Modes of Experimentation:
An Innovation Process – and Competitive – Variable

Stefan Thomke[1], Eric von Hippel[2] and Roland Franke[3]

[1] Assistant Professor, Harvard Business School
[2] Professor, MIT Sloan School of Management
[3] Consultant, Biotechnology and Pharmaceutical Industry

Abstract

The outputs of R&D, such as new research findings and new products and services are generated with the aid of specialized problem-solving processes. These processes are somewhat arcane, and have been largely ignored in studies of technical change. However, improvements to them can significantly affect the kinds of research problems that can be addressed, the efficiency and speed with which R&D can be performed, and the competitive positions of firms employing them.

In this paper we first describe the general nature of the trial-and-error problem-solving processes and strategies for experimentation used in the development of new products and services. We next discuss the rapid advances being made in problem-solving methods, and the impact such advances can have on the competitive position of adopting firms. Finally, we offer a detailed case study of the impact one novel experimental method, combinatorial chemistry, is having on the economics of the drug discovery process.

***Keywords*:** Problem-Solving; Experimentation; Technological Innovation; R&D Efficiency; Pharmaceutical Drug Development; Product Development Performance.

1. Introduction

The impact of outputs of the R&D process on firms and industries has long been acknowledged. For example, the major consequences of the development and continued improvement of semiconductors, of computerized manufacturing, etc. have been noted by many. But the outputs of R&D are themselves "manufactured" with the aid of specialized problem-solving processes. These underlying processes have been largely ignored in studies of technical change. However, the application of problem-solving processes represent an increasing proportion of economic activity (Carter 1995), and the processes themselves are improving rapidly both in terms of the kinds and efficiencies of outputs producible. These changes, in turn, are having and will increasingly have an impact on the competitive position of adopting firms.

In this paper we will explore the general nature of the problem-solving processes used in R&D, and the potential impact of novel problem-solving methods on firm R&D performance and competitive standing. We begin (section 2) by describing the problem-solving process used in R&D in general terms, and showing how different experimentation strategies can influence R&D efficiency. We next (section 3) illustrate the rapid rate of advance affecting experimental methods, and then observe that early developers or adopters can have a competitive advantage over rivals, because new methods are often difficult to acquire and use. We follow (section 4) by presenting a field study of combinatorial chemistry – a new method now being introduced into the drug discovery process that promises to make drug discovery both faster and less costly. This case is especially relevant to our topic because more effective drug discovery methods can convey a great competitive advantage to pharmaceutical firms: Present drug discovery processes are currently very lengthy and expensive, and the commercial advantages to being first with a significant new product can be very large.

2. The Problem-Solving Process

Research into the nature of problem-solving shows it to consist of trial and error, directed by some amount of insight as to the direction in which a solution might lie (Barron 1988). This general finding is supported by empirical studies of problem-solving in the specific arena of product and process development (Allen 1966, Alexander 1964, Clark and Fujimoto 1991, Iansiti 1997, Marples 1961, Smith and Eppinger 1997, von Hippel and Tyre 1995, Wheelwright and Clark 1992, Thomke 1997). Such studies show trial and error learning conducted via a process of conscious experimentation as a prominent feature. In this section we begin by discussing the general nature of trial and error problem-solving via experimentation. Then we discuss the creation of strategies for solving a given problem via a number of related experiments.

2.1. Problem-Solving via Experimentation

Experimentation using trial and error problem-solving begins with the selection or creation of one or more possible solutions. The alternatives selected may or may not include the best possible solutions - one has no way of knowing. These are then tested against an array of requirements and constraints (Dunker 1945, Marples 1961, Simon 1969). The new information provided by a trial and error experiment to an experimenter are those aspects of the outcome that he or she did not (was not able to) know or foresee or predict in advance – the "error". Test outcomes are used to revise and refine the solutions under development, and - generally - progress is made in this way towards an acceptable result.

One may view the experimental trial and error process as cycles that repeatedly "generate and test" design alternatives (Simon 1969). For example, one might conceive of, design, and build a prototype of a new, more rapidly-deploying airbag for a car ("generate alternative"); and run an experiment to evaluate its actual deployment speed ("test alternative"). If the results of a first experiment are satisfactory, one stops. However if, as is usually the case, analysis shows that the results of the initial experiment are not satisfactory, one may elect to modify one's experiment and "iterate" - try again. Modifications may involve the experimental design, the experimental

conditions, or even the nature of the desired solution. (For example, a researcher may design an experiment with the goal of identifying a new cardiovascular drug. However, experimental results obtained on a given compound might suggest a different therapeutic use, and cause researchers to change their view of an acceptable solution accordingly.)

Experimentation is often carried out using simplified versions (models) of the eventually-intended test object and/or test environment. For example, aircraft designers often conduct experiments on possible aircraft designs by testing a scale model of that design in a "wind tunnel" – an apparatus that creates high wind velocities that partially simulate the aircraft's intended operating environment. The value of using models in experimentation is twofold: to reduce investment in aspects of the real that are irrelevant for the experiment, and to "control out" some aspects of the real that would affect the experiment in order to simplify analysis of the results. Thus, models of aircraft being subjected to wind tunnel experiments generally include no internal design details such as the layout of the cabins - these are both costly to model and typically irrelevant to the outcome of wind tunnel tests, which are focused on the interaction between rapidly moving air and the model's exterior surface.

Models used in experimentation can be physical in nature, as in the example just given, or they can be represented in other forms. Computer simulation, for example, involves representing experimental objects and experimental environments in digital form, and then simulating their interaction within a computer in a type of virtual experiment (Thomke 1998). Thus, one might model an automobile and a crash barrier inside computer, perform the computations needed to simulate the crash of the model car into the model barrier, and then calculate the effects of that crash on the structure of the car via finite element analysis. One could then assess the results of this virtual experiment by viewing a visual display of the "crashed" car on a video display, and/or by looking at detailed calculations of the forces and accelerations generated during the simulated crash and the effects of these on the car's structure.

Sometimes designers will test a real experimental object in a real experimental context only after experimenting with several generations of models that isolate

different aspects of the real and/or that encompass increasing amounts of the complexity of the real. Developers of pharmaceuticals, for example, might begin by testing a candidate drug molecule against just the purified enzyme or receptor it is intended to affect, and then test it again and again against successively more complex models of the human organism (e.g., tissue extracts, tissue culture, animal models, etc.) before finally seeking to test its effect on real human patients during clinical trials.

Models do not represent reality completely (if they did, they would be the reality they are to represent). In part this is by design, and for the purposes mentioned earlier. In part the representation is incomplete because one does not know and/or cannot economically capture all the attributes of the real situation, and so could not transfer them into a model even if one wanted to. The incompleteness of a model is a source of unexpected errors when a given model being used in testing is replaced by a different model or by the real context or object for the first time (Tyre and von Hippel 1997). As an illustration, consider the airbag inflation example given earlier. If the gas used to inflate the airbag had been toxic, and the various experimental apparatus used to test the airbag had not been capable of detecting this factor, the problem would have been detected as an unexpected error only when real airbags were deployed in the real use environment.

2.2. *Parallel and Serial Strategies for Experimentation*

Researchers engaging in problem-solving via experimentation generally do not expect to solve a problem via a single experiment, and so often plan a series of experiments intended to bring them a solution to their problem in an efficient manner. The strategy they choose is in part a function of the information they have regarding the topography of the "value landscape" which they plan to explore when seeking a solution for their problem (Alchian 1950).[1]

---

[1] The concept of a "value landscape" is related to the study of evolutionary biology which regards fitness landscapes as the distribution of fitness values across a space of entities (Kauffman and Levin 1987, Wright 1932). More recently, fitness landscapes have been used in the study of organizational structure and strategy in the context of changing environments (Bruderer and Singh 1996, Tushman and O'Reilly III 1996, Levinthal 1997). In order to

A value landscape can be visualized as a flat plain with one or more hills rising from it. The total landscape represents the arena that the experimenters plan to search to identify an acceptable contains an acceptable solution to their problem. The probability of finding a solution increases as one ascends the "hills" in the landscape, and so the experimenters' goal is to devise a series of experments that will enable them to identify and explore those hills in an efficient manner. Real-world experimenters may not have much information regarding the value landscape they plan to explore when they begin their work – and may even abandon one landscape and switch to another as their work proceeds. (Explorations of the specification of "well-structured" problems have shown that problem-solvers often may pursue a solution to a problem across a range of value landscapes rather than simply seeking to search a given landscape in an efficient manner.[2]) Nonetheless, experimenters expectations regarding the topography of the value landscape(s) they have chosen are central to their construction of efficient experimental strategies.

As illustration, consider the choice between a strategy of serial experimentation versus parallel experimentation. When identification of a satisfactory solution to a problem involves more than one trial and error experiment, the information gained from a previous experiment(s) may serve as an important input to the design of the next one. Experiments which do incorporate learning derived from other experiments in a set are considered to have been conducted in series. Experiments that are conducted according

---

distinguish between biological evolution and the design and experimentation process, we instead use the term "value landscape" for the remainder of the paper. (For a good explanation of fitness and value landscapes and their respective differences, see Baldwin and Clark 1997a.)

[2] Well-structured problems have value landscapes for which one can precisely specify a process of trial-and-error that will lead to a desired solution in a "practical" amount of time (Reitman 1965, Simon 1973, Pople 1982). For example, a traveling salesman problem "of a size amenable to practical computation" is well-structured, because one can precisely specify a generator of alternative solutions and a solution testing procedure that are guaranteed to eventually identify the best solution. (A traveling salesman problem involves determining the most efficient itinerary for a salesman who must physically visit each of a given list of cities.) A real-world problem-solver facing a traveling salesman problem may solve this problem as given or may decide to modify it – thereby creating a new value landscape(s) to explore. For example, the problem-solver might modify the original problem by deciding to consider the option of contacting customers in the specified list of cities by using the Internet rather than by arranging physical visits by a single salesman.

to an established plan that is not modified as a result of the finding from other experiments are considered to have been conducted in parallel. For example, one might carry out a pre-planned "array" of experiments, analyze the results of the entire array, and then carry out one or more additional verification experiments (Montgomery 1991). The experiments in the initial array are viewed as being carried out in parallel, while those in the second round are carried out in series with respect to that initial array.

Suppose that the problem at issue is to deduce the correct combination for a combination lock. Good locks may have 10^6 or more possible combinations, of which only one is correct. They are also designed so as to give an "experimenter" (in this case, a robber) no indication as to how close he or she may be to the correct combination. That is, they are designed to display a value landscape that is absolutely flat for all combinations except the correct one, which can be visualized as rising up from the landscape like a narrow tower with vertical sides. In a value landscape with this topography, a parallel experimentation strategy would be the fastest, although not necessarily the most efficient choice (see table 1, strategies (a) and (c) and related discussion below). This is because, in this landscape configuration, each failed trial provides very little information that would be of use in a serial experimentation strategy – only the information that "the combination you just tried is not the correct one."[3]

(INSERT TABLE 1 ABOUT HERE)

In contrast, suppose that the value landscape is a hill with only a single peak and sides that extend to all edges of the landscape. (This is the shape, for example, of the

[3] Simon (1969) uses a similar example in explaining problem-solving as natural selection, noting that the example was originally supplied by W. Ross Ashby. "Suppose that the task is to open a safe whose lock has 10 dials, each with 100 possible settings, numbered from 0 to 99. How long will it take to open the safe by a blind trial-and-error search for the correct setting? Since there are 100^{10} possible settings, we may expect to examine about one half of these, on average, before finding the correct one [..]. Suppose, however, that the safe is defective, so that a click can be heard when anyone dial is turned to the correct setting. Now each dial can be adjusted

value landscape in the children's' game in which a child is guided to a particular spot via feedback from other children who say "warmer" each time a step is taken towards that spot.) In such a case a strategy of serial experimentation may be the most efficient choice, because the information gained from each step taken is so useful in guiding the direction of the next trial step that the correct solution is often found after only a few trials.

The relative efficiency of experimentation strategies can be estimated using what is known about the topography of the solution space, and what is known about the time and money costs associated with generating and testing alternatives in the solution space. Consider the following very simple search model in which the topography of the value landscape is known to consist of *n* points and to have the configuration described in the lock example discussed above – flat except for a single point representing the correct solution[4].

A parallel experimentation strategy (strategy (a) in table 1) would require all experiments and their tests to be done at the same time. Thus, one would not be able to incorporate what one has learned from one trial and apply it to the next trial. While this approach results in a very high number of experiments (*n*), it also reduces the total development time significantly as all experimental trials are done in parallel. Thus, in the case of this example, massively parallel experimentation would be the costliest but also the fastest strategy.

In contrast, a serial strategy applied to this sample problem would allow one to learn from each experimental trial and - equipped with this new knowledge - carefully select the next one. As shown in table 1 (c), a strategy even with minimal learning (i.e. not repeating a trial that has failed) can halve the total number of experiments required

---

independently and does not need to be touched again while the others are being set. The total number of settings that have to be tried is only 10 × 50, or 500."

[4] Much more sophisticated models of search have been applied to the study of the R&D and design process. See, for example, Nelson (1961), Evenson and Kislev (1976), Weitzman (1979), Roberts and Weitzman (1981), Nelson and Winter (1982), Nelson (1982), and Baldwin and Clark (1997b). The purpose of our simpler model is to help the reader in understanding the cost and time trade-offs between parallel and serial experimentation strategies.

on average, but would dramatically increase total development time relative to the purely parallel approach[5].

Of course, if there is the opportunity for greater learning from each trial, the number of trials in the series likely to be required to reach the solution (and therefore the total elapsed time) is further reduced. For example, consider a very favorable learning scenario where the *n* trials are arranged on a linear scale (e.g. *n* different pressure settings) and that after each trial, one could learn whether to move up or down on that scale. Thus one would effectively reduce the search space by 50% after each experimental cycle and rapidly progress towards an optimal solution. An experimenter would start with $^{n}/_{2}$ (the midpoint) and move to either $^{n}/_{4}$ or $^{3n}/_{4}$, depending on the outcome of the first experiment, and continue in the same fashion until the solution is found. A real-world example for such a search can be found in the practice of system problem identification: very experienced electronic technicians tend to start in the middle of a system, find the bad half, and continue to subdivide their search until the problem is found. One can easily see that the expected number of trials until success using such a serial strategy (with the kind of learning described) can be reduced to $\log_2 n$ – a dramatic reduction in cost. However, total development time would exceed that of the purely parallel strategy by the same factor (see table 1, strategy (b))[6].

Real-world experimentation strategies can be much more complex than our simple example, and will often contain a combination of serial and parallel approaches. As we will see in the discussion of drug discovery in section 4, pharmaceutical firms typically employ both, serial and parallel experimentation in the search for promising candidate drug molecules. Factors such as cost and time of generating and testing

---

[5] Assume that the set of possible experimental trials is of size *n*. After an alternative is generated, a screen tests if it is the solution (there exists only one solution in *n*). If the experimental trial results in a solution, the experimenter stops. If the experimental trial is unsuccessful, the experimenter randomly generates another alternative and continuous. The experimenter only learns which trials have failed and thus should be avoided going forward — i.e. the experimental learning is minimal.

[6] A related strategy is the very famous Newton's method that is widely used in numerical analysis and was first introduced in Newton's Principia Mathematica to solve a cubic polynomial. The method's iterative and sequential search, known for its rapid convergence, is

alternatives, and knowledge of and the topography of the solution space is affecting the degree to which parallelism is employed in the drug discovery process.

3. Advances in Experimentation and Problem-Solving – and Implications for Firms

The methods and tools available to help solve many types of problems are rapidly changing and improving. These advances are affecting all of the elements of the experimentation process that have been described in the previous section. That is, they are rapidly reducing the cost and time involved in designing and executing and analyzing many types of experiments, and are also affecting the type of experimentation strategies that may be most effective for an experimenter. In this section we first illustrate the rapid advances being made in experimental methods by noting the rapid evolution[7] of experimentation via computer simulation. Next we will discuss the implications that such advances can have for the competitive position of firms.

3.1. Advances in Methods for Experimentation and Problem-Solving

There is no general index that documents the rate of advance in problem-solving methods and tools. However, those with a professional interest in these matters generally judge that the rate of change today is very rapid. Advances in some fields, such as computer simulation, are applicable to a wide variety of subject matter. Others such as the scanning tunneling electron microscope are germane to only a narrow range of applications - although the range of application seen for a given technique often broadens significantly over time (Rosenberg 1982). The reader may find a brief overview of the rapid evolution of computer simulation techniques to be a useful way to gain a feeling for what we mean by "rapid advances in methods and tools" in the case of a generally-applicable tool. Later, in section four, we will provide a detailed description

---

guided by knowledge of the underlying function and its gradient to quickly find an accurate estimate of the numerical value being sought (Gerald and Wheatley 1984).

[7] We regard the evolution of methods as a decentralized, adaptive process that can be characterized by interplay between technology or method users *and* developers. Because both derive economic value from advances in areas such as simulation, they tend to reinforce each other and thus accelerate the overall evolution of such novel technologies or methods.

of the nature of and impact of an advance with a narrower range of application – combinatorial chemistry.

As was noted earlier, experimentation via computer simulation involves representing experimental objects and experimental environments in digital form, rather than in the form of physical objects tested within physical environments. Then, their interaction is within a computer in a type of virtual experiment. The advantages of substituting virtual experimentation via computer for experimentation with real physical objects can be very significant. For example, studying automobile structures via real car crashes clearly is quite expensive and time-consuming – a crash prototype can cost in excess of one million dollars and may take up a year to build and test. In contrast, once the proper digital models have been created, a virtual car crash can be run again and again within a computer under varying conditions at very little additional cost per run. Further, consider that a real car crash experiment happens very quickly – so quickly that the experimenter's ability to observe details is typically impaired, even given high-speed cameras and well-instrumented cars and crash dummies. In contrast, one can instruct a computer to enact a virtual car crash as slowly as one likes, and can zoom in on any structural element of the car (or minute section of a structural element) that is of interest and observe the forces acting on it and its response to those forces "during" the crash. Thus, computer simulation may not only decrease the cost and time of an experimental cycle but can also increase the depth and quality of analysis, leading to improved learning and ultimately products of higher quality (Thomke 98).

The steady (and really quite spectacular) improvement in the capabilities of digital computers over the past few decades has made it possible and desirable to carry out more and more experiments via computer simulation, rather than via physical experimentation. Computer simulation is today being used as a substitute for or supplement to physical experimentation in fields ranging from the design of drugs (e.g., rational drug design) to the design of mechanical products (e.g., finite element analysis), to the design of electronic products (e.g., simulation of digital circuitry), and to the analysis of financial positions (e.g., simulation of novel financial instruments). The ability to usefully substitute a simulation for a "real" experiment requires, of course,

more than the development of advanced computer equipment. It also requires the development of simulation models that are accurate from the point of view of a given experimental purpose. Often, a simulation model will not be fully accurate in ways that later turn out to matter. When this is recognized, virtual and physical experiments may be conducted in some combination in order to combat this source of error. (For example, auto designers will supplement data gathered from virtual car crash experiments with data from real crash experiments using real cars, in order to assure themselves that the results of the virtual experiments also hold in the real world.)

At the same time, of course, methods for conducting physical experiments are also advancing. For example, significant advances are being made in reducing the costs and time of building the various types of prototypes. Complex three-dimensional objects used to require days or weeks of work in a machine shop to fabricate. Many such shapes can now be made rapidly - in very few hours - by using computer-controlled machining equipment and/or equipment for creating objects via "three dimensional printing" (Sachs 1992). Similarly, physical prototypes of complex electrical circuitry - custom integrated circuits - used to take months to create via "full custom" methods, and weeks to create via "Application-Specific Integrated Circuits" (ASIC) technology. Now, designers can create customized circuits in minutes at their desks or lab benches using so-called "Field Programmable Gate Arrays" (FPGAs) (Villasenor and Mangione-Smith 1997).

3.2. Impact of Changes in Experimental Methods on Firm Competitiveness

The adoption of more effective experimental methods for problem-solving and the development of new products and services, such as those just described, can lead to significant competitive advantages for adopting firms relative to rivals *if* novel techniques that offer such advantages are not rapidly picked up by rivals as well. Or, as Barney (1986) and Wernerfelt (1984) put it with respect to core competencies: a core competence can be a source of long-term competitive advantage for a firm if it is difficult or impossible to buy or sell in the available factor markets, and if it is difficult to replicate.

We argue that the new and more effective experimental methods and techniques that are rapidly emerging are indeed often difficult to buy and sell, and difficult to replicate as well, and therefore that they can and do serve as a significant source of long-term competitive advantage for innovators and early adopters. The reason that this is so is that new methods require (1) the transfer of significant amounts of new information to the adopting firm, including new skills, and (2) some reorganization of a firm's R&D activities as well.

The requirement that new information must be transferred to a firm adopting a new experimental technique is in itself a barrier to adoption in many instances, because information is often costly to transfer to a new site in a form usable by a given information seeker. Transfer costs are affected by attributes of the information itself (e.g. how the information is encoded), and also by attributes of and choices made by information seekers and information providers (Arora and Gambardella 1994, Cohen and Levinthal 1990, Griliches 1957, Mansfield 1968, Nelson 1982, Pavitt 1987, Rosenberg 1982, Teece 1977, von Hippel 1994).

Thus, consider that only some of the information associated with the ability to execute new experimental methods may be embodied in equipment that can be purchased and installed by an adopting firm – a relatively easy form of transfer. For example, a firm can buy computers and computer programs that can be used to do experiments via computer simulation. But new equipment and new software provide only a portion of the information a firm needs to actually become competent at performing a new experimental method. Typically, new skills and expertise are also needed and, as Polanyi (1958) has pointed out, skill and expertise are often encoded within an expert's mind as tacit information that is difficult to transfer to another. For example, in a study of biology lab practicing an experimental method known as cell fusion, Barley and Bechky reported that "…experienced research support specialists and technicians [carrying out the cell fusion work] made use of signs that could not be found in textbooks, and that were difficult to define except ostensively. Partially for this reason, practices successful in one lab often failed in another unless technicians from the first trained technicians from the second." (Barley and Bechky 1994 p. 98-9)

Adopting novel experimental methods may also require considerable change in the organizational arrangements prevailing in the adopting firm. As Morison (1966) and Schˆn (1967) have pointed out, organizations are often built up around and adapted to existing technologies. When this is so, changes in technologies may require changes to organizational structures and routines. As an illustration, suppose that a firm wishes to replace some physical experimentation methods being carried out in its labs with computer simulation methods. To do this, it must typically hire new kinds of people and also reorganize the relationships between the various specialists who jointly carry out the experiments. In its existing organizational arrangements designed for physical experimentation, for example, the firm might have routines in place that enable researchers to work with design engineers and modelmakers to design and build the experiments that they wished to run. Next, the procedures might dictate that the completed experimental apparatus be transferred to experts at specialized test facilities who would actually run the experiments, collect the resulting data, and then supply that data to the researcher for analysis. In contrast, experimentation via computer simulation would require quite different organizational routines. In some cases these would enable the researcher to do the entire design, build, test and analyze experimental cycle in his or her own lab. In other cases, they might facilitate collaborative arrangements between the researcher and various types of experts not previously employed by the firm who specialize in different aspects of computer simulation.

With respect to the difficulty of achieving such organizational change, we note that Holmstrom and Tirole (1991) have argued that organizational arrangements cannot serve as sources of enduring competitive advantage because they can be easily replicated. However, much of the literature on organizational change suggests otherwise (see, for example, Milgrom and Roberts 1990, Henderson and Cockburn 1994). Thus, Henderson and Cockburn (1994), in a study of cardiovascular drug discovery, report that organizational capabilities found associated with improved productivity at

this type of research task are in fact often very difficult to transfer from firm to firm[8]. Further, they note that such arrangements can have an important impact on research productivity. In their study, about 30% of the observed variation in the "productivity" of firms in drug discovery (number of drugs discovered per R&D dollar invested) was due to unique organizational capabilities (represented by a variety of measures such as to the degree to which the firm actively manages the integration of knowledge across disciplinary and firm boundaries).

4. Field Study: The Impact of New Drug Discovery Methods on Pharmaceutical Drug Development

To this point we have described the general nature of problem-solving via experimentation in R&D, have observed that methods for accomplishing this task are evolving rapidly, and have argued that competence at problem-solving via experimentation can be important with respect to the competitiveness of firms that perform R&D. In this section we develop these points further via a case study of a recent improvement in experimental methods used in the drug discovery process – "combinatorial chemistry". We begin by describing the serious drug development problem currently facing pharmaceutical firms. Next, we describe the drug discovery process. Then we describe combinatorial chemistry, and finally we describe a research project that clearly illustrates the impact that this new method can have on the drug discovery process – and with it, upon the competitiveness of pharmaceutical firms.

---

[8] Henderson (1994: 624-6) illustrates difficulties associated with replicating organizational capabilities associated with better performance at drug discovery by presenting examples experienced by firms in their sample. Thus, there was a period when leading-edge drug discovery processes were shifting from simple mass screening of compounds for possible medicinal effects to a more precise form of research based on an understanding of a drug's mechanism of action. This change was being driven by the academic research community. Drug firm "Alpha," which had long-term ties to the academic community and which employed leading-edge researchers who were accepted as peers in that community, had no difficulty in quickly adopting the new approach to drug discovery. In contrast, firm Beta, which had not had a practice of employing scientists known to and respected by the academic community, found it very difficult to make the change. For example, they found it difficult to hire "better" people from academia who were experts in the new approach, because they did not have a reputation as a leading-edge place to work.

4.1. The Product Development Problem Facing Pharmaceutical Firms

If improvements in problem-solving methods are important to any firm, they should certainly be important to firms in the pharmaceutical industry. On the one hand, pharmaceutical firms face many potentially profitable opportunities to create new drugs to cure or ameliorate diseases ranging from cancer to heart disease, particularly if firms manage to receive patent protection and reach the market before their competitors do. Markets for new drugs typically involve $50-400 million in annual sales, and can reach sometimes into the billions as in the case of Zantec, a stomach acid inhibitor drug for ulcer treatment. On the other hand, the drug development process is currently one of the most time-consuming and costliest product development processes in any industry.

A widely-cited study of pharmaceutical drugs developed between 1972 and 1987 found that the expected capitalized development cost per marketed drug was on average 230.8 million (1987) dollars (DiMassi *et. al.* 1991), with total development times well above ten years. Various other studies have shown a trend that has caused much concern in the pharmaceutical industry: the cost and time of new drug development has increased significantly over the last thirty years (e.g. DiMassi *et. al.* 1994). Besides the impact of lower R&D productivity on firm cost and profitability, longer development times have also raised important public policy concerns. As the industry remains the dominant provider of life-saving and life-prolonging medicines, it is in the public interest to have promising new drugs available to patients as quickly as possible (Savello 1996).

The complete drug development and approval process involves three phases. It begins with a "preclinical" research phase devoted to the discovery and optimization of one or a few "lead" chemical compounds that appear to hold sufficient promise as drugs to merit investment in clinical testing. The second phase, clinical development, consists typically of three clinical phases to determine and document the safety and efficacy of the proposed drugs. The final phase involves regulatory New Drug Approval (NDA) review processes of the clinical trial outcome. The average cost and duration of preclinical and clinical development for drugs developed between 1972 and 1987 is provided in table 2.

(INSERT TABLE 2 ABOUT HERE)

4.2. The Drug Discovery Process

Drugs achieve their effect by binding with very specific molecular receptors or enzymes or biologically important molecules that are present in the human body or on/in disease-causing agents such as bacteria, fungi and viruses. The goal of drug discovery or drug design is therefore to discover or create a molecule that will in fact bind to a particular, say, receptor with a required degree of tenacity (binding affinity), and that will at the same time not bind to other receptors that may be structurally similar but have different functions.

The drug discovery process can involve either or a combination of two basic approaches:

(1) One can start with little or no knowledge about the structure of a disease target (receptor, enzyme, molecule) associated with a particular disease, and simply try out many candidate molecules until one finds one that happens to bind properly with the target receptor.

(2) One can strive to determine the structure of the relevant receptor with biophysical methods, and then attempt to design or select a molecule that will bind to it.

Until the 1970's methods of drug discovery necessarily relied on the first of these two approaches because the technical ability to determine the molecular structure of a protein receptor did not yet exist. Researchers at early pharmaceutical firms (often subsidiaries of chemical manufacturing firms) implemented this approach by setting up a systematic trial and error drug discovery system known as the "mass screening" system, which is still used today.

The mass screening system begins with the selection or design of a "screen" - e.g. a disease-causing bacterium or an isolated receptor that is known to be associated with the disease under study. "Masses" of chemical compounds are then applied to this screen (one at a time), with the goal of identifying compounds that cause the screen to

display a desired effect (e.g., killing of the disease-causing bacterium; evidence of binding to the receptor).

Traditionally, there have been two different sources of input materials to the mass screening process. The first source is proprietary archival libraries of known chemical compounds that have been collected by chemical and pharmaceutical firms over the years. A given major firm might have an archival library of perhaps half a million known compounds. The second source is extracts of plants, microorganisms and animals, each of which may contain perhaps up to 100,000 unknown chemical compounds.

Mass screening proceeds differently depending on which type of input is used. In the case of archival libraries, the known compounds are tested against the disease target screen one by one, and the effect of each on the screen is observed. In the case of natural extracts, the entire extract is tested against the screen. If a desired effect is observed, the compound responsible for that effect must then be isolated via a complex series of fractionations and retestings.

As an illustration for the natural compound process, consider the development of antibiotics based on "magainins." When researchers noticed that frogs living in bacteria-contaminated water did not appear to get skin infections, they suspected that a new and useful antibiotic compound in a frog's skin might be involved (Zasloff 1987). To identify it, they began by grinding up frog skin and subjecting the whole mixture – consisting of literally hundreds of thousands of different compounds – to mass screening tests for antibiotic activity. When these tests did indicate antibiotic activity, they next had to identify which compound(s) in the complex mixture was (were) the source of that activity. This was done by biochemical separation of the compounds found in frog skin into fractions, followed by a test of each fraction for the presence of antibiotic activity. The active fraction was then subject to further cycles of fractionation and test until finally the active compound was isolated.

When an active compound is finally identified via mass screening, it will generally not meet all of the criteria required to make it a "lead" candidate for a new drug. For example, it may display the needed medical effect very powerfully, but at the

same time may display unacceptable side effects such as toxicity, mutagenic effects in animals, or may not become available in the bloodstream after ingestion or injection Therefore, the lead optimization process in the drug discovery process is to create and test a number of variations ("analogs") of the originally-identified molecule, in order to find one or more that appears to have all the attributes needed for a successful new drug. One lead compound is then advanced into the clinical development phase where its effects are tested on humans. Experimentation with analogs is carried out in series, with some elements of parallelism: a few molecules are created and tested during each round with the objective of learning as much as possible between rounds - a strategy discussed in section 2.

At this point we should note that the traditional process used to create analogs to a proposed drug in order to create a lead drug compound is typically a very costly and time-consuming matter. In order to create analogs to the original compound, medicinal chemists (specialized organic chemists employed by pharmaceutical firms) maintain the basic structure of that compound, but add, exchange or remove chemical groups from it. On average it takes 7 to 10 days and approximately $7,500 to synthesize one such analog (Longman 1994). According to the statistics of the Centre for Medicines Research the average American pharmaceutical company synthesizes approximately 6,100 chemical compounds for each successful drug that makes it to the market place (Halliday, Walker and Lumley 1992b). This amounts to an average of $46 million for analoging alone, with a total time requirement of about 170 person-years.

The reason it is necessary to develop so many potential solutions to the receptor problem is because many drugs must be precisely tailored to discriminate sharply between very similar receptors. For example, researchers working to develop a drug for Alzheimer's disease are targeting a particular muscarinic receptor located in the brain. However, five subtypes to this muscarinic receptor are known to exist in the gut and elsewhere, and the desired drug must *not* affect these. Compounds displaying the needed selectivity can be very difficult to find without extensive analoging.

4.3. Rational Drug Design and Combinatorial Chemistry

In this century, the knowledge in chemistry, biology and the molecular basis of disease has increased exponentially. In the beginning of the 80s advanced methods of protein structure determination and computer supported molecular modeling became the focus of the pharmaceutical industry. This new technology was thought by many in the industry to be sufficiently advanced to allow the creation of a ***"rational drug design"*** methodology as an alternative to the traditional approach to drug discovery.

Pharmaceutical researchers using the rational drug design approach would use very advanced scientific methods such as x-ray crystallography and/or nuclear-magnetic resonance (NMR) spectroscopy to determine the three dimensional shape of a receptor or an enzyme that they wish to influence with a drug. They would then enter the structure of this receptor into a computer software package containing information on the configuration and strengths of the chemical bonds that can form between atoms. This software would then allow them to use simulation to design drug molecules that bind properly to the target receptor. Real molecules would then be created by chemists in the laboratory as specified by the computer modeling exercise, and these would be tested for the desired pharmaceutical effect. Thus rational drug design is an example of a strategy that tries to maximize the amount of learning between trials and thereby achieve a total reduction in the number of experimental trials (as explained in section 2.2).

However, the "rational" approach to drug design has proven to be problematic for two reasons. First, the molecular modeling of a drug requires very accurate data on the structure of the target receptor, and the required degree of accuracy is often very difficult for researchers to attain using present-day methods. Second, it has been found that the shape of a target receptor can change dramatically when a drug is inserted into the receptor's "binding pocket" (see figure 1). The effect of such "induced fit" shape changes is that a drug that has been designed to fit a receptor's empty binding pocket may in fact not fit at all. Induced fits between receptor and drug are too complex to be modeled by present computer simulation tools. As a consequence, rational drug design has proven not to be a full replacement for traditional drug design methods. Instead, computer-based molecular modeling exercises prescribed by the rational drug design

procedure are still followed up by medicinal chemists who create and test analogs to the rationally-designed compound, just as was done in traditional drug development.

(INSERT FIGURE 1 ABOUT HERE)

In the last few years, a new method called ***"combinatorial chemistry"*** has emerged very rapidly (figure 2), primarily due to its impact on the underlying experimentation economies (Plunkert and Ellman 1997). Combinatorial chemistry makes the synthesis of proposed drug compounds and analogs radically faster and cheaper (the basic principle of combinatorial chemistry and the underlying process technologies are described in the appendix). For example, cost reductions from about $7,500 per compound (traditional medicinal chemistry) to perhaps $1 to $10 per compound have been reported, with reductions in preparation time of comparable magnitude. Where a skilled medicinal chemist requires 7 to 10 days to create a single analog using traditional methods, a chemist can now use automated equipment and combinatorial chemistry techniques to create thousands of analog compounds - each precisely identified by an attached chemical "tag" - in a matter of days (Franke 1995). Thus, as we will see in the following section, the dramatically different economies of combinatorial chemistry are inducing drug developers to shift to mixed experimentation strategies with a strong emphasis on parallelism, while at the same time reducing total development time dramatically (similar to strategy (a) in table 1).

(INSERT FIGURE 2 ABOUT HERE)

The impact of this new capability on the drug discovery promises to be very significant. The amount of information that must be acquired about the structure of a receptor via computer modeling, crystallographic studies, etc. can be greatly reduced with the application of combinatorial chemistry methods, because these can be used to create literally hundreds or thousands of compounds that might fit the receptor. The most promising (that is, the ones with the best desired influence on the target receptor)

can then be identified via a mass-screening process. Next, one can create a new library of hundreds or thousands of analogs for each of these "round one winners" within a few days or weeks. One can then repeat this screening, selection and analoging process until one gets compounds displaying an excellent level of binding to only the target receptor. Because one has been able to create and test so many analogs, one can generate a "lead" compound more quickly and more cheaply. And, perhaps even more important, one can identify a *better* lead compound to carry forward into the very expensive clinical development phase.

4.4. Field Study: Drug Discovery at Pharmacopeia via Combinatorial Chemistry Methods

We next illustrate the impact combinatorial methods can have on the economics and experimentation strategies of the drug discovery process by describing a research project carried out by Pharmacopeia Inc. of Princeton, New Jersey. Pharmacopeia is a well-known leader in the novel field of combinatorial chemistry using solid support libraries (see appendix). We compare the costs and outcomes that were achieved by using combinatorial chemistry in this case versus estimated costs and outcomes that would have been achieved by using more traditional methods. We find that traditional methods would have been dramatically slower and costlier in this case – and would probably only have produced "lead" drug candidates with little chance of clinical success. The high cost and time required to create and test compounds using traditional methods would have severely limited the number of compounds considered - and thus reduced the related search space - and would have focused the search to a region with the least promising molecules.

The drug discovery project we report upon deals with the identification of lead drug candidates to be used in the treatment of an important eye disease ("glaucoma") that affects 1 in 100 adults. Glaucoma is a wide-spread human disease responsible for impaired vision and eventual blindness. To document this project, we interviewed Pharmacopeia's leading scientists and executives. These interviewees provided us with information about Pharmacopeia's combinatorial chemistry technology in general and detailed information on a drug development case in particular. Personal interviews

were followed with a detailed questionnaire that provided us with data on the efficiency of the drug discovery process used in this project. (For detailed scientific description of the underlying chemistry used, see Burbaum *et. al.* 1995.)

Glaucoma is caused by a build-up of pressure within the human eye which in turn causes damage to optical nerve cells. Scientific research has shown that excessive pressure can be treated with the aid of what is known as the "diuretic" effect (i.e. a reduction of liquid causes a decrease in pressure). It is also known that a certain group of drugs – known as carbonic anhydrase drugs – can precisely cause this diuretic effect, leading to stabilized pressure within the eye and long-term preservation of vision. The glaucoma project's objective was to find sulfonamide compounds that "lock" into and inhibit the function of the human carbonic anhydrase enzyme (hCAI) which regulates the production of liquid in the human eye, thus leading to a reduction of both, pressure and damage to optical nerve cells. As it is usually the case with pharmaceutical drugs, a promising lead compound had to discriminate against enzymes that have very similar structures in order to avoid unacceptable side effects.

In Pharmacopeia's case, a promising lead compound that might guide the way to an eventual new drug had to interact with hCAI, but discriminate against the bovine isozyme (bCAII) – two very similar receptors. (The bovine isozyme acts as a starting model for hCAII which is the human isozyme; once a drug that discriminates against bCAII is found, it acts as an excellent lead for discrimination against the human isozyme hCAII.) The sulfonamide compounds identified served as leads for additional phases in the drug development process. Identifying suitable lead compounds was very difficult because, as was learned later, only three compounds out of thousands tested eventually displayed the searched-for selectivity. (See figure 3; only compounds close to the abscissa and ordinate can discriminate against the respective enzymes.) Failure to discriminate against enzymes other than hCAI, however, was known to cause serious side-effects such as difficulties in breathing, convulsion, muscle cramps, and trembling.

(INSERT FIGURE 3 ABOUT HERE)

R&D Efficiency of Combinatorial Chemistry versus Traditional Drug Discovery Methods

Lead compounds for the glaucoma project were identified using the combinatorial chemistry methods we described earlier. Data on development time, cost and experimentation strategies was collected for the actual mode used (combinatorial chemistry) and an estimated case using traditional medicinal chemistry was constructed (table 3).

(INSERT TABLE 3 ABOUT HERE)

The estimated case was based on considerable experience with projects that were comparable in complexity and degree of difficulty but were developed using the traditional medicinal approach. A "rational" drug design approach would have focused on reducing the number of compounds tested in the traditional approach by maximizing the learning between successive rounds – but at considerable additional cost and time. Thus, it is unclear whether current "rational" drug design would have improved the efficiency and output of the traditional approach at all. The data from table 3 shows that the combinatorial approach was not only more cost effective but also led to a dramatically lower discovery lead time, allowing the firm to move to the clinical phase much earlier. It also identified lead compounds that showed a high degree of selectivity and thus much promise of success for the next development phases. In fact, interviewees strongly felt that with the cost and time required using traditional medicinal methods, it would have been (1) unlikely that the project had been pursued; and if pursued, (2) nearly impossible to identify a promising lead compound with the required selectivity.

Of course, combinatorial chemistry does not offer the same advantage for all projects, and is currently not applicable at all to some kinds of molecules. Thus, the method is not now very effective for the kinds of complex molecules often dealt with in studies of natural compounds. However, combinatorial chemistry's area of applicability is rapidly expanding and many companies are working on the conversion

of classical organic chemical reactions to combinatorial systems. Expert interviewees contacted during our case study estimate that the the advantage of offered by combinatorial methods over traditional experimental methods for projects where combinatorial chemistry is applicable today ranges from a 10% to an 80% reduction in the cost and time devoted to lead optimization - *and*, as was noted earlier, the development of better quality lead compounds than is customarily accomplished by traditional medicinal chemistry techniques.

5. Conclusion

In this paper, we have argued that the economics of problem-solving and the related R&D efficiency are being radically affected by the use of new and greatly improved versions of methods such as computer simulation and combinatorial chemistry. We explained how the introduction of novel methods could affect both the experimentation strategies adopted by firms (e.g., serial vs parallel experimentation) and the efficiency with which those strategies can be executed. Via a case study of the impact of combinatorial chemistry techniques on pharmaceutical drug discovery, we then illustrated the dramatic economic changes that can result from the adoption of novel experimental methods.

We also noted that novel experimental methods can importantly affect the relative competitive position of firms if techniques that offer such advantages are difficult or impossible to buy or sell in the available factor markets, and difficult to replicate as well. Many novel methods require novel skills and/or organizational arrangements to implement, and are likely to meet these criteria. (Certainly, as we explained, it is likely that the methods discussed in our case, combinatorial chemistry and rational drug design, meet the criteria. Equipment required by both are available on the market – but both also require novel skills and organizational arrangements that are not easily acquired by firms seeking to adopt them.)

In sum, we propose that strategies and modes of experimentation can be an important factor in the effectiveness of firms' innovation processes and their relative competitive positions. We therefore propose that further studies on this topic may be of

interest to both innovation researchers and innovation practitioners. For example, it would be useful to explore whether and how differences in the relative effectiveness of firm innovation processes can be traced to differences in the experimental methods employed, and the skill with which those methods are used. It would also be useful to explore the attributes of experimental methods that convey the greatest competitive advantage to firms using them. For example, it is likely that the methods that are the hardest to transfer to new users will be the ones that offer the greatest competitive advantage to method *users* – while method *sellers* are likely to appropriate the most benefit from methods that are easily transferred.

6. Acknowledgments

We would like to thank John Baldwin, John Chabala, Joseph Mollica and Nolan Sigal from Pharmacopeia, Michael Pavia from Millennium Pharmaceuticals, our colleagues Carliss Baldwin, David Bell and Steven Wheelwright and two anonymous reviewers for their help and comments on earlier drafts of this paper. The financial support of the Harvard Business School Division of Research and the Research on Science & Technology Program at the National Science Foundation are gratefully acknowledged.

References

Alchian, A., 1950, Uncertainty, evolution and economic theory, Journal of Political Economy, 58 (3).

Alexander, C., 1964, Notes on the Synthesis of Form (Harvard University Press, Cambridge, Massachusetts).

Allen, T. J., 1966, Studies of the problem-solving process in engineering design, IEEE Transactions on Engineering Management, EM-13, no.2: 72-83.

Arora, A. and A. Gambardella, The changing technology of technological change: general and abstract knowledge and the division of innovative labor, Research Policy 23, no 55, September, 1994, 523-32.

Baldwin, C. and K. Clark, 1997a, Design options and design evolution, Harvard Business School Working Paper #97-038.

Baldwin, C. and K. Clark, 1997b, The value of modularity: splitting and substitution, Harvard Business School Working Paper #97-039.

Barley, S. and B. Bechky 1994, "In the Backrooms of Science: The Work of Technicians in Science Labs" Work and Occupations, Vol 21 No. 1, February: 85-126

Barney, J., 1986, Strategy factor market: expectation, luck and business strategy. Management Science, 32(10): 1231-1241.

Barron, J., 1988, Thinking and Deciding, (Cambridge University Press).

Bruderer, E. and J. Singh, 1996, Organizational evolution, learning, and selection: a genetic-algorithm-based model, Academy of Management Journal, vol. 39: 1322-1349.

Burbaum, J., et al., 1995, A paradigm for drug discovery employing encoded combinatorial chemistry, Proceeding of the National Academy of Sciences, vol. 92 (June 1995): 6027-6031.

Carter, A., 1995, Measuring the performance of the knowledge-based economy, Working Paper (Brandeis University).

Clark, K. and T. Fujimoto, 1991, Product Development Performance, (Harvard Business School Press, Boston).

Cohen, W. and D. Levinthal, 1990, Absorptive capacity: A new perspective on learning and innovation, Administrative Science Quarterly, 35, 1: 128-152.

DiMassi, J., R. Hansen, H. Grabowsky and L. Lasagna, 1991, Cost of innovation in the pharmaceutical industry, Journal of Health Economics 10: 107-142.

DiMassi, J., M. Seibring and L. Lasagna, 1994, New drug development in the United States from 1963-1994, Clinical Pharmacology and Therapeutics 55: 609-622.

Duncker, K., 1945, On problem solving, Translated by Lynne S. Lees, Psychological Monographs 58, no.5.

Evenson, R. and Y. Kislev, 1976, A stochastic model of applied research, Journal of Political Economy, vol. 84 no. 2.

Fodor, S., J. Read, M. Pirrung, L. Stryer, A. Lu, D. and A. Solas, 1991, Light directed, spatially-addressable, parallel chemical synthesis, Science 251: 767-773.

Franke, R., 1995, The impact of combinatorial chemistry on the pharmaceutical drug discovery process, Unpublished Masters Thesis, MIT Sloan School of Management, Cambridge, MA.

Gerald, C. and P. Wheatley, 1984, Applied Numerical Analysis (Addision-Wesley, Reading, Massachusettts).

Griliches, Z., 1957, Hybrid corn: An exploration in the economics of technical change, Econometrica, 25, 4: 501-522.

Henderson, R., 1994, The evolution of integrative capability: Innovation in cardiovascular drug discovery, Industrial and Corporate Change, vol. 3 no. 3.

Henderson, R. and I. Cockburn, 1994, Measuring competence: Exploring firm effects in pharmaceutical research, Strategic Management Journal, vol. 15: 63-84..

Halliday, R., S. Walker and C. Lumley, 1992b, Pharmaceutical Medicine 2: 139-154.

Holmstrom, B. and J. Tirole, The theory of the firm," in R. Schmalensee and R. Willig, Eds, Handbook of Industrial Organization, (North Holland, Amsterdam).

Iansiti, M., 1997, Technology Integration: Making Critical Choices in a Turbulent World (Harvard Business School Press).

Kauffman, S. and S. Levin, 1987, Towards a general theory of adaptive walks on rugged landscapes, Journal of Theoretical Biology 128, 11-45.

Levinthal, D., 1997, Adaptation on rugged landscapes, Management Science, vol. 43, no. 7.

Longman, R., 1993, "The promise of combinatorial chemistry." In Windhover Information, Inc., Pharmaceutical Strategic Alliances, 5th Edition, January 1993-June 1994.

Mansfield, E., 1968, Industrial research and technological innovation: An econometric analysis, (W.W. Norton, New York).

Marples, D., 1961, The decisions of engineering design, IRE Transactions on Engineering Management, June: 55-71.

Milgrom, P. and J. Roberts, 1990, The economics of modern manufacturing: technology, strategy and organization, American Economic Review, pp. 511-528.

Montgomery, D., 1991, Design and Analysis of Experiments, (Wiley, New York).

Morison, E., 1966, Men, Machines, and Modern Times, (MIT Press, Cambridge).

Nelson, R., 1982, The role of knowledge in R&D efficiency, The Quarterly Journal of Economics (August 1982).

Nelson, R., 1961, Uncertainty, learning, and the economics of parallel research and development efforts, Review of Economics and Statistics 43: 351-364.

Nelson, R. and S. Winter, 1982, An Evolutionary Theory of Economic Change, (The Belknap Press of Harvard University Press, Cambridge, Massachusetts).

Pavitt, K., 1987, The objectives of technology policy, Science and Public Policy, 14, 4: 182-188.

Plunkert, M., and J. Ellman, 1997, Combinatorial chemistry and new drugs, Scientific American (April 1997).

Polanyi, M., 1958, Personal Knowledge: Towards a Post-Critical Philosophy, (University of Chicago Press, Chicago, IL).

Pople, H., 1982, Heuristic methods for imposing structure on ill-structured problems: the structuring of medical diagnostics, chapter 5 in Peter Szolovits, ed., Artificial Intelligence in Medicine, (Westview Press, Boulder, Colorado).

Reitman, W.R. (1965) Cognition and Thought, (Wiley, New York).

Roberts, K. and M. Weitzman, 1981, Funding criteria for research, development, and exploration projects, Econometrica, vol. 49 no. 5, September, 1261-1288.

Rosenberg, N., 1982, Inside the Black Box: Technology and Economics, (Cambridge University Press, New York).

Sachs, E. *et al*, 1992, CAD-Casting: Direct fabrication of ceramic shells and cores by three dimensional printing, Manufacturing Review vol.5, no.2, June: 117-126.

Savello, D., 1996, Speeding new medicines to patients, DDT vol.1, no.2: 45.

Schˆn, D., 1967, Technology and Change, (Delacorte Press, New York).

Simon, H. A., 1969, The Sciences of the Artificial, Second Edition (MIT Press, Cambridge, MA).

Simon, H. A., 1973, The structure of ill-structured problems. Artificial Intelligence 4: 181-201.

Smith, R., and S. Eppinger, 1997, A prediction model of sequential iteration in engineering design, Management Science, vol. 43 no. 8.

Stryer, Lubert, 1995, Biochemistry, Fourth Edition, (Freeman and Company, New York).

Teece, D., 1977, "Technology transfer by multinational firms: the resource cost of transferring technological how-how, Economic J., 87, 346: 242-261.

Thomke, S., 1997, Managing experimentation in the design of new products, Working Paper #96-037, in press Management Science.

Thomke, S., 1998, Simulation, learning and R&D performance: evidence from automotive development, Working Paper #97-089, Harvard Business School, in press Research Policy.

Tushman, M. and C. O'Reilly III, 1996, The ambidextrous organization: managing evolutionary and revolutionary change, California Management Review, vol. 38, no.4.

Tyre, M. J. and Eric von Hippel, 1997, The situated nature of adaptive learning in organizations, Organization Science.

Villasenor, J. and W. Mangione-Smith, 1997, Configurable computing, Scientific American, June, 66-71.

von Hippel, E., 1994, Sticky information and the locus of problem-solving: implications for innovations, Management Science 40 no. 4, April, 429-439.

von Hippel, E., and M. J. Tyre, 1995, How 'learning by doing' is done: problem identification in novel process equipment, Research Policy, January.

Weitzman, M., 1979, Optimal search for the best alternative, Econometrica, vol. 47 no. 3, May, 641-654.

Wernerfelt, B., 1984, A resource-based view of the firm, The Strategic Management Journal, vol. 5, 171-180.

Wheelwright, S., and K. Clark, 1992, Revolutionizing Product Development, (The Free Press, New York).

Wright, S., 1932, The roles of mutation, inbreeding, crossbreeding and selection in evolution, Proceedings of the 11th International Congress of Genetics, 1; 356-366.

Zasloff, M., 1987, Proceedings of the National Academy of Sciences USA, 84: 5443-5453.

Appendix

Background on Combinatorial Chemistry

Combinatorial chemistry is a very novel experimentation methodology and has evolved over the last decade. It consists of several new chemical synthesis strategies for the efficient generation of a large number of chemical compounds. This large number of chemical compounds, also called compound libraries, is subsequently used in pharmaceutical drug screening projects. The term "combinatorial" originates from chemical synthesis methods applied to most of these libraries.

The following is a brief description of the main process technologies that enable a large number of parallel experiments to be generated quickly and at low cost:

- ***Biochip libraries***: Photolithographic synthesis methods are used for the creation of compound libraries on the surface of a silicon chip (Fodor *et. al.* 1991). Up to 10,000 individual compounds can be synthesized on a silicon chip with little more than one square centimeter surface area.

- ***Solid support libraries***: Compounds are synthesized on the surface of polymer beads. This method allows the chemist to attach a certain type of molecule to glass beads and split the pool of glass beads to continue with different synthesis steps (see example below).

- ***Solution libraries***: Mixtures of compounds react chemically in a carefully designed system to form solution libraries with tens to thousands of different compounds within a few hours.

- ***Rapid parallel synthesis libraries***: Robotic equipment is custom-tailored to dispense chemicals into individual reaction chambers, carry out many individual chemical reactions in parallel, and extract and purify the reaction products automatically. Although this process is significantly slower than the other three technologies, it results in individually purified compounds at quantities sufficiently large for elaborate second round screening. (The other methods require chemical resynthesis which may cause a small but significant time delay.)

To illustrate how such a combinatorial chemistry works, consider the process of building solid support libraries. In the first synthesis phase, polymer beads are reacted in three different reaction vessels with chemical A in vessel 1, chemical B in vessel 2 and chemical C in vessel 3. After the reactions are completed, all the polymer beads are pooled and mixed. The mixture is now split into three equal portions and placed in vessel 1, 2 and 3. *Each* vessel now contains three mixtures: polymer beads covered with A, B and C.

In the second phase, Vessel 1 is reacted with chemical D, vessel 2 is reacted with chemical E and vessel 3 is reacted with chemical F. The result of this reaction is as follows:

- vessel 1 contains polymer beads carrying A-D, B-D, and C-D
- vessel 2 contains polymer beads carrying A-E, B-E, and C-E
- vessel 3 contains polymer beads carrying A-F, B-F, and C-F

The output of the second synthesis phase is a diversity of $3^2 = 9$ compounds. Without the splitting and mixing of the polymer beads after the first synthesis round and the combining of the individual pools, the second synthesis round would have yielded only 3 compounds. Combinatorial chemistry can increase the number of compounds by the power of the synthesis phases, resulting in a large chemical diversity very quickly. For instance the third round of the example given above would result in $3^3 = 27$ compounds (see table A.1).

(INSERT TABLE A.1 ABOUT HERE)

Chip technology and solution libraries are following different chemistries but have similar exponential increases in compounds generated with each synthesis round. Rapid parallel synthesis, however, achieves its efficiency gain through robot technology and parallel execution of many individual reactions. However, the efficiency gain is not exponential with the synthesis phase.

Table 1 (from Claris Draw)

Testing Phase[1]	Uncapitalized expected cost	Mean phase length (months)	Capitalized expected cost[2]
Preclinical	65.5	42.6	155.6
Long-term animal	5.3	33.6	8.2
Other animal	0.4	33.6	0.7
Phase I	9.3	15.5	17.8
Phase II	12.9	24.3	21.4
Phase III	20.2	36.0	27.1
Total	**113.6**		**230.8**

[1] The New Drug Approval (NDA) review period was estimated to last 30.3 months.
[2] Costs were capitalized at a 9% discount rate.

Table 2. Expected phase costs per New Chemical Entity (NCE) (in millions of 1987 dollars) (from DiMasi, Hansen, Grabowski and Lasagna 1991). All costs were deflated using the GNP Implicit Price Deflator. A 23% clinical approval rate was utilized.

Project Variable	Combinatorial Approach	Traditional Approach [1]
(1) Total development time	**3.5 months**	**5 years [2]**
(2) No. of chemists needed	4	15 [2]
(3) No. of compounds tested	~ 9000	~3750 [3]
(5) No. of (serial) rounds	1 [4]	100 (250 max.) [5]
(6) No. of compounds per round	~ 9000	~38
(7) Cost of screen per round	$10,000	$10,000
(8) Total cost (chemists only) [6]	$167,000	$18.75 Mill.
(9) ...per compound	$19 [7]	$5000

[1] Based on estimates from developers very experienced with medicinal drug development.
[2] Typical time and resources planned for a project of the given complexity and strategic importance.
[3] A skilled chemist can prepare 50 compounds per year.
[4] During a short second round, 220 compounds were prepared over a 2-week period. The compounds were a subset of the first round and did not contain new members.
[5] While 250 is theoretically possible, it doesn't allow sufficient time for learning and analysis between rounds. Thus 100 rounds is a realistic number.
[6] A skilled chemists costs approximately $250,000 per year. In the combinatorial approach, chemists were only need for 2 months.
[7] Since the marginal cost of preparing additional compounds using combinatorial chemistry is negligible, a ten-fold increase in the number of compounds prepared would result in a per unit cost of approximately $2.

Table 3. Comparing combinatorial chemistry with traditional medicinal chemistry in the discovery of promising lead compounds for the treatment of glaucoma. (The project was followed by further refinements of the lead candidates in order to increase the probability of clinical success.)

Phase	Vessel 1	Vessel 2	Vessel 3	N
1	A	B	C	3
2	A-D, B-D, C-D	A-E, B-E, C-E	A-F, B-F, C-F	9
3	A-D-G, B-D-G, C-D-G A-E-G, B-E-G, C-E-G A-F-G, B-F-G, C-F-G	A-D-H, B-D-H, C-D-H A-E-H, B-E-H, C-E-H A-F-H, B-F-H, C-F-H	A-D-I, B-D-I, C-D-I A-E-I, B-E-I, C-E-I A-F-I, B-F-I, C-F-I	27
k	...	...	...	3^k

Table A.1. Building solid support libraries. At the beginning of each round, a new reagent is introduced to each vessel. (For example, A is added to vessel 1 in round 1, D is added to vessel 1 in round 2, G is added to vessel 1 in round 3, etc.)

Figure 1 (from Claris Draw)

Figure 2. The rapid increase of scientific publications on combinatorial chemistry (source: Science Citation Index which covers 90% of the world's significant scientific and technical literature).

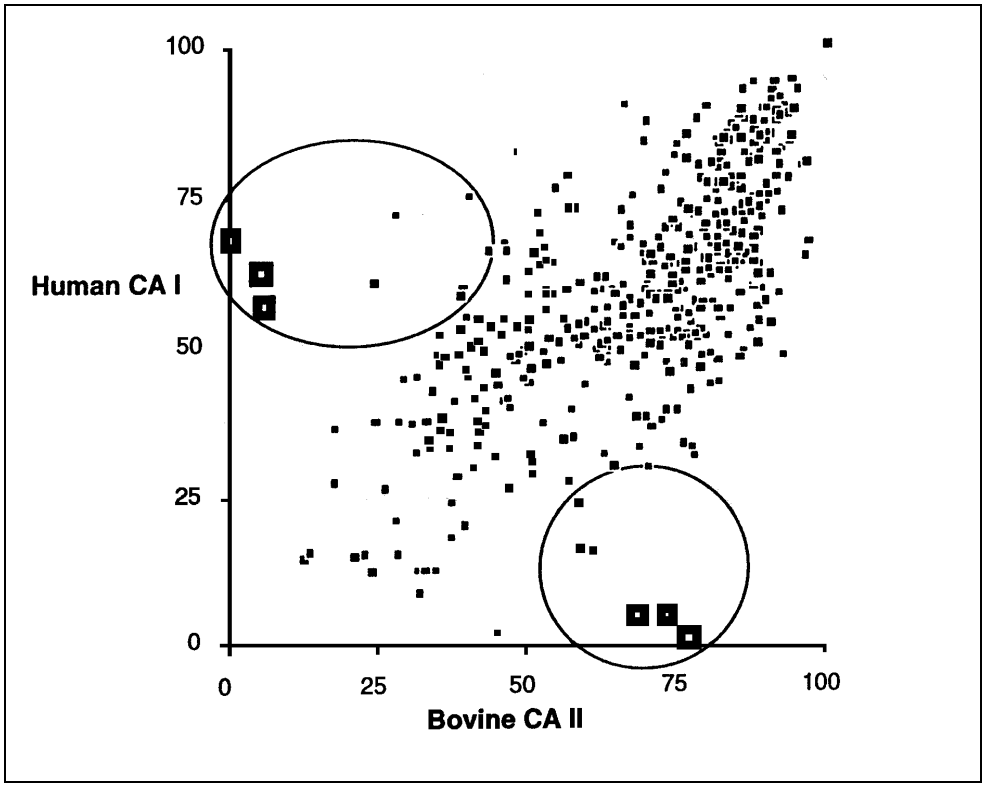

Figure 3. Each of the points represents a compound tested. The abscissa and ordinate indicate the affinity of the compounds for the human and bovine isozyme respectively. Note that, from the large number of compounds screened only three compounds per receptor (shown as large squares) displayed the desired discriminatory capability.