

This PDF is a selection from an out-of-print volume from the National Bureau of Economic Research

Volume Title: The Rate and Direction of Inventive Activity: Economic and Social Factors

Volume Author/Editor: Universities-National Bureau

Volume Publisher: UMI

Volume ISBN: 0-87014-304-2

Volume URL: <http://www.nber.org/books/univ62-1>

Publication Date: 1962

Chapter Title: The Link Between Science and Invention: The Case of the Transistor

Chapter Author: Richard Nelson

Chapter URL: <http://www.nber.org/chapters/c2141>

Chapter pages in book: (p. 549 - 584)

The Link Between Science and Invention: The Case of the Transistor

RICHARD R. NELSON

THE RAND CORPORATION

Introduction

PURPOSE AND FORM OF THE PRESENT STUDY

THIS paper is a case study of basic research in industry. It focuses on the factors affecting the allocation of research resources in a science oriented industrial research laboratory.¹

In the summer of 1948 the Bell Telephone Laboratories announced the invention of the point contact transistor—a small, efficient amplifying device. In 1951 the Laboratories announced the invention of the junction transistor, a device in most ways superior to the point contact transistor. In 1952 a Nobel prize was awarded to three of the principal participants in the research which led to these inventions.

The transistor was the result of the research of a group of extremely able scientists working close to the frontiers of scientific knowledge in one of the world's largest industrial research laboratories. As such, the invention of the transistor is a model of what many writers believe to be the coming norm. In another paper I have argued that the large industrial laboratory has by no means usurped the field of invention—the private inventor still is playing a very important role—and that many important inventions still are being made quite independently of closely preceding scientific advances.² But the role of the industrial laboratory is important and growing and it is likely that, particularly in the large laboratories, a growing share of inventive activity will be motivated by recent advances in fundamental science.

The inventive activity studied in this paper—the type which has led to such technical advances, in addition to the transistor, as nylon,

¹ This paper is one part of a continuing RAND-sponsored study of industrial research and development. The material in this paper was obtained from several published studies of the transistor and of the Bell Telephone Laboratories, from written laboratory records, from letters, and from a number of discussions with Bell Laboratories' scientists and administrators, to whom the author is deeply indebted. However, the views and opinions presented in this paper do not necessarily reflect the views of any Bell scientist, much less the consensus at Bell.

² R. R. Nelson, "The Economics of Invention: A Survey of the Literature," *Journal of Business*, April 1959.

dacron, hybrid corn, radar, and the new molecular electronics—is not generally an activity motivated by, and directed toward, the objective of a closely defined marketable product.³ The research scientists involved may have an idea of the practical fruits to which their work may lead. But the research projects undertaken are far less closely focused on a single objective than were, say, the projects which led to the automatic cotton picker, the jet engine, and Kodachrome, where the scientists and technologists were aiming at one particular target.⁴

While economic analysis of the demand for a new product and cost of development and production seems to go a long way toward explaining the direction of inventive activity when that activity is aimed at specific new products,⁵ much less is known about the factors stimulating inventive activity of the sort which led to nylon and the transistor. While perceived demand certainly plays an important role, advances in the state of scientific knowledge also seem to be extremely important. But what are the mechanisms by which changing scientific knowledge stimulates changes in inventive activity? It is hoped that this paper will shed a little more light on the subject.

Little use of formal theory will be made in this paper because of the constraint of space and because the formal theory itself is in a most unsatisfactory state. It is hoped that at some later date this study, and a number of complementary ones, can be recast within a framework of decision and organization theory. However, whatever specific form the theoretical framework may subsequently take, it is reasonably clear now that certain variables will play a major role, and these are the variables this study will stress.⁶

Decision theory examines problems relating to how a choice among alternatives should be made. When there is uncertainty, decision theory examines such questions as how much information should be acquired before a final decision is made, and how alternatives should be screened and narrowed down. The answers to the questions posed by decision theory depend on such variables as the following: the nature of the payoff function; the extent to which the best decision

³ For very short histories of several of these inventions see J. Jewkes, D. Sawers, and R. Stillerman, *The Sources of Invention*, London, Macmillan, 1958.

⁴ *Ibid.*

⁵ See Nelson, *op. cit.*

⁶ The approach will be the schizophrenic one usual in economics. On the one hand the analysis will be normative. The interest will be on how people should make decisions. On the other hand the analysis will be predictive, under the assumption that most people making the relevant decisions are quite rational.

differs depending on variables outside the decision maker's control or outside his knowledge, or both; the means and costs of reducing relevant uncertainties, i.e., the ability to learn; and the extent to which the decision maker can keep his set of alternatives open (not commit himself) until he receives more information.⁷

Organization theory examines problems relating to the interactions among a group of individuals who make decisions relevant to a possible pay-off. All the factors important in decision theory must, of course, enter the analysis of organization theory, but the focus of organization theory is on such questions as, what information should be communicated to what people, and who should make what decisions on the basis of what criteria. The answers to the questions posed by organization theory depend on the extent to which individual actions are interdependent, the nature of relevant information and the character of those who acquire it, the difficulty of effectively communicating information and orders, and other variables.⁸

The elements of the theory mentioned above will be used only informally in this paper. But pretend for the moment that the theory is in very good shape. What would be the implications for the organization of this paper, or for a paper incorporating several case studies of industrial basic research? I think the theory would suggest that it would be useful to organize the paper in three parts. First would be the case study or studies. These studies would be reasonably straightforward histories but focusing on the variables that theoretically determine the way decisions should be made and effort should be organized. The purpose of the histories would be to determine the values, or frequency distribution of values, these variables take on in industrial basic research. Second, the values of the variables would be plugged into the theory and the correct decision and organization policies deduced. Third, an organization, or a number of organizations, with a reputation for effective research management would be studied to see to what extent their policies corresponded to those which the theory indicated were sound. This would provide a quite useful check on the theory.

⁷ Decision theory presently seems to be going in many different directions. See, for example: *Decision Processes*, R. M. Thrall *et al*, editors, New York, Wiley, 1954; A. Wald, *Statistical Decision Functions*, New York, Wiley, 1950; H. A. Simon, "Theories of Decision-Making in Economics and Behavioral Science," *American Economic Review*, June 1959; R. D. Luce and H. Raiffa, *Games and Decisions*, New York, Wiley, 1957.

⁸ See, for example: J. G. March and H. A. Simon, *Organizations*, New York, Wiley, 1958; and T. Marschak, "Elements for a Theory of Teams," *Management Science*, January 1955.

Since theory is presently in such an inadequate state, this paper cannot be organized quite as sketched above. In the next section a reasonably straightforward history of the research which led to the transistor will be presented. The history will stress the pay-offs which were expected, the uncertainties involved, how new information came to light, how this new information affected the direction of the project, and the interaction of the several people working on the project. In a subsequent section a study of the organization and decision policies of the Bell Telephone Laboratories will be presented. There will be no formal theoretical section. However, the transistor history will conclude with an attempt to generalize on the characteristics of the research involved, and the organization study will be prefaced with an attempt to assess what these characteristics imply with respect to the problems of research management.

THE BELL TELEPHONE LABORATORIES AND THE TRANSISTOR

Although it is hoped that the analysis presented here has some general applicability, the paper is principally a single case study—a study of the research that led to the invention of the transistor and of the organization responsible for that research—the Bell Telephone Laboratories.

Bell Laboratories are jointly owned by the American Telephone and Telegraph Company and Western Electric, A. T. and T.'s production subsidiary. The Laboratories employ about 11,000 people, of whom about one-third are professional scientists and engineers, about one-third are technical aides, and about one-third, clerical and supporting personnel. About 85 per cent of the laboratory professionals are engaged in the development of specific devices and systems for use in the telephone system, or by the military. About 15 per cent of the professional scientists and engineers, about 500 people, constitute the research staff under William O. Baker, Vice-President in charge of research. Baker reports directly to James B. Fisk, President of the Laboratories. A large percentage of Baker's budget supports scientific research which is not tied to any specific practical objective. It is on this kind of research that this paper is focused.

The transistor was invented in the course of a research program started in 1946 at the Bell Telephone Laboratories. The transistor has several advantages over the triode vacuum tube, Lee De Forest's invention of half a century ago; it is much smaller than the vacuum tube, it requires much less energy input to do a given job, and in many

applications it is much more durable. On the negative side, the performance of the transistor is much more sensitive to varying temperature than is that of the vacuum tube; at present the transistor is not as capable as the vacuum tube at high frequencies and in handling high power, and thus far, problems of quality control have proven quite serious in transistor manufacture.

The transistor has had its most significant impact not as a component replacing vacuum tubes in established products, but as a component of products which were uneconomical before the development of the transistor. There has been an increase in the use of electronic packages where the transistor's strong points are important. Very compact computers are the most striking example. Without transistors, computers of a given capability would have to be much larger both because vacuum tubes are larger than equivalent transistors and because cooling requirements are much greater for vacuum tubes. Almost all of our new airborne navigation, bombing, and fire control systems, for example, are transistorized. So are all of our satellite computers. And without transistors our large computers, which are playing an increasingly important role in science, engineering, and management, undoubtedly would be much more expensive—probably so much so that many of their present uses would not be economically sound.

Thus the transistor has stimulated growth, including the invention and innovation on a considerable scale of products which can profitably use transistors as components. The transistor has also stimulated research and development aimed at reducing the size of complementary electronic circuit elements. Much of the work in printed circuitry, for example, certainly fits this picture. Further, as we shall see, the research which led to the transistor also produced a number of other new and improved semiconductor devices. Thus, if it be argued that one of the indexes of the importance of an invention is the amount of inventive activity it stimulates, then, by this criterion, the transistor is a major invention indeed.

But while the transistor is playing an extremely important role in our more complex electronic equipment and in equipment where size is important, like hearing aids and portable radios, it has not superseded the vacuum tube in all uses. Dollar sales of vacuum tubes are still roughly double dollar sales of transistors. Given existing costs, vacuum tubes are now more economical in the bulk of those jobs in which size and efficiency are not particularly important. And since

the birth of the transistor there have been substantial improvements in vacuum tubes, many of these improvements certainly stimulated by the competition of the transistor. Almost no invention eliminates all of its competition overnight, and the transistor is no exception. But there is reason to believe that the transistor may be one of the most important inventions of the twentieth century.

The History of the Transistor

HISTORY OF SEMICONDUCTOR RESEARCH BEFORE THE PROJECT AT BELL LABORATORIES

The transistor is a semiconductor device. The research at Bell which resulted in the invention took off from a base of knowledge about semiconductors built by several generations of scientists. Karl Popper has described the state of scientific knowledge at any time as the deposit of observations and conceptual schemes which have stood the test of time and which still are proving useful in explaining and predicting.⁹ The current state of knowledge is the result of an evolutionary process operating on ideas. Therefore, in order to understand the post-World War II research at Bell, it is important to sketch the history of prior semiconductor research.¹⁰

The element germanium is a semiconductor. So are several other elements, including silicon, and a number of compounds, such as copper oxide and zinc oxide. By 1900 many scientists and experimenters with electricity knew that these materials had quite unusual properties. In particular, it was known that these materials conducted electricity although, as the name implies, not as well as conductors like metal. It also was known that the electrical resistance of these materials decreased with temperature. That is, when these materials were warm they conducted current more easily than when they were cold. This puzzling property set semiconductors apart from other conductors, like metals, which conducted more easily when cool than when warm. Also, it was known that these materials sometimes passed current more easily in one direction than in another. In other words, they rectified current.

In this section we shall see how these phenomena, and others, gradually came to be understood. We shall see how research led to a

⁹ Karl R. Popper, *The Logic of Scientific Discovery*, Basic Books, 1959.

¹⁰ The history that follows is primarily taken from G. L. Pierson and W. H. Brattain, "History of Semiconductor Research," *Proceedings of the Institute of Radio Engineers*, December 1955.

distinction between n type and p type semiconductors, and how it came to be realized that in a semiconductor there are both electrons and "holes." We shall see how, by the start of World War II, researchers had come to an understanding of semiconductors that was satisfactory in many respects. We shall also see that one important concept, that of "minority" carriers, was being neglected. This will set the stage for analysis of the research at Bell which led to the transistor.

Necessary Terms and Concepts

In order to understand the history of semiconductor research it is necessary to be familiar with a few terms and concepts, so before proceeding with the history let us leap ahead and consider a few aspects of modern-day theory.

Modern theory views a semiconductor as a crystal containing two different types of current carriers—electrons and holes. The electrons are negatively charged and the holes may be considered as positively charged. The number of holes and electrons which are free to carry current is an increasing function of the temperature of the crystal, which explains the negative coefficient of resistivity which had puzzled researchers around the turn of the century.

The proportion of holes and electrons in a semiconductor is a very sensitive function of its purity. By doping a germanium crystal with other elements a very great variation in the hole-electron ratio can be achieved. In some semiconductors there are many more electrons (negative charge carriers) than holes (positive charge carriers). Thus electrons are the "majority" carriers and holes are the "minority" carriers, and this type of crystal is called, conveniently enough, an n (for negative) semiconductor. In other crystals, holes are the majority carriers and this sort of semiconductor is called a p (for positive) type.

If a p and an n type crystal are placed end to end they form a " p - n junction." Such a p - n junction will conduct current much more easily in one direction than in the other, in other words, it is a rectifier. The theoretical explanation of rectification in p - n junctions rests on the fact that on one side of the junction (the n side) most of the charge carriers are negative, while on the other side of the junction (the p side) most of the charge carriers are positive. The explanation makes no use of the concept of minority carriers. However, as we shall see later, in the working of the transistor minority carriers play a key role. Without an understanding of the fact that there are minority

carriers as well as majority carriers, the operation of the transistor cannot be comprehended.

Research Before World War II

Returning to our story, though by 1900 many scientists knew that the materials we now call semiconductors had interesting properties, they knew little about why. The birth of the radio industry created a practical demand for good rectifiers and quite early in the game "cat's whisker" rectifiers (semiconductors) became widely used. But, in large part because the vacuum tube rectifier was better understood and hence the direction of possible improvement more clearly indicated, the semiconductor rectifier declined in importance relative to the vacuum tube during the twenties and early thirties.

However, during this period research on semiconductors was far from stagnant. A number of experiments made it quite clear that in some semiconductors the charge carriers behaved as if they had a positive charge, and several scientists were coming to the belief that there were two quite different types of semiconductors.

During the thirties research workers in the field of radio waves and communications turned their interest to higher frequencies. The ordinary vacuum tube performed poorly at these higher frequencies, and attention returned to crystal detectors and hence to research on semiconductors. Improvements came rapidly. Techniques for producing very pure silicon were improved and metallurgists learned how to add very accurately measured quantities of impurities. At the Bell Laboratories it was learned that when silicon ingots were doped with certain elements (arsenic, phosphorous, antimony) the rectifying contact would conduct easily only when the crystal was negative relative to the metal, and that when silicon ingots were doped with certain other elements (aluminum, boron) the rectifying contact would conduct easily only when the crystal was positive relative to the metal. The first type of semiconductor came to be known as the *n* (for negative) type and the second as the *p* (for positive) type.¹¹

Thus, before the war scientists at the Bell Telephone Laboratories

¹¹ The semiconductor types were named before theoretical understanding of the differences was achieved. Note that for the first type of semiconductor the rectifying contact conducts when the metal is positive relative to the crystal, so it could have been called *p* type, and similarly the second could have been called *n* type. Had they been, language use would have worked against the understanding of semiconductors. The situation would have been much like that in early 19th century electrical theory, where Franklin's convention of defining current flow direction as from plus to minus hindered understanding of electron flow.

and elsewhere were experimenting with p and n crystals and calling them that. Also before the war many scientists were thinking of making a semiconductor amplifier. The reasoning on the prospects for an amplifier was principally in terms of a simple analogy. Vacuum tubes rectified and, with the introduction of a grid, amplified. Semiconductors rectified. Therefore, somehow, they should be able to amplify. Indeed several workers suggested that a grid should be inserted into a semiconductor diode, but due to the extreme thinness of the rectifying area (rectification occurs in a region very close to the surface contact of cat's whisker rectifier, or at the n - p junction of a junction rectifier) this proved very difficult to do.

Meanwhile the conceptual scheme which we have described earlier and which would permit the workings of the semiconductor to be much better understood was gradually taking shape. Advances in quantum mechanics during the twenties set the stage for A. H. Wilson's quantum mechanical model of a solid semiconductor which was published in 1931. The Wilson model provided the basis for a theoretical explanation of the difference between n type and p type semiconductors, but, although the model was well known, few scientists saw this until after World War II. Indeed it seems that it was not until the postwar project at Bell Laboratories that the model was extended to apply to doped germanium and silicon. Further, and this is extremely important, though scientists in the field knew, or should have known from their feel for the above theory, that every semiconductor had both positive and negative charge carriers, their attention was focused almost exclusively on majority carriers. Thus n type germanium was pictured as having electrons carrying current, p type as having holes carrying current, and the minority carriers were ignored. The working of the transistor depends on minority carriers as well as majority carriers, and until both types were considered together understanding was sorely hindered.

To one who is not a physical scientist it is interesting that by the mid-1930's a well-known article, Wilson's, contained most of the essential ingredients for a rather complete understanding of semiconductors, but that almost all scientists missed some of the essential points. Similar instances in economic theory are legion. By the start of World War II, then, scientists understood quite well certain aspects of semiconductors, and were well on the road to understanding rectification. The essential theoretical foundations had been laid, but the phenomenon of minority carriers was being neglected. To many

scientists doing research in the field the time seemed ripe for major breakthroughs.

THE PROJECT AT BELL

The Start of the Project

In this section we shall follow the research work at the Bell Telephone Laboratories which led to the transistor. We shall see how difficulties with the first solid state amplifier design which interested Shockley led to a series of experiments which resulted, quite surprisingly, in the discovery of the transistor effect and the invention of the point contact transistor. We shall see how, in attempting to explain the transistor effect, the concept of minority carriers snapped into focus, and how this concept led to the invention of the junction transistor.¹²

The Bell Telephone Laboratories were deep in research in the field of semiconductors long before World War II. Bell's tradition in quantum mechanics was very strong. Several of the experiments which demonstrated the wave nature of electrons were conducted at the Laboratories during the thirties by Clinton J. Davisson; and Davisson and G. P. Thompson of England shared a Nobel prize for their efforts. William Shockley, Walter Brattain, Dean Wooldridge, and several other top-flight solid-state physicists were brought to the Bell Laboratories during the thirties. People in the Laboratories' metallurgy department were playing a major role in the advances then being made in producing pure crystals, and J. H. Scaff and others at Bell performed the experiment which led to the naming of *n* and *p* type crystals. Also, the pre-World War II work of the Laboratories in this area had led to the development of better crystal rectifiers, the thermistor—a circuit element whose resistance decreased with temperature—and other circuit elements.

During World War II research work in semiconductors continued at the Laboratories, but the war work was, of course, device oriented and several of the Bell scientists worked elsewhere. Semiconductor knowledge and technique played an extremely important role in the war-time work on radar. At Bell and at Purdue University research on germanium and silicon was pushed with the aim of developing better rectifiers, and important breakthroughs were made in techniques of

¹² The material in this section is taken from several sources, the most important of which are the Shockley and Bardeen Nobel Prize lectures.

producing very pure crystals and doping them to very close specifications.

During the summer of 1945, as it became evident that the war would soon be over, steps were taken to smooth the transition of the Laboratories to a peacetime basis. Shockley, who had been on leave from Bell during the war, believed strongly that the Laboratories should intensify its solid-state work, and in consultation with M. J. Kelly, then Director of Research, and others, was convincing enough so that it was agreed that it would be good research strategy to bring together in one department a number of people who had been working on solid-state physics, and perhaps to draw in some new talent. Within the solid-state research group, to be headed by Shockley and S. O. Morgan, a subgroup, including Shockley, was to work on semiconductors. It was felt that the greatly increased role of solid-state devices, particularly semiconductors, in communications technology warranted an increase in Bell's effort in this area. It was felt that advances in the understanding of semiconductors, including a better grasp of the meaning of the quantum mechanical model, had set the stage for major breakthroughs, and that the techniques of making crystals to close specifications promised materials which could be produced to fit the theoretical model. Improvements in rectifiers, and thermo- and photo-electric devices were judged a quite likely result of semiconductor research. Further, Shockley strongly believed that he could make a solid-state amplifier and his enthusiasm was contagious.

Shockley and Morgan were given hunting licenses and Brattain and Pearson were talked into joining the semiconductor research group. John Bardeen was hired from the outside. Later R. B. Gibney, a physical chemist, and H. R. Moore, a circuit expert, joined the group. The prime reason for establishing a special solid-state physics research group was the belief that the interaction of physicists, chemists, and metallurgists all interested in related problems would be instrumental in speeding the advance in understanding, and that the organization of a separate group would facilitate communication and mutual help. But it would be a mistake to believe that all the work on semiconductors going on at the Laboratories after 1946 was done by this subgroup. Throughout the period people in the metallurgy department were working on ways to make better crystals and rectifiers and there was considerable interplay between Shockley's group and the metallurgists. And scattered throughout the research divisions were people

working in this area from time to time. All in all, the people playing a major role at one time or another in the work which led to the transistor discovery may have numbered about thirteen.

We have seen that the motives of the Bell Telephone Laboratories in establishing this new project were reasonably clear. Bell believed that major advances in scientific knowledge in this field were likely to be won and that advances in knowledge were likely to be fruitful in improving communications technology. One possible result was an amplifier. But improvements in rectifiers, thermistors, and other solid-state devices also were judged strong possibilities. It was the wide range of possible useful results which made the project attractive. The motives of the scientists on the project were, of course, much more complex. Several of the scientists involved were not much interested in or concerned with any practical applications their work might lead to. Their intellectual interests were focused almost exclusively on creating more knowledge about semiconductors. Others in the group were concerned with practical applications as well as with the underlying sciences. Shockley's interests were multiple. As a theorist he was fascinated by the prospect of developing a good theory of semiconductors. He also was fascinated with the prospects for a solid-state amplifier. Shockley's work was focused in a direction compatible with both these ends. That a good share of the work of the semiconductor research group was allocated so as to clear the way for an amplifier seems, in large part, to have been the result of Shockley's influence. But it is extremely difficult to say how much of this influence was "authority" and how much was Shockley's ability to interest others in what interested him.

Research Leading to the Discovery of the Transistor Effect

Stated in broad terms, the general scientific aim of the semiconductor research program was to obtain as complete an understanding as possible of semiconductor phenomena, not in empirical terms, but on the basis of atomic theory. Wilson's work was an important start for, as we have seen, a sound theoretical foundation already was partially built although it had not been fully exploited in thinking about doped germanium and silicon. With the wisdom of hindsight we know that one important roadblock to understanding of doped crystals was failure to consider minority carriers, the flow of electrons in p type germanium and holes in n type germanium. Also (though not to be discussed in this paper), the workers in the field were not

adequately treating surface states, i.e., the fact that the properties of a solid at the surface can be (and usually are) quite different from the properties in the interior.

During the first few years of the solid-state physics project (the years we are considering here), the Shockley-Morgan group operated on a budget of roughly half a million dollars a year, about enough to support twenty to thirty scientists. Probably less than half of the group worked with Shockley on semiconductors. Both because the developing theory was simpler and better understood for simple crystals, and because the metallurgists at Bell had developed ways to produce very pure germanium crystals and to introduce impurities to close specifications, the work of the semiconductor research group focused on germanium at the start and later broadened to include silicon. During 1945 and 1946 much additional experience was gained in growing crystals, quite a bit was learned about semiconductor impurities, and rectification was studied. During the early years of the project research interest was quite diffuse. Considerable work, however, was focused on a solid-state amplifier.

At the early stages of the effort at Bell, Shockley's ideas on possible ways to make amplifiers shifted from placing a grid in the area of rectification (the strict analog of the vacuum triode) to influencing the number of movable electrons in a semiconductor (and hence the current flow) with an electric field imposed from the outside without actually touching the material. Shockley's calculations, based in large part on his developing extension of the Wilson model, indicated that a device based on the latter idea would amplify. Experiments were devised (in 1946 and 1947) to see if the gadget worked as theory indicated it should. It did not. Sometimes even the sign of the effect was off. When the sign was right the magnitude of the effect was roughly a thousandth of the theoretical effect.

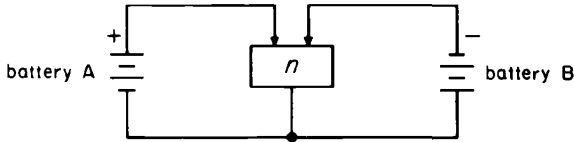
To explain the negative result Bardeen proposed that the electrons affected by the field were not free inside the silicon, but were trapped at the surface in what he called surface states. Thus the application of an electric field would not significantly affect the number of free charge carriers in the semiconductor. Other scientists, including Shockley himself, had previously suggested the possibility of surface states at the free surface of a solid, but no one had realized the importance of this phenomenon to the properties of semiconductors. Bardeen's theory very effectively explained failure of amplification in the field effect experiments and also significantly increased understanding of

rectification at the junction of a semiconductor and a metal (cat's whiskers), but for our purposes it is not necessary to describe the theory.

What is important for our purposes is that, to test the Bardeen theory and to attempt to find a way to neutralize the surface states so that a useful field effect amplifier could be built, Shockley, Bardeen, and Brattain, often with the collaboration of others in the group performed a number of experiments. The physical phenomena involved seem to have drawn a great deal of interest.

A set of experiments by Bardeen and Brattain play the key role in our story. These experiments did yield observed amplification from a field effect. But more important, in one of the experiments two contacts were placed quite close together on a germanium crystal. (Figure 1). It does not matter here just why this particular experiment

FIGURE 1



was performed. What is significant is that it was not performed with the hope of observing the most important result it yielded. For in the course of the experiment it was observed that connecting up the A battery increased the current flow in the B battery circuit. The device amplified. This was the first indication of the transistor effect.

The research workers were very well aware of the importance of their discovery. Experiments motivated, in part at least, by a desire to make a field effect amplifier work had resulted in the discovery of an amplifier working on quite different principles, an amplifier which, as it was developed and perfected in subsequent work, came to be called a "point contact" transistor.

The most obvious explanation for the current flow increase in the B circuit induced by connecting the A circuit was hole flow from the left top contact to the right top contact. The key concept in the modern explanation is the flow of minority carriers in a crystal, holes in n type germanium. Although the latter was not seen immediately, gradually it came to be accepted.

This description of the discovery of the point contact transistor in late 1947 and very early 1948 is too neat and simple, and the outline of the development of the ideas which explained it is much too orderly to be accurate. Indeed, there still is no really adequate quantitative theory explaining the working of the point contact transistor. But what is important for us is that the experiments were conducted by men who had amplification as their goal, who observed something that they were not looking for or expecting which indicated the possibility of building an amplifier of a design very different from the one they had in mind, and who explained the working of this new amplifier in terms of injected minority carriers.

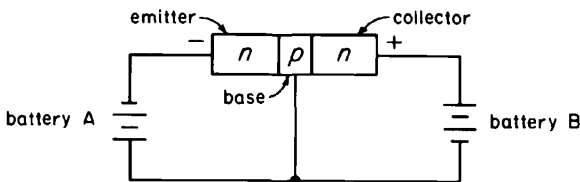
For the observation that minority carriers are important in semiconductor current flow provided the key which enabled Shockley to propose still another design for an amplifier, a junction transistor. Thus the field effect experiments led to two amplifier designs. They led directly to the point contact transistor and, through theory, to the junction transistor.

The Junction Transistor

Shockley spent most of 1949 writing his *Electrons and Holes in Semiconductors*. Much of the book is a contribution to theoretical physics, but in it he describes the principles of the junction transistor. Unlike the point contact transistor which, as we have seen, was invented partially by accident, the junction transistor was predicted theoretically and then built. Essentially the theory was the invention.

An *n-p-n* junction transistor consists of a germanium or silicon crystal with two *n* regions separated by a thin *p* region (Figure 2).

FIGURE 2



The contacts to the *n*, *p*, and *n* layers are referred to as the emitter terminal, the base terminal, and the collector terminal. Shockley showed that an increase in the voltage across the A circuit would lead

to an increase in the flow of electrons between the emitter and the collector, the electrons flowing right through the central p region. And he showed that, for appropriate battery biases, the resulting change in voltage across the B circuit would exceed the inducing change across the A circuit. Or, the transistor would amplify.

Notice that minority carriers, electrons flowing through the p type base region (injected from the n type emitter region), are of crucial importance to the working of the junction transistor. If only holes could conduct in p type germanium then the transistor would not work. As we have seen, it took the semiaccidental discovery of the point contact transistor to focus attention on minority carriers. Until the experiments which resulted in the point contact transistor, Shockley did not see clearly that they were important. But once he did, he soon could see that an $n-p-n$ junction connected as in Figure 2 would amplify.

Research After the Discovery of the Transistor Effect

The discovery of the transistor effect and the consequent invention of the point contact and junction transistors acted to focus much more closely the interests of the semiconductor research group. Funds for this group came from the broader solid-state research group budget, and although the budget of the solid-state group was not dramatically increased for a year or so, it appears that the proportion of the research effort of the Shockley-Morgan group which was directed to semiconductor research definitely increased. Further, a special semiconductor development group under J. A. Morton was organized shortly after the birth of the point contact transistor. And the allocation of research in metallurgy definitely was affected.

Within the semiconductor research group Shockley, as we have seen, directed his attention to working out the theory of holes and electrons in semiconductors. Others directed their experimental and theoretical work to minority carriers. Since an effective amplifier of the junction design requires very pure and orderly crystals, people in Shockley's group, as well as in Morton's group and in metallurgy, intensified their research on this problem. Methods for pulling single crystals from a melt of germanium were developed by G. K. Teal and J. B. Little, and a method known as zone refining was invented by W. G. Pfann. During 1949 and 1950 better and better junction transistors were made and the construction of a reliable junction transistor is conventionally dated 1951.

THE LINK BETWEEN SCIENCE AND INVENTION

It should be stressed here that the breakthrough in knowledge about semiconductors led to research and development of several other devices in addition to the transistor. The theory led directly to the development of better rectifiers and thermo-electric devices. Parametric amplifiers, amplifiers quite different from the transistors, rest on the principles treated by Shockley's theory. The Bell solar battery, developed a bit later, also rests on the theory. Although many of the new semiconductor devices were invented in laboratories other than Bell, Bell Laboratories has continued to be in the forefront of semiconductor developments. Thus, though the transistor breakthrough tended for a while to draw together the interests of the semiconductor research group, the new theory carried within itself the seeds of subsequent research diversification. For the application and value of the theory and the new light it shed on semiconductors far exceeded its specific application to junction transistors.

1948 and 1951 are given as the standard dates when handbuilt point contact and junction transistors were first publicly demonstrated. But generally—and the transistor is no exception—the road is long and difficult between the first demonstrator model of a new invention, and a reliable, producible, economic product. In the early days of the transistor, its performance was likely to change if someone slammed a door. All transistors used were laboriously constructed by hand. No one quite knew in what uses the transistor would prove economic.

Much money and talent were spent in improving the operating characteristics of transistors and making them more predictable and reliable, in developing new circuits and designs to take advantage of the transistor's strong points, and in developing an economic production technology. More money has been spent at the Bell Telephone Laboratories on these problems than was spent on the project which was described in the preceding section, but this is beyond our story.

NATURE OF THE RESEARCH ACTIVITY

Although one case study is insufficient support for confident generalization, in this section I will attempt to sketch certain characteristics of the transistor research history which seem to have relevance to the policy decisions of an organization undertaking basic research.¹³

¹³ Several other studies of basic research show much the same pattern as the transistor history. See, for example, "Nylon," *Fortune*, July 1940; I. B. Cohen, *Science, Servant of Man*, Boston, Little Brown, 1948; James B. Conant, *Science and Common Sense*, Yale

Some implications of these characteristics will be very briefly examined in the introduction to the section treating the actual research policies of the Bell Telephone Laboratories.

Uncertainty and Learning

At the start of the semiconductor research project there was considerable uncertainty whether an amplifier could be achieved at all, and if so, what was the best way of achieving it. Shockley believed he could make a solid-state amplifier but was not quite sure how, and several other people were not so optimistic. Because of the great uncertainties involved, much of the research effort was directed toward learning rather than toward the achievement of a specific and well-defined result. Shockley and others, however, certainly hoped to achieve an amplifier and allocated much of their research time accordingly. And the direction of research changed dramatically in response to what was learned.

To illustrate these points, recall the chronology of the research. At the start of the project much of the semiconductor research was quite generally oriented; only as knowledge accumulated did the amplifier interest tend to predominate. And in the course of the project, three different amplifier designs were considered. In the early stages, almost all amplifier oriented research was concerned with the field effect amplifier, but this design did not work as hoped and expected. Indeed, the results of the critical experiments in this area were sufficiently unexpected that a new theory had to be formulated to explain them—Bardeen's theory of surface states. The allocation of experimental effort was shifted to check the theory. The second amplifier design, the point contact, was discovered as a more or less surprising result of the surface-state experiments. Further, in order to explain the results of the experiments, attention turned to minority carriers, and once the importance of minority carriers was clearly realized, Shockley was able to design a junction amplifier. Thus, only toward the end of the project was the design which has proved most successful clearly perceived.

This last point needs to be stressed. There never seems to have been an attempt to list all research alternatives and to pick the best on

University Press, 1951; H. D. Smyth, *Atomic Energy for Military Purposes*, Princeton, 1946. Many other references could be cited, though few deal in any detail with specific research projects.

the basis of some formal calculation. Rather, the *discovery* of new ideas and alternatives occurred often in the program. It does not seem an adequate representation to say that what was learned from the surface-state experiments made the junction transistor a more promising alternative. Before the experiments, the path to an amplifier by way of junctions and minority carriers just was not clearly perceived.

Interaction

A number of different people with different skills contributed to the research, but the exact nature of the interaction could not have been predicted and planned for in advance. Bardeen's analysis of surface states and Brattain's experimental skill were key factors, but no one at the start of the program could have predicted that these particular experiments would be performed, much less the importance of the results. Throughout the course of the program there seems to have been a great deal of informal exchange of ideas, and quite a bit of cooperation on experimental work.

The major requirements for effective interaction seem to have been easy communication and the ability of individuals to drop what they were doing to help with problems their colleagues brought up, if these problems seemed interesting and important. There is little evidence of closely programmed and directed team work, or of any requirement for it. Unfortunately, I have not been able to uncover much in the way of specific examples to illustrate these points.

Goals

As we have seen the research program was not justified in its early stages by prospects for the invention of a specific device. Indeed, the early project reports do not even mention amplifiers. Instead the purpose of the project is stated in the early reports as the advancement of knowledge about semiconductors. Of course, this does not mean that the research workers, themselves, had no ideas about possible specific practical results; they certainly did. An amplifier was one. Rectifiers and thermoelectric devices were others. But these devices were listed in the early project reports as "for instances" of the possible practical payoffs. And, as we have seen, the research did in fact lead to a number of technical advances in addition to the transistor.

*Research Management at the Bell Telephone
Laboratories*

THE PROBLEMS OF RESEARCH MANAGEMENT

This section will examine the way that basic research decisions are made at the Bell Telephone Laboratories. But first it seems worthwhile to discuss very briefly a few of the implications, with respect to research policy, of the characteristics of basic research sketched in the preceding section.

Given the resources an industrial laboratory intends to allocate to basic research, there are a vast number of projects to which laboratory personnel can be allocated. Clearly somehow some of these projects must be selected and the number of persons working on them be decided. The nature of basic research imposes serious constraints on the policies the laboratory can pursue in making these decisions.

Criteria

In order that some projects be selected and others rejected, some criteria must be used. The choice of good criteria must rest on the realization that a successful basic research project promises major advances in scientific understanding and that major advances in scientific understanding are likely to stimulate a wide range of practical inventions, though just what these technical advances will be it is very difficult to specify in advance. A project which is likely to have payoffs in a wide range of applications should not be justified on the basis of any one in particular, especially if it is not at all clear that the specific objective selected will be achieved. Indeed, for many organizations it just does not make very good sense to sponsor basic research at all. The outcomes of most basic research efforts are too unpredictable for a firm with a small range of products and processes to have confidence that the results will be of use to it. But, if the range of products and technology is wide, the firm may have some confidence that the results will be of practical value.¹⁴

Perhaps it can be assumed, as a first approximation, that for a company with a wide technological base there is a high correlation between the value to the company of the technical advances generated by research and the scientific advances achieved by research. Certainly

¹⁴ For an extension and elaboration of this line of argument, see R. R. Nelson, "The Simple Economics of Basic Scientific Research," *Journal of Political Economy*, June 1959.

there are major qualifications to this. Probably the criterion should be amended by the requirement that the area of research be one where applications to the technology of the company seem quite possible. But for a company with a wide enough technological base to support basic research at all, the criterion of "scientific promise" is probably better than one that stresses the value of specific predicted practical inventions.

The criterion "scientific promise," subjective though it may be, does seem operational. Recall that before the start of the transistor work a number of scientists agreed that the time was ripe for major advances in knowledge of semiconductors. And the restriction that the research be in fields where applications to company technology seem possible also is operational. Recall that scientists were able to list a number of applications which might result from a major breakthrough in knowledge of semiconductors, and many of their prophecies were fulfilled.

Authority

But who is to apply the criterion? Who is to decide which scientists will work on which projects? For reasons which will be discussed later, that criterion, if adopted, argues for a decentralized decision making structure. And the nature of the basic research activity, the changing knowledge which is the scientific goal of research, also argues against centralization. Further, the type of interaction we have noted in the transistor project requires that individuals be free to help each other as they see fit. If all allocation decisions were made by a centrally situated executive, the changing allocation of research effort called for as perceived alternatives and knowledge change would place an impossible information processing and decision making burden on top management. Clearly the research scientists must be given a great deal of freedom, and the type of decision which must be cleared through a central authority must be quite limited.

It is earnestly hoped that the preceding discussion, and the study of Bell Laboratories which follows, is more than a pious restatement of the old saws, "scientific knowledge in itself pays off" and "in research, freedom is good." It is true that I believe the theme generally is valid. However, I trust that this paper sheds a little more light on why it is valid, and what its limitations are. In particular, it is clear that a policy involving decentralized decision making (research freedom), and the acceptance of the criterion of scientific promise

must somehow be complemented by policies designed to constrain research to those areas where application to the technology of the company seems promising.

Bell has developed policies which seem to cope with this problem adequately. There is evidence that several other companies noted for their success in basic research management have similar policies.

RESEARCH POLICY OF THE BELL TELEPHONE LABORATORIES

The research philosophy of the Bell Telephone Laboratories has been stated succinctly by James B. Fisk, president of the Laboratories. "Our fundamental belief is that there is no difference between good science and good science relevant to our business. Among a thousand scientific problems, a hundred or so will be interesting, but only one or two will be truly rewarding—both to the world of science and to us. What we try to provide is the atmosphere that will make selecting the one or two in a thousand a matter of individual responsibility and essentially automatic." There are two aspects then to Bell's policy toward research. First, scientific worth is assumed to be highly correlated with potential technical value. And second, the individual scientist is to be free to choose the projects he considers of greatest interest. In short, the management philosophy of Bell Laboratories corresponds to the one we have just stated. How does this policy work out in practice? How serious are the problems introduced by the possible divergence of individual scientific interests from the interests of Bell Laboratories? How does Bell cope with the problem? As we shall see, the policy is not quite as simple and clear-cut as it seems. The fact that Bell is in the communications business does play a very major role in determining the type of research that is undertaken.

Freedom of Research Choice

It is well to dwell a bit here on what is meant by freedom in research. Freedom clearly involves a range of alternatives among which choice can be made but, equally clearly, the range of alternatives in any real situation is constrained in many ways. The training of the research scientist, the equipment available to him, the state of knowledge in the field in which he is interested, all are constraints on his choice. The imagination of the research scientist, of his peers in the laboratory, and of research management also affect the range of his perceived alternatives.

If we use the expression “degree of research freedom” to mean the extent to which constraints are imposed upon the range of choice open to a research scientist as a matter of company policy, then from the point of view of the company it is rational to give the research scientist wide freedom of choice if it is believed that he will tend to do more valuable work, on the average, if he selects his projects as he sees fit than if some higher authority provides him with a much more constrained range of choice. In a way, the arguments for freedom of choice for the research scientist resemble and rest on the same assumptions as the arguments for a free enterprise economic system do.

Let us assume for the moment that institutional arrangements and the system of incentives are such that individual and group aims coincide. This is one of the prerequisites for research freedom, but it is not enough. In addition it must be assumed that the individual has more information about the choices open to him—their relative prospects and their costs—than a central authority has. Usually this assumption is violated in situations where there is strong interdependence among the actions of a large number of different individuals. Where there is such interdependence one person’s decision should be made in the light of knowledge about what everyone else is doing. This is the kind of information that a central executive has, but that, if the number of people involved is large, the individual may not have. However, in basic research it seems that the number of persons among whom coordination is required is small at any one time, and the individual scientist is likely to know quite well what his colleagues are doing. In the transistor project this kind of knowledge was facilitated by the policy of bringing most of the scientists doing semiconductor research into one group. And as to the promise of his own line of research, the working scientist is likely to know much more than the laboratory management possibly can know.

Philosophers and Gestalt psychologists have argued convincingly that people do not observe the same things unless those things are in the same context, and at the frontiers of science what is missing is a clear context. A scientist working on a research project often has a “feel” for the promise of his work but is not capable of expressing his intuition, his embryonic theory, in language. Others working in the same area may have the same feel, but often not. Yet one of the marks of a good scientist is that his hunches prove right. The approaches he sees as promising but which he cannot in the beginning describe precisely blossom into useful hypotheses and conceptual schemes.

Often, after the fact, he can state clearly a plausible reason why he thought his work promising when he started it, and why he did things the way he did. But it is more than likely that, if you had queried him *ex ante*, he would not have given such a good set of reasons for why he was doing what he was doing and why he felt it important. If this point of view is accepted, a research director is likely to be a much poorer judge than the scientist working on a project is of the scientific promise of the project. The scientist has information that the research director cannot have, and much of the information is not in a solid enough form to permit easy communication.

Allocation of Resources

Given the nature of scientific research, and an organization where individual scientists have a wide degree of freedom, the allocation of the scientific staff among competing alternatives is likely to be accomplished by an evolutionary, or a natural selection process. We have argued that uncertainty and learning are key aspects of research. Consider a laboratory at a moment of time. Scientists are working on many research projects at many different stages. Some research work has been running for a long time and many of the original objectives have been achieved. Some research work started but a short time ago is looking quite unpromising already. But some research work which has just started is looking extremely promising and interesting, and some research work is producing new answers, new prospects, and new questions. An alert scientist working on a project which is looking increasingly unpromising or on a project which appears to be running into sharply diminishing returns has very strong incentives—his professional reputation, his scientific curiosity, and his future at the laboratories—to phase out his current work and phase in research in a more promising area—a new project or a going project which has exciting prospects.

The above description sounds very much like the economist's model of a changing economy. Some industries are dying and others are thriving. An astute entrepreneur spurred by a higher expected return will leave the dying industry and enter the thriving industry. And, as in the economist's picture of a changing economy, luck plays a great role. The structure of economic demand may change again, and the dull entrepreneur who sticks with the industry his more alert colleague thought declining will *ex post* have been shown to be the wiser. Meanwhile the alert entrepreneur may find that the skills he applied in the

old industry are of little use in the new. And in the laboratory, as in the economy, there are transfer costs. The scientist who jumps from project to project, like the entrepreneur who jumps from industry to industry, may be not very wise.

However, the analogy is weak when incentives and signals are compared. In the economist's model of the market economy, perceived profit opportunities are viewed as the prime motive, and cost and demand are the factors determining profit. Cost and demand are assumed to be clearly signaled by the structure of prices. The incentives of the scientist are much more complex than the incentives of the entrepreneur of economic theory; though perhaps they are no more complex than the incentives of the entrepreneurs of real life. The furthering of a reputation as a scientist among scientists and the satisfaction of intellectual curiosity are certainly as important, perhaps more important, dimensions of the research worker's goals as are financial advance and status within an organization.

Other Advantages of Decentralization

In addition to allocating research scientists among alternative projects, the mechanism described above seems especially well suited for generating new ideas, projects, and alternatives. Even if the performance of the mechanism in allocating men among known alternatives was believed to be far from optimal, this latter attribute would be a strong argument for research freedom. It is quite likely that it is even more important to the success of a laboratory for it to generate good new ideas than it is for it to allocate very efficiently its resources among a given set of alternatives.

Decision making by evolution, in addition to providing greater flexibility and speed than would be possible under a more formal and centralized structure, has another extremely important advantage. The traditions of the scientific community are extremely strong where freedom to pursue research interests is concerned. To be told just what line of research to follow—to have it made clear that the goal of the research is company profit, not increased knowledge or benefit to mankind—to realize all too plainly that a few individual supervisors, not a wide jury of scientific peers, are to evaluate the work—strikes hard at the traditions of science. Since World War II there seems to have been a striking reduction in the intensity with which scientists cling to these traditions. But their force is still strong, and many of

the most outstanding scientists feel that to engage in industrial research would prostitute their heritage.

The Bell Telephone Laboratories, perhaps more than any other industrial laboratory, has avoided establishing a decision and control system which would run against the traditions of the scientific community. For this reason the Laboratories have a great advantage, relative to similar institutions, in recruiting people. Many scientists who would work in no other industrial laboratory, will work at Bell. This fact explains the extremely high quality of the laboratory staff that Bell has managed to maintain. Bell salaries are not particularly high by industrial laboratory standards.

Because a scientist at the Laboratories is not forced to abandon the traditions of the scientific community, Bell scientists tend to maintain very strong links with the academic world. Many Bell scientists have taught at universities, and many Bell scientists are actively sought by university faculties. The quantity and quality of Bell's scientific publications is matched only by the best of universities.

In the work which led to the transistor, Shockley, Bardeen, and others were in close touch with members of the university community working in the field of solid-state physics. There were visits and many letters. Clearly, this close link with the main stream of the scientific community is of major importance to the Laboratories for the flow both of ideas and of men.

Management Controls: Environment and Employment

Although as a broad generalization it may be reasonable to argue that the scientific merit of a project is a good index of potential profit, clearly, from the point of view of the company, advances in some areas of science are likely to be much more profitable than advances in other areas of science. The argument that the results of research are uncertain does not imply that these areas cannot be specified. Thus from the point of view of Bell Laboratories, advances in knowledge of the magnetic properties of material are almost certain to prove more important than advances in botany, and probably more important than advances in organic chemistry. At the start of the semiconductor research project it was clear that if there were significant scientific breakthroughs, the payoffs for communications equipment would be great.

Further, though basic research projects should not be justified in terms of specific practical objectives, we have seen from the case of the

transistor that it is sometimes possible to perceive certain practical advances which may result from research. And from the point of view of the company, these objectives should be given some weight in research project selection.

Clearly it is in the interests of the company to have procedures for guiding research so that the most promising areas of science are stressed, and so that practical objectives are not completely suppressed in the research project selection mechanism. The Bell Telephone Laboratories are in the communications industry, and the research work undertaken there reflects this fact, though the research scientists are subject to few, if any, more constraints than their university peers.

Areas of Research

There are no geneticists working at the Laboratories. There are few organic chemists. At present (though this may be changed shortly) there is no group working in nuclear physics. In 1946 there was but a handful of men at Bell who were working in the field of solid-state physics. Today there are many more. While the choice of the research area of the scientists in the research department is seldom subject to strong executive pressure, formal executive decisions strongly affect the allocation of research effort through the hiring process. The growing promise of the research work in the field of semiconductors definitely strengthened the force of Shockley's requests for more people to work with him. And though "promising" is a difficult word to define in this context, prospects for improving communications equipment certainly do not hurt the promise of an area of research.

The decision to hire or not to hire new men certainly does not rest exclusively on evaluation of prospects for advances in different fields applicable to communications technology, but the areas of scientific research which are growing at Bell Laboratories are in general those in which significant advances are expected to have some application to communications technology. And as we have seen, changes in prospects are generated in the course of research activity itself. The striking success of the rather small semiconductor research group led to an expansion in the number of solid-state scientists working at Bell (as well as to an increase in the proportion of the veteran Bell staff interested in and working in solid-state physics). Thus the composition of Bell's scientific staff makes the starting of research work in a field not related to communications technology unusual.

But what happens when a member of the research staff at Bell, hired because at that time he was doing work in a field of interest to the company, becomes interested in a new field (as he sometimes does)? Several factors must be considered here. First, few at the Laboratories will be so rash as to state that research in a particular area of science certainly will not result in knowledge of use in the design of communications systems. Of course, this is not to say that certain areas are not considered much more promising than others. Second, the best way to find out whether a new area of science looks promising is to do some work in the area. Third, a policy of applying strong formal pressure to dissuade a scientist from working in a certain area would, if used as a matter of course, undermine the general philosophy of research direction at Bell which has proved so successful in the past, and further, would tend to make the Laboratories a much less attractive place to top-flight scientists. Therefore laboratory policy is to avoid pressuring a man not to work in an area of interest to him.

If the new work of an individual proves of significant interest, both scientifically and in possible communications applications, then it is likely that others in the laboratory will also initiate work in the field, and that people from the outside will be brought in. Thus a new area of laboratory research will be started. If the work does not prove of interest to the Laboratories, eventually the individual in question will be requested to return to the fold, or leave. It is hoped that pressure can be informal. There seems to be no consensus about how long to let a man wander, but it is clear that young and newly hired scientists are kept under closer reins than the more senior scientists. However even top-flight people, like Jansky, have been asked to change their line of research. But, in general, the experience has been that informal pressures together with the hiring policy are sufficient to keep A. T. and T. and Western Electric more than satisfied with the output of research.

Interest in Devices

The hiring process is also an important mechanism for keeping the laboratory alert to the possibilities for new devices. The Bell Telephone Laboratories tend to attract scientists who are also interested in devices. Shockley is a case in point. Further, problems relating to practical devices are often very interesting ones. Bardeen and Brattain found the problem of why Shockley's field effect amplifier did not work extremely exciting as a scientific one. And results worthwhile from the purest scientific point of view came from their efforts

to explain the failure of the amplifier. The fact that the research department is housed in the same building with the development and systems engineering departments also seems important in keeping the research people alert to the device needs of the Telephone Company. There are many interdepartmental seminars and formal talks. But perhaps informal contact is the most important link.

Device consciousness on the part of the research staff can lead to a serious misallocation of research effort if device oriented work is permitted to cut back sharply the effort directed toward advancing scientific knowledge. For this reason, though the research scientists are encouraged to be device conscious, laboratory policy discourages research people from working on a device beyond the building of a simple working model.

Thus, early in 1948, very shortly after the discovery of the point contact transistor, a development group under the direction of J. A. Morton was established. By 1950 there were more people engaged in development work than in research. The ultimate role of Morton's group was to create new devices for the systems development groups to use in creating new systems for Bell Telephone and for the military. The early objectives of the development group were to improve the point contact amplifier, examine and test a wide range of possible applications, and to produce a number of transistors to achieve these purposes. Later the development group began to concern itself with problems of cost.

But the research group that created the transistor did not participate in this work. The history of the solid-state research group after the transistor experiments is a good illustration of the evolutionary allocation mechanism which has been described. In 1951, with the advances achieved in junction transistors, the solid-state physics group began to grow rapidly, and in addition to the work on the junction transistor and underlying physics, work in other areas was intensified. As germanium and silicon were better understood, research in that area ran into diminishing returns and research on more complex semiconductors grew more promising. Similarly, research on magnetics and dielectrics was increased. In 1952 a separate transistor physics department was organized, headed by Shockley. Semiconductors research, once a subsection headed by Shockley under the solid-state physics group of Shockley and Morgan, which in turn was under physical research, is now a separate major department.

Concluding Thoughts

There is a limit to what can be learned from one case study. Generalizations are hazardous. However, it seems worthwhile to speculate on several aspects of industrial research which recently have been the subject of much public interest.

“TEAMWORK” IN RESEARCH

Much of the recent literature on industrial research has stressed the “team” aspect of this activity.¹⁵ Many writers have expressed sharp discomfort with the idea of the research team, and have argued that industrial research which stresses teamwork is likely to be scientifically arid. As I have emphasized before, the transistor is just one case, but nonetheless it seems worthwhile to examine what teamwork meant in the case of the transistor.

First of all, we have seen that it meant interaction and mutual stimulation and help. Shared intense interest in the general field, ease of communication, differences in the viewpoints and experience of different scientists—these elements naturally call for interaction and make interaction fruitful for scientific advance. The purpose of bringing together the people doing work in solid-state physics was to achieve this end, and from the history of the project it seems that the close interaction of several people definitely contributed to the advance achieved. But several people outside the team also interacted in an important way. In particular, the metallurgists’ work in developing ways to produce pure crystals was both an important link at many places in the chain of research and an activity in large part stimulated by the solid-state research effort of the team.

Second, we have seen that teamwork in the case of the transistor did not mean a closely directed project with an assigned division of labor in the form of tasks and schedules for each of the team members. There were no closely defined goals shared by all members of the group. All were interested in understanding semiconductors but, at least at the start, all were not excited by amplifiers. The project was marked by flexibility—by the ability to shift directions and by the rather rapid focusing of attention by several people on problems and phenomena unearthed by others.

Third, we have seen that teamwork in the case of the transistor did mean a more concentrated effort than probably would have developed

¹⁵ “Team” is used here in the popular sense, not in the sense of the theory of teams.

had the individuals involved not been brought together in one group. Shockley's interest in amplifiers definitely tended to draw the research interests of the group toward his focus. The research work of the transistor team, despite its very loose formal structure, was definitely affected by the fact that the men were working very closely together. In the case of the transistor this pulling together of interests proved extremely fruitful. Yet in many cases a more diversified attack might prove rewarding.

One wonders whether an invention like the transistor which came about as a direct result of an advance in scientific theory ever can be a team effort, if invention refers to a basic idea, and team refers to a group of people whose work is closely coordinated and planned by a team leader. It appears that the type of coordination required in an organization emphasizing fundamental science and the practical devices that might be created from advances in fundamental science is achieved quite effectively without planned coordination of effort and central control. Indeed it seems unlikely that a more formal control structure would be flexible enough to achieve anything like the type of coordination that marked the work of the research group on semiconductors. No one could have predicted the course of new knowledge or planned ahead to assure the right type of help at the right time. The informality of the decision structure played a very important role in permitting speedy cooperative response to changing ideas and knowledge. Thus the transistor was a team invention, but not in the sense of the term which has grown fashionable in recent years.

THE ROLE OF THE LARGE CORPORATION

In contrast to the public dismay about "teamwork" in research, many people have argued that the research laboratories of the large corporation are the only possible source of such inventions as the transistor. To what extent is an invention like the transistor dependent upon a sponsoring organization like the Bell Telephone Laboratories? How did the size of the Laboratories and the size of the corporations owning the Laboratories affect the project? Could the transistor have been developed at a significantly different institution, say, a laboratory owned by a much smaller company? My feeling is that devices like the transistor, based on fundamental new scientific knowledge, can come from small industrial laboratories or from universities, but that a large industrial laboratory like Bell does have a comparative advantage in this business.

The ingredients which seem to have played such an important role in the success of the semiconductor research project include the close interaction of a group of top-flight scientists, a great deal of freedom in the course of research, and an extremely strong interest on the part of at least one member of the group in inventing a practical device. This is high-priced talent at work, and when the project was initiated no guarantees were given as to the profits which would result. And for about two years before the transistor was invented, the group was kept free from pressure to produce practical results. This is not the type of project a small industrial laboratory is likely to be able to afford.

On the other hand, a university with strong science departments would have the resources to support work like the semiconductor research project. During World War II, and for a time after the war, work was going on at Purdue University under Lark-Horowitz which might have led to the transistor. The idea that an amplifier could be constructed from semiconductors was quite widely held by people working in the field. But, if my interpretation of the Bell project is correct, device minded scientists are essential if a laboratory is to develop devices such as the transistor. Shockley's theory, on which rests the design of the junction transistor, was in large part motivated by the desire to design a transistor. Further, it seems likely that if the identical theory had been developed by someone not interested in amplification, it would have taken some time for another worker to see that an amplifier could be designed on the basis of the theory. The popular image of the university scientist is that of a pure scholar, not interested in practical devices. If all university scientists were so, universities could hardly be the source of inventions like the transistor. For better or for worse, many university scientists do not fit that image, but are quite interested in devices. It is my feeling that the university with strong science departments, including scientists with a major interest in devices—not the laboratory of the small industrial corporation—is the major alternative source to the large industrial laboratory for inventions like the transistor.

ECONOMIC FACTORS AFFECTING THE DIRECTION OF
SCIENTIFIC RESEARCH

It has often been argued that the rate and direction of advance in pure science must be considered as an autonomous factor in any theory of the factors affecting the rate and direction of inventive activity. And it is likely that any analysis describing the motives of

the scientists who played the key roles in advancing quantum mechanics (Bohr, de Broglie, Schrodinger, Heisenberg, Pauli, Dirac, and others) as involving, in an important way, prospects for advancing practical technology, would be farfetched. But our case study has shown that the laboratories of large corporations may be effective institutions for drawing scientific research to areas where the importance of practical advances is great. And recall that the science of thermodynamics was, in large part, called forth by the development of steam engines, not vice versa. Advances in technology itself certainly affect the direction of scientific research.

The tremendous increase in the number of students taking undergraduate and graduate training in solid-state physics clearly has been strongly influenced by improved income prospects resulting from the increased use of solid-state physicists in industry and government sponsored research and development. This is not the whole explanation. Solid-state physics is much more scientifically fashionable now than it was. But even what is "fashionable" is strongly affected by applications. Since the birth of the transistor the proportion of articles relating to solid-state physics has significantly increased, and many scientists believe that the correlated change in the allocation of research effort was, in considerable part, due to the invention of the transistor and the consequent spotlighting of the field. Of course the same statements could be made about nuclear physics and the bomb. The direction of scientific advance is not independent of economic and practical factors.

WHAT IS BASIC RESEARCH ?

One of the most important things which can be learned from the history of the transistor is that the distinction between basic research and applied research is fuzzy. In the transistor project the results included both an advance in fundamental physical knowledge and the invention and improvement of practical devices. The scientists involved, though many of them were not interested in devices, were able to predict roughly the nature of the practical advances; indeed in some instances they were able to predict quite closely. And several of the scientists were motivated by the hope both of scientific advance and practical advance. Thus the project was marked by duality of results, and of motives. Yet by the standards of the National Science Foundation the Bell semiconductor research work most certainly would be considered basic research.

I have a feeling that duality of interests and results is far from unusual. I wonder how many scientists—university scientists—doing basic research do not think now and then about the possible practical applications of their work. I wonder how many are completely uninfluenced in their choice of research by consideration of possible practical benefit to mankind? By posing these questions I am not implying that the answer to both is “none.” But the answer may be “only a small proportion.”

I have the feeling that many scientists in industrial research laboratories, including—perhaps, particularly including—men with considerable executive authority, are defensive and internally torn about the dual nature of the research work. Unlike the management of most other industrial laboratories, the administration at the Bell Telephone Laboratories is strongly imbued with a belief in the value of fundamental research to the telephone company and with the understanding that people doing fundamental research, to be effective, must be kept free from day-to-day practical problems. Since the telephone laboratories were among the first industrial laboratories to undertake fundamental research, and since Bell's record has been so outstanding, the organization likes to toot its horn now and then. People in the Laboratories like to stress the extent to which much of their research is fundamental, and how fundamental research yields big payoffs. When it is suggested that much of the fundamental research is not conducted with “pure” motives, there is a tendency to get defensive, for the scientific community has long been accustomed to separate fundamental research from applied research on the basis of purity of motive. If research is conducted solely to advance man's knowledge it is fundamental; if it is conducted to help achieve a practical objective, it is applied, and somehow less intellectually respectable. And many people at Bell who are defenders of the fundamental research program are cut on this intellectual saw.

Shockley, however, suffers from no such intellectual split. In his Nobel lecture he states, “Frequently, I have been asked if an experiment I have planned is pure or applied science; to me it is more important to know if the experiment will yield new and probably enduring knowledge about nature. If it is likely to yield such knowledge, it is, in my opinion, good fundamental research; and this is more important than whether the motivation is purely esthetic satisfaction on the part of the experimenter on the one hand or the improvement of the stability of a high-power transistor on the other.”

Much of the history of science bears out Shockley's view. Some of the greatest scientific advances resulted from the research of men who were much concerned with the practical implications of their work, and their work was in large part motivated by the desire to benefit mankind.

The fuzzy nature of the boundary between basic and applied research does not imply, however, that some projects are not clearly more basic than some others. Nor does it imply that there are no important distinctions between the industrial research laboratory and the university laboratory. Basic research undertaken at industrial laboratories must somehow be related to expected company profit. And many of the greatest scientists have taken pride in the "purity" of their research motives and in the impracticality of fundamental science. Much of the foundation of modern physics was laid by men who chose science as a career in part because it offered an escape from the materialistic world. Although these men were often able to predict roughly the practical implications of their research, such practical aspects were unimportant, or even repugnant, to them. For these men the industrial research laboratory would seem a poor home. An industrial laboratory is not and probably should not be the equivalent of a university science department. It is risky to speculate how much of the "pure" research conducted at universities is "purer" than the "pure" research conducted at the Telephone Laboratories, but I believe that Einstein would have been unhappy at Bell, and Bell would be most unhappy if most of their research scientists were as intellectually pure as Einstein. It is probably best for all that this is so.

