WILSON

An Introduction to Scientific Research

AN INTRODUCTION TO SCIENTIFIC RESEARCH

E. BRIGHT WILSON, JR.

Theodore William Richards Professor of Chemistry, Harvard University

McGRAW-HILL BOOK COMPANY, INC.

NEW YORK · TORONTO · LONDON

1952

AN INTRODUCTION TO SCIENTIFIC RESEARCH

Library of Congress Catalog Card Number: 52-7448

VI

PREFACE

Few readers of this preface will need to be convinced of the enormous importance of scientific research in our present-day civilization or of the magnitude of the effort which is being expended in this activity. Anyone who has tried to do research also knows that it is in general a highly inefficient endeavor. An exploration into the unknown cannot be planned in advance with the precision of a mass-production process. Nevertheless, some investigators are far more effective than others and make fewer wrong decisions at the innumerable crossroads which are reached daily during the course of a typical research problem. We have no way of acquiring the inborn wisdom which is mostly responsible for their success, but perhaps there are a few techniques which we can learn from them.

This book is an attempt to collect in one place and to explain as simply as possible a number of general principles, techniques, and guides for procedure which successful investigators in various fields of science have found helpful. The emphasis is entirely on the practical rather than the philosophical or psychological aspects. Topics have been included only if they appeared to be useful to working scientists in more than one field. As a consequence the coverage is necessarily broad rather than deep.

Naturally a physical chemist cannot claim to be able to write a book equally useful in all the sciences. Nevertheless, many of the topics treated have been found useful by others in such diverse fields as agriculture, industrial and military research, biology, and medicine as well as in the physical sciences.

Much of the material should be understandable to a college senior, but the book is more specifically intended for students beginning research and for those more experienced research workers who wish an introduction to various topics which were not included in their training. The mathematical treatments have been kept as elementary as possible but are given where they seemed required.

In carrying out these objectives, I have acted simply as a collector of ideas from many areas, in most of which I claim no expertness. I have tried to present these ideas from the viewpoint of a practicing scientist and to illustrate them with as many actual examples as possible. Many of these are examples of the dire consequences of ignoring the maxims herein set down. In this field I do claim a certain authority; many of the blunders were original contributions of my own. I hope the reader will excuse an occasional statement which may strike him as too pontifical; it is only the fervor of the recent convert who has learned some things the hard way.

Naturally such a book as this could not have been written without a great deal of help from others, in the form of suggestions, criticism, and just plain instruction. I cannot begin to acknowledge here the aid of all those who have helped. There were, however, many who spent hours educating me in various areas and correcting my mistakes. These include D. J. Finney, P. A. P. Moran, M. G. Kendall, J. W. Tukey, and F. Mosteller. Most of the book was written while I was a Guggenheim Fellow and Fulbright grantee at Oxford, and I should like to acknowledge the indispensable assistance I received from the Guggenheim Foundation and the U.S. State Department. Moreover I am indebted to the chemists at Oxford for their very notable hospitality, particularly to Dr. and Mrs. J. W. Linnett.

Individual chapters have been read and criticized by some of the above and by Prof. G. B. Kistiakowsky, Dr. R. H. Hughes, Dr. R. M. Fristrom, and Dr. D. Eggers. The help of these and a number of other individuals is gratefully acknowledged, but of course they are in no way responsible for any errors of fact or judgment which the book may contain.

I am indebted to Prof. Ronald A. Fisher, Cambridge, to Dr. Frank Yates, Rothamsted, and to Messrs. Oliver & Boyd, Ltd., Edinburgh and London, for permission to reprint Table 9.2 from their book *Statistical Tables for Biological, Agricultural, and Medical Research.* I also wish to thank the other authors and publishers who have granted permission for the use of various material.

Finally I should like to acknowledge the considerable importance of the suggestions, criticism, and sympathy received from my wife, Emily Buckingham Wilson.

E. Bright Wilson, Jr.

Cambridge, Mass.
August, 1952

INTRODUCTION

This book is intended to assist scientists in planning and carrying out research. In a sense, therefore, it deals with scientific method, but not from the usual philosophical viewpoint. Rather, the aim has been the practical one of gathering together a number of principles, maxims, procedures, and general techniques which have been found useful in a range of sciences. The principle of selection has been to include only topics which would help someone decide what to do next and which are of a broad nature not too specific to a particular science.

Scientific work, by its very nature, cannot be reduced to a routine process, but this makes it certain that there is much room for improvement in efficiency. Furthermore, skilled and experienced workers usually learn many procedures only after years of actual practice. No book can completely replace experience, but much knowledge gained this hard way can be transferred to others via the printed page.

The topics have been arranged more or less in the order in which they arise during the course of an investigation, starting with the choice of a problem and ending with the publication of the results. The different chapters and, to a considerable extent, the different sections of each chapter have been made as nearly independent of one another as seemed feasible so that the reader is encouraged to pick and choose the items he most needs. Where necessary, cross references have been provided. Some readers may find sections which are too difficult or too detailed for their immediate needs. The relative independence of the sections should permit skipping to later topics.

Since a considerable number of subjects have been included, many of them are not covered at all thoroughly. In fact, in nearly every section it would be correct to include the statement: "This section is designed to introduce the reader to a topic about which whole books have been written." Therefore, at the end of each chapter general references have been provided which should permit these topics to be followed up further. The references are not meant to give the history of each subject; therefore, secondary rather than primary sources are often listed. The diligent reader should be able to go from these to the primary sources if he so desires. Certain material of a more detailed character has also been placed in the Notes at the ends of the chapters. In this way the text is kept free from footnotes and references.

CONTENTS

CHAPTER 7. CLASSIFICATION, SAMPLING, AND MEASUREMENT

CHAPTER 8. THE ANALYSIS OF EXPERIMENTAL DATA

CHAPTER 9. ERRORS OF MEASUREMENT

CHAPTER 13. REPORTING THE RESULTS OF RESEARCH

CHAPTER 1

THE CHOICE AND STATEMENT OF A RESEARCH PROBLEM

Many scientists owe their greatness not to their skill in solving problems but to their wisdom in choosing them. It is therefore worth considering the points on which this choice can be based.

1.1. Problems in Pure Science

It is hard to justify the choice of a problem in the field of pure research. Why should one choice be better than another?

One of the most important criteria is this: it should interest the investigator strongly. Scientific research, not being a routine process but requiring originality and creative thought, is very sensitive to the psychological state of the scientist. An uninterested worker is unlikely to produce the new ideas necessary for progress. One famous scientist has expressed this idea by saying that the problem should be important in the larger picture of one's view of the world.

Usually it is desirable to have new ideas of some sort before undertaking a problem, especially in a field which has been extensively investigated before. It is true that very simple and apparently obvious solutions have eluded experienced investigators and then been discovered by a new worker much later. However, it is much more often the case that an old problem is solved because some new tool, experimental or theoretical, has become available from another source. For example, the field of microwave spectroscopy has always been an attractive one, but until the invention of magnetron and klystron oscillators, it could not be exploited.

It is reasonable to ask what connection a given problem has with other branches of science. One problem may be important because it leads somewhere, while another may be trivial because it is a dead end.

On the other hand, it is almost always worth while to explore a region which is really new. Unexpected results can generally be relied upon under these circumstances. The synthesis of one more straight-chain hydrocarbon may be of doubtful value, in the absence of some particular purpose, but the discovery of a new class of compounds is likely to have repercussions in many directions.

It needs to be borne in mind that nature is far too vast to hope to chart its expanse in complete detail. It is therefore important that every task undertaken should be selected because it is likely to tell something about a wide area, rather than merely the immediate neighborhood. It is very easy, for example, to choose chemical compounds or biological species for study because they are available or experimentally convenient. Naturally both these practical considerations have to be kept in mind, but every effort should be made to select substances which are significant or fit into a larger pattern of inquiry.

The most rewarding work is usually to explore a hitherto untouched field. These are not easy to find today. However, every once in a while some new theory or new experimental method or apparatus makes it possible to enter a new domain. Sometimes it is obvious to all that this opportunity has arisen, but in other cases recognition of the opportunity requires more imagination.

When it is not a question of preliminary exploration of virgin territory, it is usually best to undertake experiments which are designed to test well-thought-out hypotheses. Experiments for experiment's sake are much less likely to lead anywhere. The results are often not useful later because, when a new hypothesis arises, its test may require data taken under somewhat different conditions.

Far too often projects are undertaken solely as a matter of experimental convenience. It is true that a new technique should first be applied to those situations which are experimentally the simplest, but as soon as possible, topics should be chosen because of their larger significance and because they fit into a pattern leading to a better understanding of the whole subject.

Another question which is worth asking before undertaking a new problem is: "Why should I, among all the scientists of the world, be the one to do this job?" There are many possible answers to this. Your experience may be just right, either experimentally or theoretically. You may possess unique equipment or a group of colleagues especially well equipped to advise you. You may have an original new idea and satisfactory equipment and experience. The problem may interest you so strongly that you are willing to invest the time in mastering a new field and take the risk of not being able to contribute something new after you have mastered it. If none of these things is true, it is rather unlikely that anything very startling will ensue.

A research worker in pure science who does not have at all times more problems he would like to solve than he has time and means to investigate them probably is in the wrong business. He may be an excellent experimenter and may have all the qualities required for success in applied science, but he lacks qualities of mind important for pure science. This is not at all to imply that applied science is easier, less demanding, or in any way inferior to pure science; it requires its own special abilities, but they are somewhat different.

From time to time the proposal is put forward that pure science should be planned "by some master board of strategists" which would direct workers to those fields where gaps were thought to exist. The utter folly of this idea is apparent to anyone with the slightest knowledge of the history of science. How could any board have directed anyone to discover radio, or X rays, or penicillin when at the time no one even suspected that these things existed?

Abandoning a Problem. The scientist who gives up too easily is unlikely to reap any great harvest, but on the other hand it is also possible to be too tenacious. It is a wise man who knows when to abandon a research or a field of research. No one can ever exhaust any field completely, but there always comes a point where further work, with existing techniques and ideas, is relatively less profitable than the same effort turned in other directions. Perhaps even earlier there comes a time when the field had better be turned over to new blood. No one can be so obstructive of progress as the "expert" who has worked all his life on a single subject.

1.2. Problems in Applied Science

Statement of the Problem. Waste in applied science may originate in an imperfect statement of the problem. Sometimes the problem that is enunciated is really a spurious one, the observations which gave it birth being faulty. Sometimes the problem, although real, is trivial. In other cases the problem is such that, even if it were solved, the solution would not be utilized. A careful statement of the problem often brings to light these conditions. Frequently, a small amount of time spent restating the problem in different ways, redefining it, or expressing its limits, points the way to its solution.

In applied science problems are often assigned to research workers by higher authorities, but this does not absolve the scientist from responsibility for examining the statement of the problem with great care. In many cases this will lead to new ways of stating it and to further conferences concerning the exact aims of the proposed investigation.

It is very important for the investigators who are to carry out the actual research to know as much as possible about the background of the problem, how it arose, why it is important, and what will be done with the results. Unfortunately, the nature of a research director's job is such that it takes positive action on his part to prevent the growth of a state of mind quite contrary to the above principles. It is easy to develop the habit of making decisions about programs and handing these to subordinates as dictates, without passing on the information upon which the decisions were based. This is not merely bad for morale; it very frequently leads to foolish and useless undertakings which a closer meeting of minds would have avoided. This difficulty arises because the exacting and many-sided nature of a research director's job makes it very hard for him to behave otherwise. Thus the necessity is very great for strong measures to prevent a chasm from growing between the director and his staff.

Equally it is the duty of the research worker to attempt to gain a sympathetic understanding of the reasons for the initiation of a proposed problem. He should also appreciate that even with the best of intentions no one can ever pass on all the background for every decision so that something has to be left to a feeling of confidence in higher authority.

Secrecy. The greatest difficulties, and consequently the greatest losses in efficiency, occur when secrecy, either military or industrial, is involved. Science simply does not flourish under such circumstances. Admittedly secrecy is sometimes necessary, but less often than is usually believed. When it is required, very special efforts are called for in order to be sure that each worker knows what he is trying to do and why it is important to do it. It is especially necessary that good coordination be arranged so that information acquired by one group is efficiently passed on to others who need it.

During the Second World War a considerable number of laboratories separately spent much time developing transient measuring equipment with trigger circuits, sweep circuits, timing circuits, amplifiers, oscilloscopes, and cameras. These were used for such varied purposes as measuring the blast from bombs and studying the effect of DDT on cockroaches. Because of security restrictions, most of these groups knew nothing of the work of the others so that each had to make the same mistakes and go through the same time-consuming stages. In retrospect there would seem to have been no reason why instrument work of such a nature should have been classified as confidential. Limitation of classification to those items which specifically need it makes it easier to avoid carelessness in the handling of truly secret material.

It is the interdependence of apparently unrelated topics that is so baffling to the nonscientist—who often has to make decisions strongly affecting science. It is hard for him to understand that a research on monomolecular films on water, such as Langmuir's, can lead to improved equipment for showing the motion picture *Gone with the Wind.*

Fundamental Work. Another problem facing the research director is how to divide his resources between direct, *ad hoc* attacks on immediate problems and longer range fundamental studies. So often the immediate problems appear to be so very urgent that long-range work is disrupted. Everyone with wartime or industrial experience knows that in a large fraction of the cases the urgency disappears or is eclipsed by a newer emergency before the first problem is solved. This is clear proof that insufficient thought was given to the original statement of the problem.

There is one school which holds that every applied research laboratory should set aside perhaps 20 to 30 per cent of its resources for long-range fundamental work in the field with which it is concerned. The choice of this work should be left largely to the more experienced research workers themselves but should have the aim of gaining a fuller understanding of the field. Research directed toward a particular product or practical application is not fundamental investigation in the sense meant here, regardless of how long-range it may be.

Thus this view would support the idea that in the research laboratory of a steel company there should be a certain proportion of the staff carefully protected from other calls on their time and free to investigate such topics as the quantum theory of metals, the crystal structures of intermetallic compounds, chemical kinetics of reactions in the solid state, etc. The purpose of this freedom would not be philanthropy but a hardheaded realization that any basic knowledge pertaining to steel would almost surely be used later to solve practical and urgent problems in a much more rapid and satisfactory way than the usual empirical cut-and-try procedures which must be employed when understanding is lacking.

This philosophy has justified itself in many organizations and is not to be confused with the idea that some members of a laboratory should be permitted to do whatever pure science interests them personally, regardless of its connection with the company's business. This greater freedom can be useful in attracting high-grade men, it certainly adds to the store of human knowledge, it often has good publicity value, and it sometimes pays in the end through unexpected discoveries, but it is not as easy to justify to a profit-seeking board of directors as the policy of allowing freedom to acquire basic knowledge in the field of the company's product.

It is, however, almost beyond the strength of human nature to resist the temptation to take people off fundamental work "temporarily" to tackle immediate problems. For this reason some concerns have set up separate organizations purposely placed beyond the reach of these calls. It would almost seem that no other method can be relied upon.

Taking Stock. A large applied research laboratory should devote a definite part of its resources to reviewing its past accomplishments and failures and seeking the reasons for these. Have a disproportionate number of problems been undertaken and then abandoned? Have too many problems been carried on the books? Have problems been shifted around too much from one group to another? Have there been many problems which were successfully solved but the results never used? Have problems been solved but found later not to have been stated in the right form originally? Have sufficient new ideas arisen from the scientific staff itself? Are new techniques constantly being developed and absorbed? How do the methods being used compare with those of other laboratories in the same field? Is the staff able to increase its knowledge continuously so that certain types of problems have become demonstrably easier and easier to solve? These and many other questions should be raised and investigated at frequent intervals. Otherwise it is very easy to drift along with a very inefficient system of organization and management, since output is not easily measurable.

1.3. The Cost of Experiments

In designing a bridge, an engineer naturally chooses the most economical design which satisfies all the specifications, including the aesthetic requirements. In designing an elaborate experiment, questions of cost are all too frequently ignored completely. This is partly because of the great difficulty of making good estimates of the time required to carry out a given investigation, but it is also partly a traditional attitude that somehow science is above vulgar monetary considerations.

With the increasing cost of research, it becomes necessary to take economic factors into account, however difficult this may be. Certainly there is no excuse for doing a given job in an expensive way when it can be carried through equally effectively with less expenditure. It is much more difficult to decide whether a given project should be carried out at all, considering its probable cost. In applied research there sometimes exist fairly definite criteria, such as the possible monetary benefits of a successful research, coupled with a rough estimate of the chance of success.

In pure science no estimate of monetary value is usually available or in fact desirable. Here cost still enters in deciding between alternative problems. Naturally this is not the only factor, but it is certainly wrong to disregard it altogether.

Cost estimates should include not only direct expenditures for materials but also salaries and overhead, even if these are not directly charged to the project. Many scientists are quite unaware of the magnitude of overhead costs and are shocked when they see figures for them. Generally speaking, overhead includes rent, heat, electricity, etc., and administrative expenses of the laboratory. These costs usually amount to 40 to 100 per cent of the total direct salaries and wages.

A common source of waste is the misuse of the time of salaried personnel with scientific training. There are many jobs which less highly trained assistants could carry out equally well. One reason for this misuse of talent is the low salary scale so often paid scientists. When a Ph.D. is paid little more than a mechanic, there is no economic reason for not allowing the Ph.D. to do a mechanic's job, whereas from the larger