J. HOOVER MACKIN
University of Texas

Rational and Empirical Methods of Investigation in Geology[1]

Most of us are concerned, and some of us have strong feelings, pro or con, about what has been happening to geology in the past 25 years: greatly increased use of nongeologic techniques in the solution of geologic problems, such as dating by radioisotope methods; the tendency for what were special fields of interest to become nearly or wholly independent disciplines, with separate journals and jargon; and most of all, because it penetrates every field, what may be called the swing to the quantitative.

At meetings of our societies, when the elder brethren gather together in hotel rooms after the technical sessions, the discussion usually comes around to these changes. There are apt to be sad postmortems for certain departments, once powerful, which are now, owing to the retirement or flight of their older stalwarts, largely staffed by dial twisters and number jugglers. It is stated, as a scandalous sign of the times, that in certain departments geologic mapping is considered to be, not research, but a routine operation—something like surveying from the point of view of an engineer—and therefore not suitable as a basis

[1] A preliminary draft of this paper was given as an address at the banquet of the Branner Club during the meeting of the Cordilleran Section of the Geological Society in Los Angeles, April 17, 1962. The text has benefitted in substance and form from criticisms by the other authors of papers in this volume. I would like also to express my gratitude to the following, who have read parts or all of the manuscript: Charles Bell, Richard Blank, Howard Coombs, Ronald DeFord, Ken Fahnestock, Peter Flawn, John Hack, Satish Kapoor, William Krumbein, Luna Leopold, Mark Meier, H. W. Naismith, and Dwight Schmidt. Special thanks are due Frank Calkins, who did his best to make the paper readable.

for the doctoral thesis. There is almost always at least one sarcastic remark per evening along the line of what our equation-minded youngsters think is the function of the mirror on a Brunton compass; a comment or two on their ignorance or disregard of the older literature; some skepticism as to whether the author of a new monograph on the mechanism of mountain building had ever been *on* a mountain, *off* a highway; and so on. This is partly banter, because we are aware that these are merely the usual misgivings of every older generation about the goings-on of every younger generation. But sometimes there is evidence of real ill-feeling, which in part at least reflects a defensive attitude; and there may be a few who seem to think that the clock ought to be stopped—that nothing new is good.

Though I am one of the elders, I often cross the hall to a concurrent session of another group, our avant-garde, where there is an almost evangelical zeal to quantify, and if this means abandoning the classical geologic methods of inquiry, so much the better; where there are some who think of W. M. Davis as an old duffer with a butterfly-catcher's sort of interest in scenery; where there is likely to be, once in a while, an expression of anger for the oldsters who, through their control of jobs, research funds, honors, and access to the journals, seem to be bent on sabotaging all efforts to raise geology to the stature of a science; where, in the urgency for change, it seems that nothing old is good.

This picture is not overdrawn, but it applies only to a small number: the blacks and the whites, both sure of their ground. Most geologists are somewhere in the gray between, and are beset with doubts. As for myself, I have sometimes thought that the swing to the quantitative is too fast and too far, and that, because a rather high percentage of the conclusions arrived at by certain methods of manipulating numerical data are superficial, or wrong, or even ludicrous, these methods must be somehow at fault, and that we do well to stay with the classical geologic methods. But at other times I have been troubled by questions: why the swing has been so long delayed in geology as compared with physics and chemistry; and whether, with its relative dearth of quantitative laws, geology is in fact a sort of subscience, as implied by Lord Kelvin's pronouncement that what cannot be stated in numbers is not science. (For original wording, and a thoughtful discussion, see Holton, 1952, p. 234.) Even more disturbing is the view, among some of my friends in physics, that a concern with cause-and-effect relations merely confuses the real issues in science; I will return to this matter later. If only because of the accomplishments of the scientists who hold these views, we must wonder whether our accustomed ways of thinking are outmoded, and whether we should not drastically change our habits of thought, or else turn in our compasses and hammers and fade away quietly to some haven reserved for elderly naturalists.

Preparation for a talk on quantitative methods in geomorphology, as a visiting lecturer at the University of Texas last year, forced me to examine

these conflicting appraisals of where we stand.[2] I suggest that two changes, quite different but closely interlocking, are occurring at the same time and have become confused in our thinking.

One of these changes includes an increase in the rate of infusion of new ideas and techniques from the other sciences and from engineering, an increase in precision and completeness of quantitative description of geologic features and processes of all kinds, and an increased use of statistics and mechanical methods of analyzing data. This change fits readily within the framework of the classical geologic method of investigation, the most characteristic feature of which is dependence on reasoning at every step; "Quantitative Zoology," by Simpson, Roe, and Lewontin (1960) shows the way. In so far as it merely involves doing more completely, or with more refinement, what we have always been doing, it is evolutionary; and it is axiomatic that it is good. Some of us may find it hard to keep abreast of new developments, but few oppose them even privately, and even the most reactionary cannot drag his feet in public without discredit to himself.

The other change is the introduction, or greatly increased use, of an altogether different method of problem-solving that is essentially empirical. In its purest form this method depends very little on reasoning; its most characteristic feature, when it functions as an independent method, is that it replaces the reasoning process by operations that are largely mechanical. Because in this respect and others it is foreign to our accustomed habits of thought, we are inclined to distrust it. By "we" I mean, of course, the conservatives of my generation.

At least a part of the confusion in our thinking comes from a failure to distinguish between the evolutionary quantification, which is good, and the mechanical kind of quantification, which I think is bad when it takes the place of reasoning. It is not easy to draw a line between them because the empirical procedures may stand alone, or they may function effectively and usefully as parts of the classical geologic method; that is, they may replace, or be combined in all proportions with, the reasoning processes that are the earmarks of that method. When this distinction is recognized it becomes evident that the real issue is not qualitative versus quantitative. It is, rather, rationality versus blind empiricism.

---

[2] I was only dimly aware, until some library browsing in connection with methodology in the other sciences, of the extent of the scholarly literature dealing with the history and philosophy of science. And I was surprised, as was Claude Albritton (1961), to find that with a few noteworthy exceptions (for example, Conant, 1951, p. 269–295) geology is scarcely mentioned in that literature. I should like to make it plain at the outset that I am not a scholar—I have only sampled a few anthologies of the history of science. I should emphasize also that I do not presume to speak for geology; what I say expresses the viewpoint of a single field geologist.

Although the timing has been influenced by such leaders as Chayes, Hubbert, Leopold, Krumbein, and Strahler, we are now in the swing to the quantitative because of the explosive increase in the availability of numerical data in the last few decades (Krumbein, 1960, p. 341), and because basic descriptive spadework has now advanced far enough in many fields of geology to permit at least preliminary formulation of significant quantitative generalizations. The quantification of geology will proceed at a rapidly accelerating rate no matter what we do as individuals, but I think the rate might be quickened a little, and to good purpose, if the differences between the two groups on opposite sides of the hall, at least those differences that arise from misunderstanding, could be reduced. An analysis of certain quantitative methods of investigation that are largely empirical will, I hope, serve to bring out both their merits and limitations, and may convince some of our oldsters that although disregard of the limitations may produce questionable results, it does not follow that there is anything wrong with quantification, as such, nor with blind empiricism, as such. But this is not very important—time will take care of the oldsters, soon enough. This essay is for the youngsters—the graduate students—and its purpose is to show that as they quantify, which they are bound to do, it is neither necessary nor wise to cut loose from the classical geologic method. Its message is the not very novel proposition that there is much good both in the old and the new approaches to problem-solving. A brief statement of what I am calling the rational method will point up the contrast between it and the empirical method, with which we are principally concerned.

The Rational Method

I'm sure that most American geologists are acquainted with our three outstanding papers on method: G. K. Gilbert's "Inculcation of the Scientific Method by Example," published in 1886; T. C. Chamberlin's "Method of Multiple Working Hypotheses," published in 1897; and Douglas Johnson's "Role of Analysis in Scientific Investigation," published in 1933. I do not need to describe the so-called scientific method here; for present purposes I need only remind you that it involves an interplay of observation and reasoning, in which the first observations suggest one or more explanations, the working hypotheses, analysis of which leads to further observation or experimentation. This in turn permits a discarding of some of the early hypotheses and a refinement of others, analysis of which permits a discarding of data now seen to be irrelevant to the issue, and a narrowing and sharpening of the focus in the search for additional data that are hidden or otherwise hard to obtain but which are of special diagnostic value; and so on and on. These steps are spelled out in formal terms in the papers just mentioned, and it was useful to do that, but those who use the method all the time never follow the steps in the order stated; the method has become a habit of thought that checks reasoning against other lines of reasoning, evidence against other kinds of evidence, reasoning against evidence, and evidence against reasoning, thus testing both the evidence and the reasoning for relevancy and accuracy at every stage of the inquiry.

It now seems to be the vogue to pooh-pooh this method, as differing in no essential way from the method of problem-solving used by the man in the street. I've been interested in watching the way in which men in the street, including some medical doctors—practitioners, not investigators—arrive at conclusions, and I can only suggest that the scientists who insist that all persons arrive at conclusions in the same way should reexamine their conviction. There are, of course, rare intellects that need no disciplining, but for most of us with ordinary minds, facility in the operations that I have just outlined must be acquired by precept, example, and practice.

The objective of the scientific method is to understand the system investigated—to understand it as completely as possible. To most geologists this means understanding of cause and effect relations within the system (Garrels, 1951, p. 32). Depending on the nature of the problem and its complexity, quantitative data and mathematical manipulations may enter the investigation early or late. In general, the larger the problem, the more many-sided it is, the more complicated by secondary and tertiary feedback couples, and the more difficult it is to obtain the evidence, the more essential it is to the efficient prosecution of the study that the system first be understood in *qualitative* terms; only this can make it possible to design the most significant experiments, or otherwise to direct the search for the critical data, on which to base an eventual understanding in quantitative terms.

A problem—any problem—when first recognized, is likely to be poorly defined. Because it is impossible to seek intelligently for explanations until we know what needs explaining, the first step in the operation of the scientific method is to bring the problem into focus. This is usually accomplished by reasoning, i.e., by thinking it through, although we will see shortly that there is another way. Then, if it is evident that the problem is many-sided, the investigator does not blast away at all sides at once with a shotgun; he shoots at one side at a time with a rifle—with *the* rifle, and *the* bullet, that he considers best suited to that side.

This means that the investigator admits to his graphs, so to speak, only items of evidence that are relevant to the particular matter under investigation, and that are as accurate as practicable, with the probable limits of sampling and experimental error expressed graphically. In reading answers from the graph, he does no averaging beyond that required to take those limits into account. And once an item of information has been admitted to the graph, it cannot be disregarded; as a rule, the items that lie outside the clusters of points are at least as significant, and usually much more interesting, than those that lie within the clusters. It is from inquiry as to why these strays are where they are that most new ideas—most breakthroughs in science—develop.

The scientific method tries to visualize whole answers—complete theoretical structures—at the very outset; these are the working hypotheses that give direction to the seeking-out and testing of evidence. But one never rushes ahead of the data-testing process to a generalization that is regarded as a conclusion. This is not because there is anything ethically wrong with quick generalizing. It is only that, over a period of 500 years, investigators have found that theoretical structures made in part of untested and ill-matched building blocks are apt to topple sooner or later, and that piling them up and building on them is therefore not an efficient way to make progress. The need to test the soundness of each building block *before* it gets into the structure—to determine the quality and the relevance of each item of evidence *before* it gets onto the graph—is emphasized by Douglas Johnson (1933). His approach was the antithesis of that to which we may now turn.

The Empirical Method

What I have long thought of as the engineering method or the technologic method (we shall soon see that it needs another name) deals almost exclusively with quantitative data from the outset, and proceeds directly to a quantitative answer, which terminates the investigation. This method reduces to a minimum, or eliminates altogether, the byplay of inductive and deductive reasoning by which data and ideas are processed in the scientific method; this means that it cannot be critical of the data as they are gathered. The data are analyzed primarily by mathematical methods, which make no distinction between cause and effect; understanding of cause and effect relations may be interesting, but it is not essential, and if explanations are considered at all, there is usually only one, and it is likely to be superficial. All of the reasoning operations that characterize the so-called scientific method depend on a fund of knowledge, and on judgment based on experience; other things being equal, the old hand is far better at these operations than the novice. But the operations of the "engineering method" are much less dependent on judgment; in applying this method the sharp youngster may be quicker and better than the experienced oldster. For this reason and because of its quick, positive, quantitative answers, it makes a strong appeal to the younger generation. I would like now to explain the logic of this method, as it operates in engineering.

Many engineers feel that unless a relation can be stated in numbers, it is not worth thinking about at all. The good and sufficient reason for this attitude is that the engineer is primarily a doer—he designs structures of various types, and supervises their building. In the contract drawings for a bridge he must specify the dimensions and strength of each structural member. Nonengineers may be able to think of a drawing that indicates the need for a rather strong beam at a given place in the bridge. But a young man who has spent five years in an engineering school is incapable of thinking seriously of a "rather strong" beam; all of the beams of his mind's eye have numerical properties. If the strength of a beam cannot be put in numerical terms, thinking about it is mere daydreaming.

The matter of stresses in a steel structure is fairly cut and dried. But the engineer is confronted with many problems for which there are no ready answers; he must deal with them—he must complete his working drawings—against a deadline. If he is charged with the task of designing a canal to carry a certain flow of irrigation water without either silting or erosion of the bed, or with the immensely more complex task of developing and maintaining a 10-foot navigable channel in a large river, he cannot wait until he or others have developed a complete theory of silting and scouring in canals and rivers. It may be 50 or a 100 years before anything approaching a complete theory, in quantitative terms, can be formulated; and his drawings, which must be entirely quantitative, have to be ready within a few days or weeks for the contractors who will bid the job. So he has to make certain simplifying assumptions, even though he realizes that they may be wide of the mark, and he has to make-do with data that are readily available, even though they are not entirely satisfactory, or with data that can be obtained quickly from experiments or models, even though the conditions are significantly different from those existing in his particular canal or river.

He is accustomed to these expedient operations, and he is not much concerned if, in plotting the data, he mixed a few oranges with the apples. In fact, he wouldn't worry much if a few apple *crates* and a few orange *trees* got onto his graph. He cannot scrutinize each item of evidence as to quality and relevancy; if he did, none but the simplest of structures would ever get built. He feels that if there are enough points on a scatter diagram, the bad ones will average out, and that the equation for the curve drawn through the clustered points will be good enough for use in design, always with a goodly factor of safety as a cushion. And it almost always is. This method is *quantitative*, *empirical*, and *expedient*. As used by the engineer, it is logical and successful.

It is of course used by investigators in many fields other than engineering. Friends in physics and chemistry tell me that it accounts for a large percentage of the current research in those sciences. A recent paper by Paul Weiss (1962) with the subtitle "Does Blind Probing Threaten to Displace Experience in Biological Experimentation?" calls attention to its increasing use in biology. The approach and examples are different, but the basic views of Dr. Weiss correspond so closely with those expressed in this essay that I am inclined to quote, not a passage or two, but the whole paper. Because this is impracticable, I can only urge that geologists interested in this phase of the general problem —whither are we drifting, methodologically?—read it in the original.

In view of its widespread use in science, what I have been calling the engineering or technologic method certainly should not be identified, by name, with engineering or technology as such. And on the other side of the coin, the so-called scientific method is used more consistently and effectively by many engineers and technologists than by most scientists. Besides being inappropriate on this score, both terms have derogatory or laudatory connotations which beg some questions. So, with serious misgivings that will be left unsaid, I will from here on use the term "rational method" for what we are accustomed to think of as the scientific method, and what I have been calling the engineering method will be referred to as the empirical method.[3]

Actually, the method that I am trying to describe is *an* empirical method; it is shotgun or scatter-diagram empiricism, very different from the one-at-a-time, cut-and-try empiricism of Ehrlich who, without any reasoned plan, tried in turn 606 chemical substances as specifics for syphilis. The 606th worked. Both the scatter-diagram and the one-at-a-time types can be, at one extreme, purely empirical, or, if you prefer, low-level empirical. As Conant (1952, pp. 26–30) points out, the level is raised—the empirical approaches the rational —as the gathering and processing of the data are more and more controlled by reasoning.

Use of Examples

The expositions of the rational method by Gilbert, Chamberlin, and Johnson all depend on the use of examples, and having tried several other ways, I am sure that this is the only way to make clear the workings of the empirical method. I have chosen to use actual examples, because these are far more effective than anything I could invent. They could have been selected from any field in geology. My examples are from recent publications dealing with the geologic work of rivers; I know of no other field in which the two approaches to problem-solving stand in such sharp contrast. "Horrible examples" are available, analysis of which would have a certain entertainment value; I shall draw my examples from publications that rank as important contributions. The principal example is from a paper that is unquestionably the outstanding report in this field, "The hydraulic geometry of stream channels," by Luna Leopold and Thomas Maddock (1953). I have discussed the methodology of geologic investigation with Leopold on numerous occasions, and we have, in effect, agreed to disagree on some points.[4]

[3] So many friends have objected to these terms that I should say that I am fully aware that they are unsatisfactory, chiefly because they have different connotations in different fields of study. I use them in their plain English meaning. They seem to me to be less objectionable than any other terms, but I will not take issue with those who think otherwise.

Examples are essential in a discussion of methods, but it is difficult to work with them. The problems of fluvial hydraulics are so complex that if the examples are to be comprehensible they must be simplified, and we must treat them out of context. This may irritate the few who are familiar with these matters at the technical level; I can only ask their indulgence on the ground that I am steering a difficult course between nonessential complexity and over-simplification. I should acknowledge, moreover, that I am an interested party; about 15 years ago I published an article in this field (Mackin, 1948). Finally, and most important, I will be deliberately looking at the way data are handled from the point of view of the conservative geologist, unaccustomed to this manner of handling data and highly critical of it. But I will come around full circle in the end, to indicate that the operations I have been criticizing are those of a valid method of investigation which is here to stay.

Downstream Change in Velocity in Rivers

All of us have seen the white water of a rushing mountain stream and the smooth-surfaced flow of the streams of the plains, and we are prepared by the contrast to suppose that the velocity of the flow decreases downstream. We are aware, moreover, that slope commonly decreases downstream and that velocity tends to vary directly with slope. Finally, we know from observation that the grain size of the load carried by rivers tends to decrease downstream, and that the grain size of the material carried by a river varies directly with some aspect of the velocity. For these reasons, we have always taken it for granted that velocity decreases downstream.

So in 1953, when Leopold and Maddock stated that velocity in rivers *increases* downstream, the statement came as a first-rate shock to most geologists. Three graphs (Fig. 1) from that article are good examples of the sort of evidence, and the manner of handling evidence, on which this generalization is based. They are log-log plots of several parameters; at the top, width of channel against discharge in cubic feet per second; in the middle, depth against discharge; and at the bottom, velocity against discharge. Each point represents data obtained from a U. S. Geological Survey gaging station in the Yellowstone-Big Horn drainage system. The points at the far left, such as 13 and 16, are on small headwater tributaries, and those at the far right, such as 19, are on the main stem of the Yellowstone. The upper and middle graphs show that, as should be expected, both width and depth increase with increase in dis-

[4] Leopold states his position elsewhere in this volume.

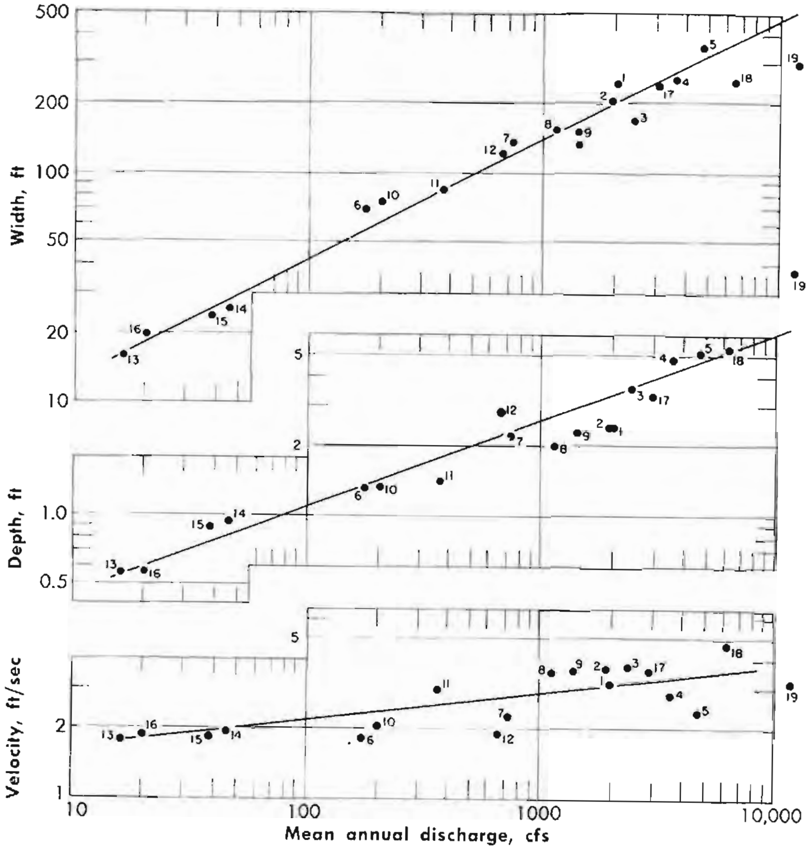


FIG. 1. Width, depth, and velocity in relation to discharge, Bighorn and Yellowstone Rivers, Wyoming and Montana (Leopold and Maddock, 1953, Fig. 6).

charge; the line in the lower graph also slopes up to the right; that is, velocity increases with increase in discharge.

Some may wonder why we have moved over from increase in velocity downstream, which is the exciting issue, to increase in velocity with increase in discharge. While it is true that discharge increases downstream in most rivers, it is at best only an approximate measure of distance downstream—the distance that would be traveled, for example, by the grains composing the load. The answer given in the Leopold-Maddock paper is that there were not enough gaging stations along the rivers to provide a sufficient number of points. Use of discharge, rather than distance, makes it possible to bring onto one graph the main stream and its tributaries of all sizes; or, for that matter, since "main stream" is a relative term, all the neighboring streams in an area large enough to provide enough points to bring out the significant relationships.

This explanation does not quite answer the question, unless expediency is an answer, but it raises another question.

Velocity at any given place—at any gaging station, for example—varies with variations in discharge from time to time during the year; as discharge and depth increase, usually in the spring, velocity at a given place increases very markedly. We may ask, then, *what* discharge is represented by the points on the lower graph? The question is pertinent, because we know that in most rivers much of the year's transportation of bed load—the sand and gravel that move along the bed—is accomplished during a relatively brief period of maximum discharge. But these graphs show mean annual discharges, and the velocities developed at those discharges. The reason for using mean annual discharge is said to be that this parameter is readily available at a large number of gaging stations. This explanation does not answer the question: what is the relevance of mean annual discharge in an analysis of the geologic work of rivers?

This general question, which applies to each of the stations considered individually, takes on another meaning when the relations between mean annual discharge and maximum discharge on streams are considered. Reference to Water Supply Paper 1559 (1960, p. 169) indicates that at point 13 (Fig. 1), which represents a gaging station on the North Fork of Owl Creek, the average annual discharge for the 14-year period of record was 15 cfs (cubic feet per second), whereas the maximum discharge during the same period was 3200 cfs; that is, the maximum was about 213 times the average. The same paper (p. 234) indicates that at point 19, which represents the Yellowstone River at Sidney, Montana, the average annual discharge over a 46-year period was 13,040 cfs, whereas the maximum during the same period was 159,000 cfs; here the maximum was about 12 times the average. The noteworthy thing about this graph—the thing that makes it so exciting—is that it shows that velocity *increases* downstream although we know from observation that grain size *decreases* downstream. The significance of the graph is more readily understood when we remember: (1) that the larger grains move only at times of maximum discharge; (2) that this graph shows mean annual discharge; and (3) that in the small rivers on the left side, the maximum discharge may be more than 200 times as great as the discharge shown on the graph, while in the big rivers on the right, it is less than 20, and usually less than 10 times the discharge shown, that is, that the critical ratios on the two sides are of a different order of magnitude. The slope of the line is an important statistical fact, but it does not bear directly on transportation of bed load by rivers.

One more thought in this connection. The depth at average annual discharge at point 13, on the North Fork of Owl Creek, is shown in the middle graph as being something less than 0.6 foot. I know the general area, and, although I have no measurements at this gaging station, it is my recollection that the larger boulders on the bed of the North Fork are more than 0.6 foot in diameter—the boulders on the bed have diameters that are of the same order of magnitude as the depth at which the very low velocity shown for this point was calculated. Similar relationships obtain for other small headwater streams, the points for which anchor down, so to speak, the left end of the line.

Let us look briefly at one more aspect of the case. The velocity is lower near the bed of a river than near the surface. Rubey (1938) and others have shown that the movement of bed load is determined, not by the average velocity, but by the velocity near the bed. And it has also been established that the relation between average velocity and what Rubey calls "bed velocity" varies markedly with depth of water, roughness of channel, and other factors. We may reasonably ask, then, *what* velocity is represented by the points on the graph? The answer is spelled out clearly by Leopold and Maddock (1953, p. 5).

> Velocity discussed in this report is the quotient of discharge divided by the area of the cross section, and is the mean velocity of the cross section as used in hydraulic practice . . . This mean velocity is not the most meaningful velocity parameter for discussing sediment transport, but it is the only measure of velocity for which a large volume of data is available. Although the writers recognize its limitations, the mean velocity is used here in lieu of adequate data on a more meaningful parameter.

There are various other similar questions about this graph, some of which are discussed by the authors in the clear and candid style of the last quotation. I will not develop these questions, or the secondary and tertiary questions that spring from the answers. Some of you may be thinking: never mind the individual points; what about the trends? It could be argued that if the conclusions are internally consistent; if they match those for other river systems; if, in short, these procedures get results, this alone justifies them.

Let's look at the results. Figure 2 is the velocity-discharge graph of Fig. 1, modified by use of symbols to identify related points and with dashed lines for individual rivers.

Points 1, 2, 3, 4, and 5, are on the main stem of the Big Horn River. Points 1, 2, and 3 are in the Big Horn Basin; point 4 is about 50 miles downstream from 3, and 5 is about 20 miles downstream from 4. The dashed line, which fits these points quite well, slopes down to the right; it means that on the main stem of the Big Horn, velocity *decreases* downstream.

Points 6, 7, 8, and 9 are on the Wind River, which is actually the upper part of the Big Horn River. I do not know whether points 8 and 9 represent the same types of channel conditions as 1, 2, 3, 4, and 5, as suggested by their positions, or whether they should be grouped with 6 and 7, as called for by the

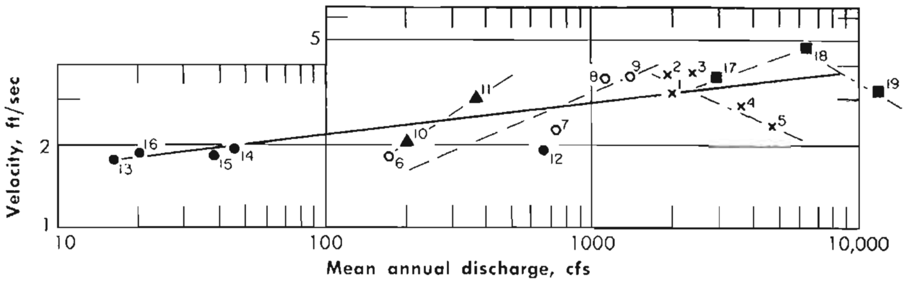


FIG. 2. Same as velocity-discharge graph in Fig. 1, with dashed lines for certain rivers; triangles, Greybull River; open circles, Wind River; X's, Bighorn River; solid squares, Yellowstone River.

geographic usage of the names. Let us say, then, that in what the geographers call the Wind River, velocity at average annual discharge increases downstream.

Points 10 and 11, both on the Greybull River, also suggest by their relative position that velocity increases downstream; the line slopes up to the right. But point 10 is near the mouth of the Greybull, about 40 airline miles downstream from 9; the average annual discharge decreases downstream (Leopold, 1953, p. 612) partly because of withdrawal of water for irrigation. Velocity actually decreases very markedly downstream on the Greybull.

Points 17, 18, and 19 are on the main stem of the Yellowstone. It appears from this graph that velocity increases downstream between 17 and 18, and decreases downstream between 18 and 19.

The generalization that velocity increases downstream, at a rate expressed by the slope of the solid line on this graph, is a particular type of empirical answer. It is what the nonstatistician is likely to think of as an "insurance company" type of statistic—a generalization applying to this group of rivers collectively, but not necessarily to any member of the group. Of the river segments represented on the graph, about half increase in velocity downstream, and about half decrease in velocity downstream. As shown by the different slopes of the dashed lines, in no two of them is the rate of change in velocity the same.

This is really not very surprising. The solid line averages velocity-discharge relations in river segments that are, as we have seen, basically unlike in this respect. Moreover, slope, which certainly enters into velocity, is not on the graph at all. For these reasons the equation of the solid line is not a definitive answer to any geologic question.

But—and here I change my tune—this graph was not intended to provide a firm answer to any question. It is only one step—a preliminary descriptive step—in an inquiry into velocity changes in rivers from head to mouth. This is accomplished by plotting certain conveniently available data on a scatter diagram.

I have indicated earlier how this procedure, which is empirical, expedient, and quantitative, serves the practicing engineer very well in getting answers that are of the right order of magnitude for use in design in deadline situations. Here we see the same procedure operating as a step in a scientific investigation. It is used in this graph to learn something about velocity relations in rivers from a mass of data that were obtained for a different purpose; the purpose of U. S. Geological Survey gaging stations is to measure discharge, not velocity. This gleaning of one kind of information from measurements—particularly long-term records of measurements—that are more or less inadequate because they were not planned to provide that kind of information, is a very common operation in many scientific investigations, and is altogether admirable.

There is another point to be made about this graph. Before the work represented by it was done, there had been no comprehensive investigations of velocity in rivers from head to mouth; this study was on the frontier. In these circumstances, some shots in the dark—some *shotgun* shots in the dark—were quite in order. The brevity with which this point can be stated is not a measure of its importance.

Finally, I wish to emphasize that Leopold and Maddock did not regard the solid line as an answer—its equation was *not* the goal of their investigation. They went on in this same paper, and in others that have followed it, to deal with velocities developed at peak discharges and with many other aspects of the hydraulic geometry of river channels. It is for this reason, and the other two reasons just stated, that I can use the graph as I have without harm to its authors.

But our literature is now being flooded by data and graphs such as these, without any of the justifications, engineering or scientific, that I have outlined. In many instances the graph is simply a painless way of getting a quantitative answer from a hodge-podge of data, obtained in the course of the investigation, perhaps at great expense, but a hodge-podge nevertheless because of the failure of the investigator to think the problem through prior to and throughout the period of data gathering. The equations read from the graphs or arrived at by other mechanical manipulations of the data are presented as terminal scientific conclusions. I suggest that the equations may be terminal engineering conclusions, but, from the point of view of science, they are statements of problems, not conclusions. A statement of a problem may be very valuable, but if it is mistaken for a conclusion, it is worse than useless because it implies that the study is finished when in fact it is only begun.

If this empirical approach—this blind probing—were the only way of quantifying geology, we would have to be content with it. But it is not; the quantitative approach is associated with the empirical approach, but it is not wedded to it. If you will list mentally the best papers in your own field, you will discover that most of them are quantitative and rational. In the study of rivers

I think of Gilbert's field and laboratory studies of Sierra Nevada mining debris (1914, 1917), and Rubey's analysis of the force required to move particles on a stream bed (1938). These geologists, and many others that come to mind, have (or had) the happy faculty of dealing with numbers without being carried away by them—of quantifying without, in the same measure, taking leave of their senses. I am not at all sure that the percentage of geologists capable of doing this has increased very much since Gilbert's day. I suggest that an increase in this percentage, or an increase in the rate of increase, is in the direction of true progress.

We shall be seeing more and more of shotgun empiricism in geologic writings, and perhaps we shall be using it in our own investigations and reports. We must learn to recognize it when we see it, and to be aware of both its usefulness and its limitations. Certainly there is nothing wrong with it as a tool, but, like most tools, how well it works depends on how intelligently it is used.

Causes of Slope of the Longitudinal Profile

We can now turn to a matter which seems to me the crux of the difference between the empirical and the rational methods of investigation, namely, cause-and-effect relations.[5] I would like to bring out, first, an important difference between immediate and superficial causes as opposed to long-term, geologic causes; and second, the usefulness, almost the necessity, of thinking a process through, back to the long-term causes, as a check on quantitative observations and conclusions.

Most engineers would regard an equation stating that the size of the pebbles that can be carried by a river is a certain power of its bed velocity as a complete statement of the relationship. The equation says nothing about cause

[5] I am aware that my tendency to think in terms of cause and effect would be regarded as a mark of scientific naïveté by some scientists and most philosophers. My persistence in this habit of thought after having been warned against it does not mean that I challenge their wisdom. Perhaps part of the difficulty lies in a difference between what I call long-term geologic causes and what are sometimes called ultimate causes. For example, a philosopher might say, "Yes, it is clear that such things as discharge and size of pebbles may control or cause the slope of an adjusted river, but what, then, is the cause of the discharge and the pebble size? And if these are effects of the height of the mountains at the headwaters, what, then, is the cause of the height of the mountains"? Every cause is an effect, and every effect is a cause. Where do we stop? I can only answer that I am at the moment concerned with the geologic work of rivers, not with the cause of upheaval of mountains. The question, where do we stop?, is for the philosopher, who deals with all knowledge; the quest for ultimate causes, or the futility of that quest, is in his province. The investigator in science commonly stays within his own rather narrow field of competence and, especially if time is an important element of his systems, he commonly finds it useful to think in terms of cause and effect in that field. The investigator is never concerned with ultimate causes.

and effect, and the engineer might be surprised if asked which of the two, velocity or grain size, is the cause and which is the effect. He would almost certainly reply that velocity controls or determines the size of the grains that can be moved, and that therefore velocity is the cause. To clinch this argument, he might point out that if, by the turn of a valve, the velocity of a laboratory river were sufficiently increased, grains that previously had been at rest on the bed would begin to move; that is, on the basis of direct observation, and by the commonsense test of relative timing, the increase in velocity *is* the cause of the movement of the larger grains. This is as far as the engineer needs to go in most of his operations on rivers.

He might be quite willing to take the next step and agree that the velocity is, in turn, partly determined by the slope. In fact, getting into the swing of the cause-and-effect game, he might even volunteer this idea, which is in territory familiar to him. But the next question—what then, is the cause of the slope?—leads into unfamiliar territory; many engineers, and some geologists, simply take slope for granted.

Our engineer would probably be at first inclined to question the sanity of anyone suggesting that the size of the grains carried by a river determines the velocity of the river. But in any long-term view, the sizes of the grains that are supplied to a river are determined, not by the river, but by the characteristics of the rocks, relief, vegetative cover, and other physical properties of its drainage basin. If the river is, as we say, graded (or as the engineer says, adjusted), this means that in each segment the slope is adjusted to provide just the transporting power required to carry through that segment all the grains, of whatever size, that enter it from above. Rivers that flow from rugged ranges of hard rock tend to develop steep slopes, adjusted to the transportation of large pebbles. Once they are developed, the adjusted slopes are maintained indefinitely, as long as the size of the pebbles and other controlling factors remain the same. Rivers that are supplied only with sand tend to maintain low slopes appropriate to the transportation of this material.

If the sizes of the grains supplied to a given segment of an adjusted river are abruptly increased by uplift, by a climatic change, or by a work of man, the larger grains, which are beyond the former carrying power, are deposited in the upper part of the segment; the bed is raised thereby and the slope is consequently steepened. This steepening by deposition continues until that particular slope is attained which provides just the velocity required to carry those larger grains, that is, until a new equilibrium slope is developed, which the river will maintain thence forward so long as grain size and other slope-controlling conditions remain the same.

Thus in the long view, velocity is adjusted to, or determined by, grain size; the test of relative timing (first the increase in grain size of material supplied to the river, and then, through a long period of readjustment, the increase in velocity) marks the change in grain size as the cause of the change in velocity. Note that because the period of readjustment may occupy thousands of years, this view is based primarily on reasoning rather than on direct observation. Note also that we deal here with three different frames of reference spanning the range from the empirical to the rational.

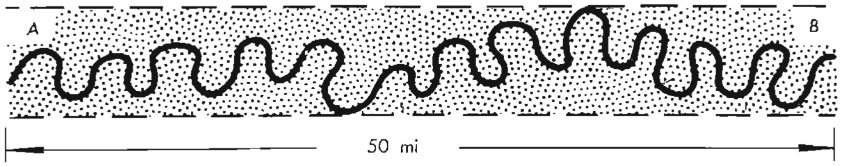
The statement that grain size tends to vary directly with bed velocity is an equation, whose terms are transposable; neither time nor cause and effect are involved, and this first frame may be entirely empirical. The numerical answer is complete in itself.

The short-term cause-and-effect view, that grain size is controlled by bed velocity, is in part rational, or if you prefer, it represents a higher level of empiricism. As I see it, this second frame has a significant advantage over the first in that it provides more fertile ground for the formulation of working hypotheses as to the mechanical relations between the flow and the particle at rest or in motion on the bed, leading to purposeful observation or to the design of experiments.

The third frame, the long-term view, that velocity is controlled by grain size, has a great advantage over the short-term view in that it provides an understanding of the origin of slope, which the short-term view does not attempt to explain. It is largely rational, or if you prefer, it represents a still higher level of empiricism.

Because I think that the objective of science is an understanding of the world around us, I prefer the second and third frames to the first, but I hope that it is clear that I recognize that all the frames are valid; the best one, in every instance, is simply the one that most efficiently gets the job done that needs doing. The important thing is to recognize that there *are* different frames; and that they overlap so completely and are so devoid of boundaries that it is easy to slip from one to the other.

The difference between the rational and the empirical approach to this matter of river slope, and the need for knowing what frame of reference we are in, can be clarified by a little story. One of the earliest theories of the origin of meanders, published in a British engineering journal in the late eighteen hundreds, was essentially as follows: divested of all geographic detail. Two cities *A* and *B*, both on the valley floor of a meandering river, are 50 airline miles apart. City *B* is 100 feet lower than city *A*; hence the average slope of the valley floor is two feet per mile (Fig. 3). But the slope of the river, measured round its loops, is only one foot per mile. The British engineer's theory was, in effect, though not expressed in these words, that the river said to itself, "How, with a slope of one foot per mile, can I manage to stay on a valley floor with a slope of two feet per mile? If I flow straight down the middle of the valley floor, starting at *A*, I will be 50 feet *above* the valley floor at *B*, and that simply will not do." Then it occurred to the river that it could meet this



$$M = \frac{S_y}{S_r}$$

FIG. 3. Diagram illustrating an hypothesis for the origin of meanders.

problem by bending its channel into loops of precisely the sinuosity required to keep it *on* the valley floor, just as a man might do with a rope too long for the distance between two posts. And it worked, and that's why we have meanders.

Note that this theory not only explains meandering qualitatively, but puts all degrees of meandering, from the very loopy meanders of the ribbon-candy type to those that are nearly straight, on a firm quantitative basis—the sinuosity or degree of meandering, M, equals the slope of the valley floor, S_y, over the slope of the river, S_r.

There is nothing wrong with this equation, so long as it only describes. But if its author takes it to be an explanation, as the British engineer did, and if he slips over from the empirical frame into the rational frame, as he may do almost without realizing it, he is likely to be not just off by an order of magnitude, but upside-down—to be not only wrong but ludicrous. This explanation of meanders leaves one item out of account—the origin of the valley floor. The valley floor was not opened out and given its slope by a bulldozer, nor is it a result of special creation prior to the creation of the river. The valley floor was formed by the river that flows on it.

Causes of Downvalley Decrease in Pebble Size

It is a matter of observation that there is commonly a downvalley decrease in the slopes of graded rivers, and it is also a matter of observation that there is commonly a downvalley decrease in the size of pebbles in alluvial deposits. A question arises, then, as to whether the decrease in slope is caused in part by the decrease in pebble size, or whether the decrease in pebble size is caused in part by the decrease in slope, or whether both of these changes are independent or interdependent results of some other cause. My third and last example applies the empirical and rational approaches to a part of this problem, namely, what are the causes of the decrease in pebble size? The reasoning is somewhat more involved than in the other examples; in this respect it is more truly representative of the typical geologic problem.

The downvalley decrease in pebble size could be caused by either of two obvious, sharply contrasted mechanisms: (1) abrasional wear of the pebbles as they move along the bed of the stream, and (2) selective transportation, that is, a leaving-behind of the larger pebbles. The question is, which mechanism causes the decrease, or, if both operate, what is their relative importance?

There is no direct and satisfactory way of obtaining an answer to this question by measurement, however detailed, of pebble sizes in alluvial deposits. The most commonly used approach is by means of laboratory experiment. Usually fragments of rock of one or more kinds are placed in a cylinder which can be rotated on a horizontal axis and is so constructed that the fragments slide, roll, or drop as it turns. The fragments are remeasured from time to time to determine the reduction in size, the corresponding travel distance being calculated from the circumference of the cylinder and the number of rotations. This treatment does not approximate very closely the processes of wear in an actual river bed. Kuenen (1959) has recently developed a better apparatus, in which the fragments are moved over a concrete floor in a circular path by a current of water. Whatever the apparatus, it is certain that the decrease in pebble size observed in the laboratory is due wholly to abrasion, because none of the pebbles can be left behind; there is no possibility of selective transportation.

When the laboratory rates of reduction in pebble size per unit of travel distance are compared with the downvalley decrease in pebble size in alluvial deposits along most rivers, it is found that the decrease in size along the rivers is somewhat greater than would be expected on the basis of laboratory data on rates of abrasion. If the rates of abrasion in the laboratory correctly represent the rates of abrasion in the river bed, it should be only necessary to subtract to determine what percentage of the downvalley decrease in grain size in the alluvial deposits is due to selective transportation.

Field and laboratory data bearing on this problem have been reviewed by Scheidegger (1961) in his textbook, "Theoretical Geomorphology," which is about as far out on the quantitative side as it is possible to get. Scheidegger (p. 175) concludes that ". . . the most likely mechanism of pebble gradation in rivers consists of pebbles becoming contriturated due to the action of frictional forces, but being assigned their position along the stream bed by a sorting process due to differential transportation."

If I understand it correctly, this statement means that pebbles are made smaller by abrasion, but that the downvalley decrease in pebble size in alluvial deposits is due largely (or wholly?) to selective transportation.

On a somewhat different basis—the rate of reduction of pebbles of less resistant rock, relative to quartzite, in a downvalley direction in three rivers east of the Black Hills—Plumley (1948) concludes that about 25 per cent of the reduction in these rivers is due to abrasion, and about 75 per cent is due to selective transportation.

These two conclusions as to the cause of the downstream decrease in pebble size, solidly based on measurements, agree in ascribing it mainly to selective transportation. Let us try a different approach—let us think through the long-term implications of the processes.

Downstream decrease in pebble size by selective transportation requires that the larger pebbles be left behind permanently. The three-inch pebbles, for example, move downstream to a certain zone, and are deposited there because they cannot be transported farther. The two-inch pebbles are carried farther downstream, to be deposited in an appropriate zone as the slope decreases. These zones may have considerable length along the stream, they may be poorly defined, and they may of course overlap, but there is a downstream limit beyond which no pebbles of a given size occur in the alluvial deposits because none could be carried beyond that limit, which is set by transporting power.

Consider a river carrying a bed load of sand and gravel under steady-state conditions such that the slope and altitude in a given segment are maintained indefinitely without change, and let it be assumed for simplicity that the channel is floored and walled by rock (Fig. 4a). The load moves chiefly

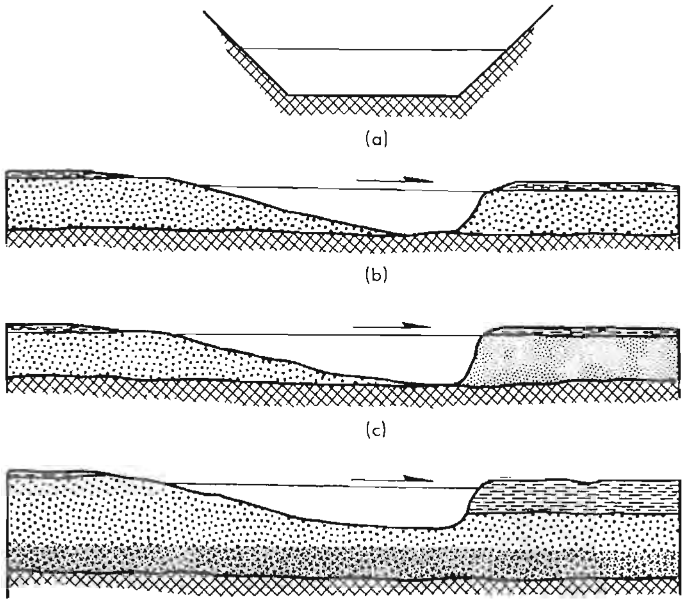


FIG. 4. Diagram illustrating exchange in graded and aggrading rivers.

during high-water stages and lodges on the bed during low-water stages. The smaller pebbles are likely to be set in motion sooner than the larger pebbles during each rising stage, they are likely to move faster while in motion, and they are likely to be kept in motion longer during each falling stage. In this sense, the transportation process is selective—if a slug of gravel consisting of identifiable pebbles were dumped into the segment, the smaller pebbles would outrun the larger, and this would cause a downstream decrease in the sizes of these particular pebbles in the low-water deposits. But in the steady-state condition, that is, with a continuous supply of a particular type of pebble or of pebbles of all types, all the pebbles deposited on the bed during the low stages must be placed in motion during the high stages; if the larger pebbles were permanently left behind during the seasonal cycles of deposition and erosion, the bed would be raised, and this, in turn, would change the condition. A non-aggrading river flowing in a channel which is floored and walled by rock cannot rid itself of coarse material by deposition because there is no place to deposit it where it will be out of reach of the river during subsequent fluctuations of flow; every pebble entering a given segment must eventually pass on through it. The smaller pebbles move more rapidly into the segment than the larger pebbles, but they also move more rapidly out of it. In the steady-state condition, the channel deposits from place to place in the segment contain the same proportions of the smaller and larger pebbles as though all moved at the same rate. Selective transportation cannot be a contributing cause of a downstream decrease in pebble size in our model river because there can be no selective deposition.

In a real river that maintains the same level as it meanders on a broad valley floor, bed load deposited along the inner side of a shifting bend is exchanged for an equal volume of slightly older channel deposits eroded from the outside of the bend. If these channel deposits were formed by the same river, operating under the same conditions and at the same level over a long period of time (Fig. 4b), the exchange process would not cause a reduction in the grain size of the bed load; insofar as selective transportation is concerned, the relation would be the same as in our model river. But if, by reason of capture or climatic change or any other change in controlling conditions, the older alluvial deposits in a given segment are finer grained than the bed load now entering that segment (Fig. 4c), exchange will cause a decrease in pebble size in a downstream direction, at least until the older deposits have been completely replaced by deposits representing the new regime. Exchange also causes a reduction in grain size if the river, maintaining the same level, cuts laterally into weak country rock that yields material finer in grain size than the load that is being concomitantly deposited on the widening valley floor.

The selective transportation associated with the process of exchange in the graded river, while by no means negligible, is much less effective as a cause of downstream decrease in pebble size than the selective transportation that characterizes the aggrading river. The essential difference is shown in Fig. 4(d); some of the deposits formed by one swing of the aggrading river across its valley floor are not subject to reworking in later swings, because the channel is slowly rising. The largest pebbles in transit in a given segment in a high-water stage are likely to be concentrated in the basal part of the deposit formed during the next falling stage. Thus the aggrading river rids itself of these pebbles, selectively and permanently, and there is a corresponding downstream decrease in pebble sizes in the deposits.

If upbuilding of the flood plain by deposition of overbank material keeps pace with aggradational rising of the channel, the shifting meanders may exchange channel deposits for older alluvium consisting wholly or in part of relatively fine-grained overbank material (Fig. 4d). But in rapidly aggrading rivers this rather orderly process may give way to a fill-spill mechanism in which filling of the channel is attended by the splaying of channel deposits over adjoining parts of the valley floor. On some proglacial outwash plains this type of braiding causes boulder detritus near the ice front to grade into pebbly sand within a few miles; there is doubtless some abrasional reduction in grain size in the proglacial rivers, but nearly all the decrease must be due to selective transportation.

Briefly then, thinking the process through indicates that the downstream decrease in grain size in river deposits in some cases may be almost wholly due to abrasion, and in others almost wholly due to selective transportation, depending primarily on whether the river is graded or aggrading and on the rate of aggradation. It follows that no generalization as to the relative importance of abrasion versus selective transportation in rivers—all rivers—has any meaning.

A different way of looking at this problem has been mentioned in another connection. As already noted, selective transportation implies permanent deposition, for example, the three-inch pebbles in a certain zone, the two-inch pebbles in another zone farther downstream, and so on. If this deposition is caused by a downstream decrease in slope, as is often implied and sometimes stated explicitly (Scheidegger, p. 171), then what is the cause of the decrease in slope? We know that the valley floor was not shaped by a bulldozer, and we know that it was not formed by an act of special creation before the river began to flow. As we have seen in considering the origin of meanders, rivers normally shape their own valley floors. If the river is actively aggrading, this is usually because of some geologically recent change such that the gradient in a given segment is not steep enough to enable the river to move through that segment all of the pebbles entering it; in this (aggrading) river, the size of the pebbles that are carried is controlled in part by the slope, and the larger pebbles are left behind. But if the river is graded, the slope in each segment is precisely

that required to enable the river, under the prevailing hydraulic conditions, whatever they may be, to carry the load supplied to it. The same three-inch pebbles that are the largest seen on the bed and banks in one zone will, after a while, be the two-inch pebbles in a zone farther downstream.

We cannot wait long enough to verify this conclusion by direct observation of individual pebbles, because the pebbles ordinarily remain at rest in alluvial deposits on the valley floor for very long intervals of time between jogs of movement in the channel. We are led to the conclusion by reasoning, rather than by direct observation. In the long-term view, the graded river is a transportation system in equilibrium, which means that it maintains the same slope so long as conditions remain the same. There is no place in this self-maintaining system for permanent deposits: if the three-inch pebbles entering a given zone accumulated there over a period of geologic time, they would raise the bed and change the slope. As the pebbles, in their halting downvalley movement in the channel, are reduced in size by abrasion, and perhaps also by weathering while they are temporarily at rest in the valley floor alluvium, the slope, which is being adjusted to their transportation, decreases accordingly.

Does this reasoning settle the problem? Of course not! It merely makes us take a more searching look at the observational data. Since it is theoretically certain that the mechanisms which cause pebbles to decrease in size as they travel downstream operate differently, depending on whether the river is graded or aggrading, there is no sense in averaging measurements made along graded rivers with those made along aggrading rivers. However meticulous the measurements, and however refined the statistical treatment of them, the average will have no meaning.[6]

The reasoning tells us that, first of all, the rivers to be studied in connection with change of pebble size downstream must be selected with care. Because a steady-state condition is always easier to deal with quantitatively than a shifting equilibrium, it would be advisable to restrict the study, at the outset, to the deposits of graded rivers; when these are understood, we will be ready to deal with complications introduced by varying rates of aggradation. Similarly, it will be well, at least at the beginning, to eliminate altogether, or at least reduce to a minimum, the complicating effects of contributions from tributaries or other local sources: this can be done by selecting river segments without large tributaries, or by focusing attention on one or more distinctive rock types from known sources. There are unavoidable sampling problems, but some of these

[6] I owe to Frank Calkins the thought that, like most hybrids, this one would be sterile. The significance of this way of expressing what I have been saying about the averaging of unlike things is brought out by Conant's (1951, p. 25) definition of science as "an interconnected series of concepts and conceptual schemes that have developed as a result of experimentation and observation and are fruitful of further experimentation and observations. In this definition the emphasis is on the word 'fruitful'."

can readily be avoided; for example, there are many river segments in which the alluvial deposits are not contaminated by lag materials. Any attempt to develop sampling procedures must take into account, first of all, the fact that the channel *deposits* in a given segment of a valley differ significantly in gradation of grain size from the material moving through the channel in that segment in any brief period; the investigation may deal with the bed load (trapped in a box, so to speak), or with the deposits, or with both; but if both bed load and the deposits are measured, the measurements can only be compared, they cannot be averaged. Certainly we must investigate, in each river individually, the effects of weathering of the pebbles during periods of rest.

We must also take another hard look at the abrasion rates obtained by laboratory experiments, and try to determine in what degree these are directly comparable with abrasion rates in rivers. It is clearly desirable to develop other independent checks, such as those given by Plumley's measurement of rates of downstream reduction in sizes of pebbles of rock types differing in resistance to abrasion. Finally, it goes without saying that the reasoning itself must be continuously checked against the evidence, and one line of reasoning must be checked against others, to make sure that the mental wheels have not slipped a cog or two.

When we eventually have sufficient data on rates of downstream decrease of pebble size in alluvial deposits along many different types of rivers (considered individually), it will be possible to evaluate separately, in quantitative terms, the effect of special circumstances influencing the process of exchange in graded rivers, rates of aggradation in aggrading rivers, and the other causes of downstream decrease in pebble size. These generalizations will apply to all river deposits, modern as well as ancient, and it may even be that we can draw sound inferences regarding the hydraulic characteristics of the ancient rivers by comparing their deposits with those of modern rivers, in which the hydraulic characteristics can be measured.

This rational method of problem-solving is difficult and tortuous, but the history of science makes it clear, again and again, that if the system to be investigated is complex, the longest way 'round is the shortest way home; most of the empirical shortcuts turn out to be blind alleys.

Whither Are We Drifting, Methodologically?

I would like now to return to some of the questions asked at the outset. Must we accept, as gospel, Lord Kelvin's pronouncement that what cannot be stated in numbers is not science? To become respectable members of the scientific community, must we drastically change our accustomed habits of thought, abandoning the classic geologic approach to problem-solving? To the extent that this approach is qualitative, is it necessarily loose, and therefore bad?

Must we now move headlong to quantify our operations on the assumption that whatever is quantitative is necessarily rigorous and therefore good?

Why has the swing to the quantitative come so late? Is it because our early leaders, men such as Hutton, Lyell, Agassiz, Heim, Gilbert, and Davis, were intellectually a cut or two below their counterparts in classical physics? There is a more reasonable explanation, which is well known to students of the history of science. In each field of study the timing of the swing to the quantitative and the present degree of quantification are largely determined by the subject matter: the number and complexity of the interdependent components involved in its systems, the relative ease or difficulty of obtaining basic data, the susceptibility of those data to numerical expression, and the extent to which time is an essential dimension. The position of geology relative to the basic sciences has been stated with characteristic vigor by Walter Bucher (1941) in a paper that seems to have escaped the attention of our apologists.

Classical physics was quantitative from its very beginning as a science; it moved directly from observations made in the laboratory under controlled conditions to abstractions that were quantitative at the outset. The quantification of chemistry lagged 100 years behind that of physics. The chemistry of a candle flame is of an altogether different order of complexity from the physics of Galileo's rolling ball; the flame is only one of many types of oxidation; and oxidation is only one of many ways in which substances combine. There had to be an immense accumulation of quantitative data, and many minor discoveries—some of them accidental, but most of them based on planned investigations—before it was possible to formulate such a sweeping generalization as the law of combining weights.

If degree of quantification of its laws were a gage of maturity in a science (which it is not), geology and biology would be 100 to 200 years behind chemistry. Before Bucher (1933) could formulate even a tentative set of "laws" for deformation of the earth's crust, an enormous descriptive job had to be well under way. Clearly, it was necessary to know what the movements of the crust *are* before anybody could frame explanations of them. But adequate description of even a single mountain range demands the best efforts of a couple of generations of geologists, with different special skills, working in the field and the laboratory. Because no two ranges are alike, the search for the laws of mountain growth requires that we learn as much as we can about every range we can climb and also about those no longer here to be climbed; the ranges of the past, which we must reconstruct as best we can by study of their eroded stumps, are as significant as those of the present. Rates of growth and relative ages of past and present ranges are just as important as their geometry; the student of the mechanics of crustal deformation must think like a physicist and also like a historian, and these are very different ways of thinking, difficult to combine. The evidence is hard to come by, it is largely circumstantial, and there is never enough of it. Laboratory models are helpful only within narrow limits. So it is also with the mechanism of emplacement of batholiths, and the origin of ore-forming fluids, and the shaping of landforms of all kinds, and most other truly geologic problems.

It is chiefly for these reasons that most geologists have been preoccupied with manifold problems of description of geologic things and processes—*particular* things and processes—and have been traditionally disinclined to generalize even in qualitative terms. Because most geologic evidence cannot readily be stated in numbers, and because most geologic systems are so complex that some qualitative grasp of the problem must precede effective quantitative study, we are even less inclined to generalize in quantitative terms. Everybody knows the story of Lord Kelvin's calculation of the age of the earth.

These things are familiar, but they are worth saying because they explain why geology is only now fully in the swing to the quantitative. Perhaps it would have been better if the swing had begun earlier, but this is by no means certain. A meteorologist has told me that meteorology might be further ahead today if its plunge to the quantitative had been somewhat less precipitous—if there had been a broader observational base for a qualitative understanding of its exceedingly complex systems before these were quantified. At any rate, it is important that we recognize that the quantification of geology is a normal evolutionary process, which is more or less on schedule. The quantification will proceed at an accelerating pace, however much our ultraconservatives may drag their feet. I have been trying to point out that there is an attendant danger: as measurements increase in complexity and refinement, and as mathematical manipulations of the data become more sophisticated, these measurements and manipulations may become so impressive in form that the investigator tends to lose sight of their meaning and purpose.[7]

This tendency is readily understandable. Some of the appealing features of the empirical method have already been mentioned. Moreover, the very act of making measurements, in a fixed pattern, provides a solid sense of accomplishment. If the measurements are complicated, involving unusual techniques

[7] The subtitle of a recent article by Krumbein (1962), "Quantification and the advent of the computer open new vistas in a science traditionally qualitative" makes evident the overlap of our interests. Professor Krumbein's article deals explicitly with a mechanical method of processing data; the fact that there is no mention of the use of reasoning in testing the quality and relevance of the data to the specific issue being investigated certainly does not mean that he thinks one whit less of the "rational method" than I do. Similarly, I hope that nothing that I have said or failed to say is construed as meaning that I have an aversion to mechanical methods of analyzing data; such methods are unquestionably good if they bring out relationships not otherwise evident, or in any other way advance the progress of the rational method of investigation. When mechanical processes *replace* reasoning processes, and when a number *replaces* understanding as the objective, danger enters.

and apparatus and a special jargon, they give the investigator a good feeling of belonging to an elite group, and of pushing back the frontiers. Presentation of the results is simplified by use of mathematical shorthand, and even though nine out of ten interested geologists do not read that shorthand with ease, the author can be sure that seven out of the ten will at least be impressed. It is an advantage or disadvantage of mathematical shorthand, depending on the point of view, that things can be said in equations, impressively, even arrogantly, which are so nonsensical that they would embarrass even the author if spelled out in words.

As stated at the outset, the real issue is not a matter of classical geologic methods versus quantification. Geology *is* largely quantitative, and it is rapidly and properly becoming more so. The real issue is the rational method versus the empirical method of solving problems; the point that I have tried to make is that if the objective is an understanding of the system investigated, and if that system is complex, then the empirical method is apt to be less efficient than the rational method. Most geologic features—ledges of rock, mineral deposits, landscapes, segments of a river channel—present an almost infinite variety of elements, each susceptible to many different sorts of measurement. We cannot measure them all to any conventional standard of precision—blind probing will not work. Some years ago (1941) I wrote that the "eye and brain, unlike camera lens and sensitized plate, record completely only what they intelligently seek out." Jim Gilluly expresses the same thought more succinctly in words to the effect that most exposures provide answers only to questions that are put to them. It is only by thinking, as we measure, that we can avoid listing together in a field book, and after a little while, averaging, random dimensions of apples and oranges and apple crates and orange trees.

Briefly, then, my thesis is that the present swing to the quantitative in geology, which is good, does not necessarily and should not involve a swing from the rational to the empirical method. I'm sure that geology is a science, with different sorts of problems and methods, but not in any sense less mature than any other science; indeed, the day-to-day operations of the field geologist are apt to be far more sophisticated than those of his counterpart—the experimentalist—in physics or chemistry. And I'm sure that anyone who hires out as a geologist, whether in practice, or in research, or in teaching, and then operates like a physicist or a chemist, or, for that matter, like a statistician or an engineer, is not living up to his contract.

The best and highest use of the brains of our youngsters is the working out of cause and effect relations in geologic systems, with all the help they can get from the other sciences and engineering, and mechanical devices of all kinds, but with basic reliance on the complex reasoning processes described by Gilbert, Chamberlin, and Johnson.

REFERENCES CITED

Albritton, C. C., Jr., 1961, Notes on the history and philosophy of science (1) A conference on the scope and philosophy of geology: J. Graduate Research Center, Southern Methodist Univ., vol. 29, no. 3, pp. 188–192.

Bucher, W. H., 1933, The deformation of the earth's crust: Princeton, N. J., Princeton Univ. Press, 518 pp.

———, 1941, The nature of geological inquiry and the training required for it: Am. Inst. Mining Metall. Eng., Tech. Pub. 1377, 6 pp.

Chamberlin, T. C., 1897, The method of multiple working hypotheses: J. Geol., vol. 5, pp. 837–848.

Conant, J. B., 1951, Science and common sense: New Haven, Yale Univ. Press, 371 pp.

———, 1952, Modern science and modern man: New York, Columbia Univ. Press, 111 pp.

Garrels, R. M., 1951, A textbook of geology: New York, Harper, 511 pp.

Gilbert, G. K., 1886, The inculcation of the scientific method by example, with an illustration drawn from the Quaternary geology of Utah: Am. J. Sci., vol. 31, (whole no. 131), pp. 284–299.

———, 1914, The transportation of debris by running water: U. S. Geol. Survey, Prof. Paper, 86, 263 pp.

———, 1917, Hydraulic-mining debris in the Sierra Nevada: U. S. Geol. Survey, Prof. Paper 105, 154 pp.

Holton, G., 1952, Introduction to concepts and theories in physical science: Reading, Mass., Addison-Wesley, 650 pp.

Johnson, Douglas, 1933, Role of analysis in scientific investigation: Geol. Soc. Am., B., vol. 44, pp. 461–494.

Krumbein, W. C., 1960, The "geological population" as a framework for analysing numerical data in geology: Liverpool and Manchester Geol. J., vol. 2, pt. 3, pp. 341–368.

———, 1962, The computer in geology: Science, vol. 136, pp. 1087–1092.

Kuenen, P. H., 1959, Fluviatile action on sand, Part 3 of Experimental Abrasion: Am. J. Sci., vol. 257, pp. 172–190.

Leopold, L. B., 1953, Downstream changes of velocity in rivers: Am. J. Sci., vol. 251, pp. 606–624.

———, and Maddock, T., Jr., 1953, The hydraulic geometry of stream channels and some physiographic implications: U. S. Geol. Survey, Prof. Paper 252, 57 pp.

Mackin, J. H., 1941, Drainage changes near Wind Gap, Pennsylvania; a study in map interpretation: J. Geomorphology, vol. 4, pp. 24–53.

———, 1948, Concept of the graded river: Geol. Soc. Am., B., vol. 59, pp. 463–512.

Plumley, W. J., 1948, Black Hills terrace gravels; a study in sediment transport: J. Geol., vol. 56, pp. 526–577.

Rubey, W. W., 1938, The force required to move particles on a stream bed: U.S. Geol. Survey, Prof. Paper 189-E, pp. 120–140.

Scheidegger, A. E., 1961, Theoretical geomorphology: Berlin, Springer-Verlag, 333 pp.

Simpson, G. G., Roe, Anne, and Lewontin, R. C., 1960, Quantitative zoology: New York, Harcourt-Brace, 440 pp.

U. S. Geological Survey, 1960, Surface water supply of the United States, 1958; Part 6A, Missouri River Basin above Sioux City, Iowa: U. S. Geol. Survey Water-Supply Paper 1559, 434 pp.

Weiss, Paul, 1962, Experience and experiment in biology: Science, vol. 136, pp. 468–471.