the same thing is true of much of business practice. The M form, quality circles, just in time, all are terms that cover a lot of variety. Part of that variety is intentional, involving tailoring to fit particular contexts, but a good part stems from the fact that the individual actors involved are not clear themselves regarding exactly what they are doing. This makes for a situation in which learning, individual or collective, is not easy.

It also is true that in both of these areas efficacy is hard to assess. But I propose that a considerable portion of that problem results from the fact that practices are not well identified and controlled, and it is very difficult and perhaps counterproductive to do much about that.

V. Remarks on the Nature and Power of the Applications Oriented Sciences

I want to conclude this essay by briefly discussing an apparently alternative theory of why progress has been so uneven across fields of human activity, and proposing that actually the theory I have been proposing is compatible, and deeper. That theory is that the pace of progress in a technical field, or on a class of problems, is dependent on the strength and rate of advance of the sciences that bear on that arena of problem solving.

In my analysis in this essay, I have highlighted the role of strong relevant theory in enabling effective problem solving. In many areas of problem solving, an important part of that theory is "scientific". There is good evidence that strong relevant scientific theory is a major factor enabling practical problem solving efforts to be effective. Several recent studies have shown that fields of technology that have advanced rapidly draw on strong fields of science (see e.g. Nelson and Wolff, 1991).

However, those studies also have shown that the fields of science drawn on by industry in efforts to advance technology tend to be the applications oriented sciences -- fields like electrical engineering, pathology, computer and materials science (Klevorick et al, 1995). The sciences here are in "Pasteur's Quadrant", to use a concept introduced by Donald Stokes (1997). They aim for understanding that will enable better practice. They may draw from other fields of science that have less of a practice orientation, as electrical engineering draws from theoretical physics. But the output of the science is designed to be useful. I do not want to deny the relevance to practice of theory developed without any initial interest in the guiding of practice, as Maxwell's theory of electromagnetic radiation, or Einstein's theory of special relativity. However, in most fields of effective practice that I know about the underlying body of theory, of science, that actually guides problem solving has been created in large part to enable good practice.

I want to argue that the development of a strong applications oriented science supporting an arena of problem solving requires exactly that ability to see, to discriminate, to control action, in practice that I have been stressing above. As the examples I gave earlier indicate, the causal arrow here certainly goes in both directions. It may take a strong theoretical understanding to be able to make the appropriate discriminations among courses of action, to enable them to be evaluated confidently, and to permit practical paths of action to be tightly controlled. But on the other hand, the applications oriented sciences, the engineering disciplines, fields like materials science, while drawing on the more basic sciences, are basically about practice and problem solving relating to practice. Their strength requires that it be possible to pin down practice, analyze it, study how performance relates to the scientific principles of the field. (See Vincenti, 1990 for a fascinating study of aeronautical engineering).

Earlier I proposed that a broad theory seldom provides the detailed knowledge one needs to decide on appropriate practical action. Certainly the advance of engineering design, and medical practice, has been facilitated by the development of better broad theory. But much of understanding relevant to practice in these fields is at some distance away from high theory, and rather involves detailed knowledge of what works and what doesn't.

Even in the applications oriented fields, much of scientific research proceeds off-line. As I noted earlier, effective off-line R and D requires that what is learned in an artificial, and generally highly simplified, context indicates what will work on-line. For this to happen, on-line operation may need to be strongly controlled. But in some fields, this may not be possible. When it is not, there inevitably is going to be a disjunction between what goes on in research, and what goes on in practice. Or, to put it another way, the "science" may not be very useful to the advance of practice.

I am arguing that, while significant advance of a technology requires a strong underlying body of relevant science, the development of a strong field of underlying science depends on whether the technology can be tightly specified and controlled, and evaluated. I do not think this fact is sufficiently appreciated

Thus many thoughtful people have proposed that the slow pace of advance in educational practice could be increased significantly if more R and D were allocated to that objective, particularly if the underlying sciences could be improved as a result of more resources dedicated to that purpose. But for decades it has been well documented that educational practices that seem to work well in an experimental setting do not seem to transfer very well to regular classrooms. Partly this is an accuracy of copying problem, that is virtually inevitable when a practice can be described only in a coarse grained way. It has been proposed that R and D would be more effective if what was tried in an experimental setting were more closely specified and controlled, and copied more tightly. But for reasons I discussed above, this is more easily proposed than done.

It also has been proposed that if we had a better theory of how children learn, both off-line experiment and on-line practice would go better, and that the returns to more basic research funding here could be considerable. The first part of that proposal certainly is true, but the second part may be dubious. It is quite unlikely that we can learn much about how children learn in their natural world from even a strong theory of how children learn in a carefully controlled experimental setting. Perhaps the reason that, except for a few broad generalizations, we don't have a sharp simple theory of how children learn in their natural world, or in school, is that this is very complicated, and different children learn in different ways, and those ways may be very context specific. The scientific methods that have been so effective when the structures studied can be closely controlled don't work very well when they cannot.

It is my strong impression that a very similar situation obtains regarding business practice. To my knowledge, there is not much in the way of experimental laboratories where pilot versions of proposed new practices, or changes from prevailing practice, are experimented with before trying these out on-line in real time. It is well recognized, I think, that the kinds of experiments that social psychologists do that seem to hint at ways to improve practice are usually so simplified relative to the actual context that only a little can be learned from them. More generally, basic theory in this area strikes me as a long distance away from pointing clearly to ways to improve practice.

In this essay I have been trying to point out the strong connections between the effectiveness of problem solving in a field, the strength of the theory that guides problem solving, and the ability to observe evaluate and control practice. In the contemporary world, the applications oriented sciences are an important part of this dynamic system. But the causal arrows go in more complex ways than proposed implicitly by those that advocate that supporting science is almost always the best way to gain practical progress.

1	0	0 +++++++	1	1	0
1	4 X ****++	0	0	0	1 *+++++++
0	1	0 Xx	2 **+++++	1	1
6 X ******+	1 Xx	1	0	0	1
1	1	5 *****++	1	3 Xx ***	0
0	1	0	0 +++++	1	0

Figure 3

1	0	0 ++++++	1	1	0
1	4 X ++++++	0	0	0	1 X ++++++++
0	1	0 Xx	2 +++++++	1	1
6 X +++++++	1 Xx	1	0	0	1
1	1	5 +++++++	1	3 Xx +++	0
0	1	0	0 +++++	1	0

Figure 2

1	0	0 =========	1	1	0
1	4 X =========	0	0	0	1 X =========
0	1	0 X	2 =========	1	1
6 X =========	1 X	1	0	0	1
1	1	5 =========	1	3 X =========	0
0	1	0	0 =========	1	0

Figure 1

References

Dosi, G., 1982, Technological Paradigms and Technological Trajectories: A suggested Interpretation of the Determinants of the Direction of Technological Change", Research Policy, pp 147-162

Gavetti, G., 2005, "Cognition and Hierarchy: Rethinking the Microfoundations of Capabilities Development", Organization Science,

Gavetti, G., and Levinthal, D., 2000, "Looking Forward and Looking Backward: Cognitive and Experiential Search", Administrative Science Quarterly, pp 113-137

Klevorick, A., Levin, R., Nelson, R., and Winter, S., 1995, "On the Sources and Significance of Inter-industry Differences in Technological Opportunities", Research Policy, pp 185-205

Leonard-Barton, D., 1992, "Core Capabilities and Core Rigidities: A Paradox in Managing New Product Development", Strategic Management, pp111-125

Levinthal, D., 1997, "Adaption on Rugged Landscapes", Management Science, pp 934-950

Nelson, R., 1982, "The Role of Knowledge in R and D Efficiency", Quarter;u Journal of Economics, pp 453-470

Nelson, R., 2003, "On the Uneven Evolution of Human Know-how", Research Policy, pp 909-922

Nelson, R., and Winter, S., 1983, An Evolutionary Theory of Economic Change, Harvard Un. Press, Cambridge

Nelson, R., and Wolff, E., 1997, "Factors Behind Cross Industry Differences in Technical Progress", Structural Change and Economic Dynamics, pp 205-220

Newell, A., and Simon, H., 1972, Human Problem Solving, Prentice Hall, Englewood Cliffs,

Nightingale, P., 2004, "Technological Capabilities, Invisible Infrastructure, and the Un-Social Construction of Predictability: The Overlooked Fixed Costs of Useful Research, Research Policy, pp 1259-1284

Porter, R., 1997, The Greatest Benefit to Mankind, Norton, New York

Sobel, D., 1996, Longitude: The True Story of the Lone Genius Who Solved the Greatest Scientific Problem of His Time, Penguin, London

Stokes, D., 1997, Pasteur's Quadrant, Brookings, Washington

Tripsis, M, and Gavetti, G., 2000, "Capabilities, Cognition, and Inertia: Evidence From Digital Imaging", Strategic Management, pp1147-1161

Vincenti, W., 1990, What Engineers Know and How They Know It, Johns Hopkins Press, Baltimore