The *Art* of Doing SCIENCE and Engineering

Learning to Learn

Richard W. Hamming

The Art of Doing Science and Engineering

The Art of Doing Science and Engineering

Learning to Learn

Richard W.Hamming
U.S. Naval Postgraduate School
Monterey, California

GORDON AND BREACH SCIENCE PUBLISHERS
Australia • Canada • China • France • Germany • India •
Japan • Luxembourg • Malaysia • The Netherlands •
Russia • Singapore • Switzerland • Thailand •
United Kingdom

This edition published in the Taylor & Francis e-Library, 2005.

"To purchase your own copy of this or any of Taylor & Francis or Routledge's collection of thousands of eBooks please go to www.eBookstore.tandf.co.uk."

Amsteldijk 166
1st Floor
1079 LH Amsterdam
The Netherlands

British Library Cataloguing in Publication Data

Hamming, R.W. (Richard Wesley), 1915–
The art of doing science and engineering: learning to learn
1. Science 2. Engineering
I. Title
500

ISBN 0-203-45071-X Master e-book ISBN

ISBN 0-203-45913-X (Adobe eReader Format)
ISBN 90-5699-501-4 (Print Edition)

INTRODUCTION

This book is concerned more with the future and less with the past of science and engineering. Of course future predictions are uncertain and usually based on the past; but the past is also much more uncertain—or even falsely reported—than is usually recognized. Thus we are forced to *imagine* what the future will probably be. This course has been called"Hamming on Hamming" since it draws heavily on my own past experiences, observations, and wide reading.

There is a great deal of mathematics in the early part because almost surely the future of science and engineering will be more mathematical than the past, and also I need to establish the nature of the foundations of our beliefs and their uncertainties. Only then can I show the weaknesses of our current beliefs and indicate future directions to be considered.

If you find the mathematics difficult, skip those early parts. Later sections will be understandable provided you are willing to forgo the deep insights mathematics gives into the weaknesses of our current beliefs. General results are always stated in words, so the content will still be there but in a slightly diluted form.

1
Orientation

The purpose of this course is to prepare you for your technical future. There is really no technical content in the course, though I will, of course, refer to a great deal of it, and hopefully it will generally be a good review of the fundamentals you have learned. Do not think the technical content is the course—it is only illustrative material. Style of thinking is the center of the course. I am concerned with educating and not training you.

I will examine, criticize, and display styles of thinking. To illustrate the points of style I will often use technical knowledge most of you know, but, again, it will be, I hope, in the form of a useful review which concentrates on the fundamentals. You should regard this as a course which complements the many technical courses you have learned. Many of the things I will talk about are things which I believe you ought to know but which simply do not fit into courses in the standard curriculum. The course exists because the department of Electrical and Computer Engineering of the Naval Postgraduate School recognizes the need for both a general education and the specialized technical training your future demands.

The course is concerned with "style", and almost by definition style cannot be taught in the normal manner by using words. I can only approach the topic through particular examples, which I hope are well within your grasp, though the examples come mainly from my 30 years in the mathematics department of the Research Division of Bell Telephone Laboratories (before it was broken up). It also comes from years of study of the work of others.

The belief anything can be "talked about" in words was certainly held by the early Greek philosophers, Socrates (469–399), Plato (427–347), and Aristotle (384–322). This attitude ignored the current *mystery cults* of the time who asserted you had to "experience" some things which could not be communicated in words. Examples might be the gods, truth, justice, the arts, beauty, and love. Your scientific training has emphasized the role of words, along with a strong belief in *reductionism,* hence to emphasize the possible limitations of language I shall take up the topic in several places in the book. I have already said "style" is such a topic.

I have found to be effective in this course, I must use mainly first hand knowledge, which implies I break a standard taboo and talk about myself in the first person, instead of the traditional impersonal way of science. You must forgive me in this matter, as there seems to be no other approach which will be as effective. If I do not use direct experience then the material will probably sound to you like merely pious words and have little impact on your minds, and it is your minds I must change if I am to be effective.

This talking about first person experiences will give a flavor of "bragging", though I include a number of my serious errors to partially balance things. Vicarious learning from the experiences of others saves making errors yourself, but I regard the study of successes as being basically more important than the study of failures. As I will several times say, there are so many ways of being wrong and so few of being right, studying successes is more efficient, and furthermore when your turn comes you will know how to succeed rather than how to fail!

I am, as it were, only a coach. I cannot run the mile for you; at best I can discuss styles and criticize yours. You know you must run the mile if the athletics course is to be of benefit to you—hence *you* must think carefully about what you hear or read in this book if it is to be effective in changing you—which must obviously be the purpose of any course. Again, you will get out of this course only as much as you put in, and if you put in little effort beyond sitting in the class or reading the book, then it is simply a waste of your time. *You* must also mull things over, compare what I say with your own experiences, talk with others, and make some of the points part of your way of doing things.

Since the subject matter is “style”, I will use the comparison with teaching painting. Having learned the fundamentals of painting, you then study under a master you accept as being a great painter; but you know you must forge your own style out of the elements of various earlier painters plus your native abilities. You must also adapt your style to fit the future, since merely copying the past will not be enough if you aspire to future greatness—a matter I assume, and will talk about often in the book. I will show you my style as best I can, but, again, you must take those elements of it which seem to fit you, and you must finally create your own style. Either you will be a leader, or a follower, and my goal is for you to be a leader. You cannot adopt every trait I discuss in what I have observed in myself and others; you must select and adapt, and make them your own if the course is to be effective.

Even more difficult than what to select is that what is a successful style in one age may not be appropriate to the next age! My predecessors at Bell Telephone Laboratories used one style; four of us who came in all at about the same time, and had about the same chronological age, found our own styles and as a result we rather completely transformed the overall style of the Mathematics Department, as well as many parts of the whole Laboratories. We privately called ourselves “The four young Turks”, and many years later I found top management had called us the same!

I return to the topic of education. You all recognize there is a significant difference between *education* and *training*.

> Education is what, when, and why to do things, Training is how to do it.

Either one without the other is not of much use. You need to know both what to do and how to do it. I have already compared mental and physical training and said to a great extent in both you get out of it what you put into it—all the coach can do is suggest styles and criticize a bit now and then. Because of the usual size of these classes, or because you are reading the book, there can be little direct criticism of your thinking by me, and you simply have to do it internally and between yourselves in conversations, and apply the things I say to your own experiences. You might think education should precede training, but the kind of educating I am trying to do must be based on your past experiences and technical knowledge. Hence this inversion of what might seem to be reasonable. In a real sense I am engaged in “meta-education”, the topic of the course is education itself and hence our discussions must rise above it—“meta-education”, just as metaphysics was supposed to be above physics in Aristotle’s time (actually “follow”, “transcend” is the translation of “meta”).

This book is aimed at your future, and we must examine what is likely to be the state of technology (Science and Engineering) at the time of your greatest contributions. It is well known that since about Isaac Newton’s time (1642–1727) knowledge of the type we are concerned with has about doubled every 17 years. First, this may be measured by the books published (a classic observation is libraries must double their holdings every 17 years if they are to maintain their relative position). Second, when I went to Bell

Telephone Laboratories in 1946 they were trying to decrease the size of the staff from WW-II size down to about 5500. Yet during the 30 years I was there I observed a fairly steady doubling of the number of employees every 17 years, regardless of the administration having hiring freezes now and then, and such things. Third, the growth of the number of scientists generally has similarly been exponential, and it is said currently almost 90% of the scientists who ever lived are now alive! It is hard to believe in your future there will be a dramatic decrease in these expected rates of growth, hence you face, even more than I did, the constant need to learn new things.

Here I make a digression to illustrate what is often called "back of the envelop calculations". I have frequently observed great scientists and engineers do this much more often than "the run of the mill" people, hence it requires illustration. I will take the above two statements, knowledge doubles every 17 years, and 90% of the scientists who ever lived are now alive, and ask to what extent they are compatible. The model of the growth of knowledge and the growth of scientists assumed are both exponential, with the growth of knowledge being proportional to the number of scientists alive. We begin by assuming the number scientists at any time t is

$$y(t) = a\exp\{bt\}$$

and the amount of knowledge produced annually has a constant k of proportionality to the number of scientists alive. Assuming we begin at minus infinity in time (the error is small and you can adjust it to Newton's time if you wish), we have the formula

$$\frac{1}{2} = \frac{\int_{-\infty}^{t-17} kae^{bt}\,\mathrm{d}t}{\int_{-\infty}^{t} kae^{bt}\,\mathrm{d}t}$$

$$= \frac{(ka/b)\mathrm{e}^{b(t-17)}}{(ka/b)\mathrm{e}^{bt}} = \mathrm{e}^{-17b} = \frac{1}{2}$$

hence we know b. Now to the other statement. If we allow the lifetime of a scientist to be 55 years (it seems likely that the statement meant living and not practicing, but excluding childhood) then we have

$$\frac{\int_{t-55}^{t} ae^{bt}\,\mathrm{d}t}{\int_{-\infty}^{t} ae^{bt}\,\mathrm{d}t} = \frac{\mathrm{e}^{bt} - \mathrm{e}^{\{bt-b(55)\}}}{\mathrm{e}^{bt}} = 1 - \mathrm{e}^{-55b}$$

$$= 1 - (\tfrac{1}{2})^{55/17} = 1 - 0.106\ldots = 0.894\ldots$$

which is very close to 90%.

Typically the first back of the envelop calculations use, as we did, definite numbers where one has a feel for things, and then we repeat the calculations with parameters so you can adjust things to fit the data better and understand the general case. Let the doubling period be D, and the lifetime of a scientist be L. The first equation now becomes

$$\tfrac{1}{2} = \mathrm{e}^{bD},$$

and the second becomes:

$$\tfrac{9}{10} = 1 - \mathrm{e}^{bL} = 1 - (\tfrac{1}{2})^{L/D},$$

$$(\tfrac{1}{2})^{L/D} = \tfrac{1}{10},$$

$$\frac{L}{D} = \frac{\log 10}{\log 2} = \frac{1}{0.30103} = 3.3219\ldots$$

With D=17 years we have 17×3.3219=56.47…years for the lifetime of a scientist, which is close to the 55 we assumed. We can play with ratio of L/D until we find a slightly closer fit to the data (which was approximate, though I believe more in the 17 years for doubling than I do in the 90%). Back of the envelop computing indicates the two remarks are reasonably compatible. Notice the relationship applies for all time so long as the assumed simple relationships hold.

The reason back of the envelop calculations are widely used by great scientists is clearly revealed—you get a good feeling for the truth or falsity of what was claimed, as well as realize which factors you were inclined not to think about, such as exactly what was meant by the lifetime of a scientist. Having done the calculation you are much more likely to retain the results in your mind. Furthermore, such calculations keep the ability to model situations fresh and ready for more important applications as they arise. Thus I recommend when you hear quantitative remarks such as the above you turn to a quick modeling to see if you believe what is being said, especially when given in the public media like the press and TV. Very often you find what is being said is nonsense, either no definite statement is made which you can model, or if you can set up the model then the results of the model do not agree with what was said. I found it very valuable at the physics table I used to eat with; I sometimes cleared up misconceptions at the time they were being formed, thus advancing matters significantly.

Added to the problem of the growth of new knowledge is the obsolescence of old knowledge. It is claimed by many the half-life of the technical knowledge you just learned in school is about 15 years—in 15 years half of it will be obsolete (either we have gone in other directions or have replaced it with new material). For example, having taught myself a bit about vacuum tubes (because at Bell Telephone Laboratories they were at that time obviously important) I soon found myself helping, in the form of computing, the development of transistors—which obsoleted my just learned knowledge!

To bring the meaning of this doubling down to your own life, suppose you have a child when you are x years old. That child will face, when it is in college, about y times the amount you faced.

y factor of increase	x years
2	17
3	27
4	34
5	39
6	44
7	48
8	51

This doubling is not just in theorems of mathematics and technical results, but in musical recordings of Beethoven's Ninth, of where to go skiing, of TV programs to watch or not to watch. If you were at times awed by the mass of knowledge you faced when you went to college, or even now, think of your children's troubles when they are there! The technical knowledge involved in your life will quadruple in 34 years, and many of you will then be near the high point of your career. Pick your estimated years to retirement and then look in the left-hand column for the probable factor of increase over the present current knowledge when you finally quit!

What is my answer to this dilemma? One answer is you must concentrate on fundamentals, at least what *you think* at the time are fundamentals, and also develop the ability to learn new fields of knowledge when they arise so you will not be left behind, as so many good engineers are in the long run. In the position I found myself in at the Laboratories, where I was the only one locally who seemed (at least to me) to have a firm grasp on computing, I was forced to learn numerical analysis, computers, pretty much all of the physical sciences at least enough to cope with the many different computing problems which arose and whose solution could benefit the Labs, as well as a lot of the social and some the biological sciences. Thus I am a veteran of learning enough to get along without at the same time devoting all my effort to learning new topics and thereby not contributing my share to the total effort of the organization. The early days of learning had to be done while I was developing and running a computing center. You will face similar problems in your career as it progresses, and, at times, face problems which seem to overwhelm you.

How are you to recognize "fundamentals"? One test is they have lasted a long time. Another test is from the fundamentals all the rest of the field can be derived by using the standard methods in the field.

I need to discuss science vs. engineering. Put glibly:

> In science if you know what you are doing you should not be doing it.
> In engineering if you do not know what you are doing you should not be doing it.

Of course, you seldom, if ever, see either pure state. All of engineering involves some creativity to cover the parts not known, and almost all of science includes some practical engineering to translate the abstractions into practice. Much of present science rests on engineering tools, and as time goes on, engineering seems to involve more and more of the science part. Many of the large scientific projects involve very serious engineering problems—the two fields are growing together! Among other reasons for this situation is almost surely we are going forward at an accelerated pace, and now there is not time to allow us the leisure which comes from separating the two fields. Furthermore, both the science and the engineering you will need for your future will more and more often be created after you left school. Sorry! But you will simply have to actively master *on your own* the many new emerging fields as they arise, without having the luxury of being passively taught.

It should be noted that engineering is not just applied science, which is a distinct third field (though it is not often recognized as such) which lies between science and engineering.

I read somewhere there are 76 different methods of predicting the future—but very number suggests there is no reliable method which is widely accepted. The most trivial method is to predict tomorrow will be exactly the same as today—which at times is a good bet. The next level of sophistication is to use the current rates of change and to suppose they will stay the same—linear prediction in the variable used. Which variable you use can, of course, strongly affect the prediction made! Both methods are not much good for long-term predictions, however.

History is often used as a long-term guide; some people believe history repeats itself and others believe exactly the opposite! It is obvious:

> The past was once the future and the future will become the past.

In any case I will often use history as a background for the extrapolations I make. I believe the best predictions are based on understanding the fundamental forces involved, and this is what I depend on mainly. Often it is not physical limitations which control but rather it is human made laws, habits, and organizational rules, regulations, personal egos, and inertia, which dominate the evolution to the future. You have not been trained along these lines as much as I believe you should have been, and hence I must be careful to include them whenever the topics arise.

There is a saying,"Short term predictions are always optimistic and long term predictions are always pessimistic". The reason, so it is claimed, the second part is true is for most people the geometric growth due to the compounding of knowledge is hard to grasp. For example for money a mere 6% annual growth doubles the money in about 12 years! In 48 years the growth is a factor of 16. An example of the truth of this claim that most long-term predictions are low is the growth of the computer field in speed, in density of components, in drop in price, etc. as well as the spread of computers into the many corners of life. But the field of Artificial Intelligence (AI) provides a very good counter example. Almost all the leaders in the field made long-term predictions which have almost never come true, and are not likely to do so within your lifetime, though many will in the fullness of time.

I shall use history as a guide many times in spite of Henry Ford, Sr. saying, "History is Bunk". Probably Ford's points were:

1. History is seldom reported at all accurately, and I have found no two reports of what happened at Los Alamos during WW-II seems to agree.
2. Due to the pace of progress the future is rather disconnected from the past; the presence of the modern computer is an example of the great differences which have arisen.

Reading some historians you get the impression the past was determined by big trends, but you also have the feeling the future has great possibilities. You can handle this apparent contradiction in at least four ways:

1. You can simply ignore it.
2. You can admit it.
3. You can decide the past was a lot less determined than historians usually indicate and individual choices can make large differences at times. Alexander the Great, Napoleon, and Hitler had great effects on the physical side of life, while Pythagoras, Plato, Aristotle, Newton, Maxwell, and Einstein are examples on the mental side.
4. You can decide the future is less open ended than you would like to believe, and there is really less choice than there appears to be.

It is probable the future will be more limited by the slow evolution of the human animal and the corresponding human laws, social institution, and organizations than it will be by the rapid evolution of technology.

In spite of the difficulty of predicting the future and that:

> Unforeseen technological inventions can completely upset the most careful predictions,

you must try to foresee the future you will face. To illustrate the importance of this point of trying to foresee the future I often use a standard story.

It is well known the drunken sailor who staggers to the left or right with n independent random steps will, on the average, end up about $\sqrt{n}$ steps from the origin. But if there is a pretty girl in one direction, then his steps will tend to go in that direction and he will go a distance proportional to n. In a lifetime of many, many independent choices, small and large, a career with a vision will get you a distance proportional to n, while no vision will get you only the distance $\sqrt{n}$. In a sense, the main difference between those who go far and those who do not is some people have a vision and the others do not and therefore can only react to the current events as they happen.

One of the main tasks of this course is to start you on the path of creating in some detail *your vision of your future*. If I fail in this I fail in the whole course. You will probably object that if you try to get a vision now it is likely to be wrong—and my reply is from observation I have seen the accuracy of the vision matters less than you might suppose, getting anywhere is better than drifting, there are potentially many paths to greatness for you, and just which path you go on, *so long as it takes you to greatness,* is none of my business. You must, as in the case of forging your personal style, find your vision of your future career, and then follow it as best you can.

No vision, not much of a future.

To what extent history does or does not repeat itself is a moot question. But it is one of the few guides you have, hence history will often play a large role in my discussions—I am trying to provide you with some perspective as a possible guide to create your vision of your future. The other main tool I have used is an active imagination in trying to see what will happen. For many years I devoted about 10% of my time (Friday afternoons) to trying to understand what would happen in the future of computing, both as a scientific tool and as shaper of the social world of work and play. In forming your plan for your future you need to distinguish three different questions:

What is possible?
What is likely to happen?
What is desirable to have happen?

In a sense the first is Science—what is possible. The second in Engineering—what are the human factors which chose the one future that does happen from the ensemble of all possible futures. The third, is ethics, morals, or what ever other word you wish to apply to value judgments. It is important to examine all three questions, and in so far as the second differs from the third, you will probably have an idea of how to alter things to make the more desirable future occur, rather than let the inevitable happen and suffer the consequences. Again, you can see why having a vision is what tends to separate the leaders from the followers.

The standard process of organizing knowledge by departments, and subdepartments, and further breaking it up into separate courses, tends to conceal the homogeneity of knowledge, and at the same time to omit much which falls between the courses. The optimization of the individual courses in turn means a lot of important things in Engineering practice are skipped since they do not appear to be essential to any one course. One of the functions of this book is to mention and illustrate many of these missed topics which are important in the practice of Science and Engineering. Another goal of the course is to show the essential unity of all knowledge rather than the fragments which appear as the individual topics are taught. In your future anything and everything you know might be useful, but if you believe the problem is in one area you are not apt to use information that is relevant but which occurred in another course.

The course will center around computers. It is not merely because I spent much of my career in Computer Science and Engineering, rather it seems to me computers will dominate your technical lives. I will repeat a number of times in the book the following facts: Computers when compared to Humans have the advantages:

Economics	—far cheaper, and getting more so
Speed	—far, far faster

Accuracy	—far more accurate (precise)
Reliability	—far ahead (many have error correction built into them)
Rapidity of control	—many current airplanes are unstable and require rapid computer control to make them practical
Freedom from boredom	—an overwhelming advantage
Bandwidth in and out	—again overwhelming
Ease of retraining	—change programs, not unlearn and then learn the new thing consuming hours and hours of human time and effort
Hostile environments	—outer space, underwater, high radiation fields, warfare, manufacturing situations that are unhealthful, etc.
Personnel problems	—they tend to dominate management of humans but not of machines; with machines there are no pensions, personal squabbles, unions, personal leave, egos, deaths of relatives, recreation, etc.

I need not list the advantages of humans over computers—almost every one of you has already objected to this list and has in your mind started to cite the advantages on the other side.

Lastly, in a sense, this is a religious course—I am preaching the message that, with apparently only one life to live on this earth, you ought to try to make significant contributions to humanity rather than just get along through life comfortably—that the life of trying to achieve excellence in some area is in itself a worthy goal for your life. It has often been observed the true gain is in the struggle and not in the achievement—a life without a struggle on your part to make yourself excellent is hardly a life worth living. This, it must be observed, is an opinion and not a fact, but it is based on observing many people's lives and speculating on their total happiness rather than the moment to moment pleasures they enjoyed. Again, this opinion of their happiness must be my own interpretation as no one can know another's life. Many reports by people who have written about the "good life" agree with the above opinion. Notice I leave it to you to pick your goals of excellence, but claim only a life without such a goal is not really living but it is merely existing—in my opinion. In ancient Greece Socrates (469–399) said:

> The unexamined life is not worth living.