

The Future of Generalized Systems of Classification

By PHYLLIS A. RICHMOND

IT IS FASHIONABLE at present to dismiss the future development of *generalized* classification as if there could be no such thing—as if no synthesis were ever possible again because no one can see any unifying factor or factors in the proliferation of subjects with which we have to deal. This is equivalent to saying that the sky is less blue because the blind man does not see it. There will be new generalized classification systems in the future for the simple reason that we have to have them. And anything we have to have sooner or later is found.

George Gaylord Simpson, the vertebrate paleontologist, points out that classification in zoology has proceeded alternately by stages of analysis and synthesis, carried out by men who might be categorized as "lumpers" and "splitters."¹ In library classification, by analogy, the older systematizers, Dewey, Bliss and others, would be "lumpers." Those who make faceted classifications and special subject analysis systems of all kinds are "splitters." When the "splitters" have finished analyzing the new alignment of parts that make up the sum total of human knowledge—an analysis which must be performed if we are to make any use at all of the mechanical and electronic aids now available to us—there is no reason to believe that a new synthesis cannot or will not be made. It is historically true that the impasse of one age is solved in the next, or the one after the next. The only fault, a very human one present in every generation, is that impatient men grow hopeless or even

Mrs. Richmond is Supervisor of River Campus Science Libraries in the University of Rochester.

antagonistic because they cannot see the solution in their own lifetimes.

In classification, as in any other discipline, one must lay groundwork for future development. Even if it is possible to be clairvoyant to the extent of predicting correctly what form future development must take, it is still necessary to take steps to get there. An Einstein, for instance, will hit upon "some great unifying idea from which one can deduce consequences that can ultimately be bought into agreement . . . with observed and measurable phenomena."² But even an Einstein does not build on nothing. He takes account of all the abortive, inconclusive, tentative, incomplete, or even rejected solutions of his predecessors. Thus fortified with knowledge of the major blind alleys, at least, he can lay out his course. It is the duty of those now working in classification to discover and explore all possible approaches so that valuable negative as well as positive evidence will be available for the ultimate synthesizer.

This is not to say that all current work in classification will be fruitless or that things will necessarily proceed so slowly that a synthesis is impossible before the end of the century. It is to say that avenues not yet considered must be explored and that work already begun must be carried much further. Ideas, no matter how bizarre, deserve considera-

¹George Gaylord Simpson, "The Principles of Classification and a Classification of Mammals," *Bulletin of the American Museum of Natural History*, LXXXV (October 1945), 22-24.

²Morris R. Cohen, *American Thought, a Critical Sketch* (Glencoe, Illinois: Free Press, 1954), p. 81.

tion, because, with any research, we never know exactly where we are heading. Nor do we know which line of development will bear the ultimate prize.

Essentially classification is the process of taking identifiable entities and relating them in one way or another to each other. These items may be anything the mind distinguishes as an entity in itself. Since the mind is rather free in coalescing raw material of all types into entities, this means that anything "thinkable" is classifiable.

Past syntheses almost exclusively have been constructed in a dendritic pattern—an inverted tree—from general to specific. The tree has been delineated as a plane figure and treated as if division were the only method of construction. Many of the writers who look with horror upon classification as a tool for organizing knowledge see only this type of deductive system and identify all classification with it.

A different approach is to build new generalized classification systems from collections of particulars, gathered by observation, ordered by reasoned criteria, and formed by induction into generalities as an ascending scale. These generalities can then be tested through hypothesis and deduction, and finally conclusions derived from the process, which, when tested and verified by further observation, will provide a solid basis for a more accurate representation of the world in which we live. This is essentially the scientific method, defined in very simple terms. Note that the main classes which will result from such a process are not predicted or predictable in advance.

Such a system, however, does not present a perfect picture of knowledge because it, too, usually ends in a dendritic pattern, although sometimes the net results are expressed in chains or even in chains made up from intermediary matrix or lattice analysis patterns. Something vital has been left out. This is a

study of the very nature of the material to be described. In the final analysis, it is this nature which will determine the system of classification. Any system which ignores actual makeup of individual subjects, with all their ramifications, is doomed from the start.

The universe we live in is apparently open and genuinely infinite, both infinitely big and infinitely small. Data, laws, methods, theories in all fields are partially and imperfectly known. On one hand, the possibility of discovery seems unending. On the other hand, the use of creative imagination appears limitless. While the idea of progress as a fundamental pervading force has been somewhat shaken as an accepted philosophic view in the last two decades, there is no corresponding diminution in either creativity or discovery, both of which are highly individualistic matters and probably not related to progress as an ideal.

It is the job of classification to show the waxing and waning of ideals as well as ideas, since the spirit of the times, its *Zeitgeist*, adds dimension to any aspect of the sum total of human knowledge. For reasons not entirely clear, this *Zeitgeist* has usually been ignored, especially in classification involving the sciences, medicine and technology, although it exists in these areas just as much as in the humanities. The totality of the intellectual atmosphere in each era determines to a large extent what shall be accomplished in that era. This totality, which is a dimension external to the individual, can be both narrow, representing the sum total of knowledge in a given field of research, or it can be broad, reflecting the generally-held attitudes, motivation, philosophical outlook, mores, and such of a period of time.

There is also an internal factor which adds dimension to classification. This is the individual's own world outlook, his *Weltanschauung*, which affects his personal motivation, intuitive capacity, and

imagination. This results from learning, influence of other people, inspiration, and experience. It is a process of accretion, that is to say, lifetime learning added up, and of inner maturation. The long process of growing up as a part of being alive is most noticeable in research in the humanities, where outstanding work is almost never done by young men.

The addition of further dimensions to the material to be classified greatly broadens the base of classification. The necessity for describing a body of knowledge with no known boundaries, which is the real work of classification, calls for a type of description which also can have no boundaries. In a discussion of the future of physics, Albert Einstein wrote:

The belief in an external world independent of the perceiving subject is the basis of all natural science. Since, however, sense perception only gives information of this external world or of "physical reality" indirectly, we can only grasp the latter by speculative means. It follows from this that our notions of physical reality can never be final. We must always be ready to change these notions—that is to say, the axiomatic substructure of physics—in order to do justice to perceived facts in the most logically perfect way. Actually a glance at the development of physics shows that it has undergone far-reaching changes in the course of time.³

This is true of all subjects, and it is this truth that has made all classifications obsolete over a period of time. So far as generalized classification is concerned, we have not even discovered most of the main categories, to say nothing of the thousands of related categories which must exist. At best, our classifications are an approximation, and subject to constant revision as new information becomes available.

Does this mean that all classification is a waste of time, since it can never be finished? The answer is "yes" only if by

³ Albert Einstein, *The World As I See It* (New York: Covici, Friede, 1934), p. 60.

classification we mean a process of assigning a fixed and relatively unalterable "address" to each categorical item. An ideal classification should be entirely flexible. Until all research and scholarly efforts cease, there will be no way of avoiding major alterations in subject organization, and, in fact, there should be no question of having anything but constant change. In making systems, provision for this change should be all the way through as required, not just at marginal spots where it will be the least upsetting. As the Indian classification theorist, Ranganathan, has ably put it:

A classificatory language must be nimble enough to keep step with the field of knowledge. Its expectation of life is determined by the degree of its self-perpetuating quality, *i.e.*, by the smallness of the dependence of the classifier on the classificationist [maker of the classification] to seize the correct class numbers of the new formations. The degree of its self-perpetuating quality is determined by the notational devices with which it is armed.⁴

In other words, the classification must be flexible enough to cope with unlimited additions to itself, and it must not be defeated by the notation used with it.

The notation used in a classification is an extremely important feature. Most notations are inadequate and actually strangle the classification system in the process of trying to reflect its internal structure. Something that will break this stranglehold and at the same time will broaden the class structure by freeing it from dependence upon limited criteria is a must for the future. It may even prove advisable to avoid notation in a classification or to use a random number system, with no effort to reflect the structure.

Classification is and must be a mirror of the intellectual as well as the external world it seeks to represent. If this com-

⁴ S. R. Ranganathan, "Self-Perpetuating Scheme of Classification," *Journal of Documentation*, IV (March 1949), 240.

bined world is a reality of four or five or more dimensions, the classification must have four or more dimensions, with some kind of spatial representation for visualization purposes.

Such a spatial representation is a *sine qua non* of adequate symbolic description of anything approaching reality. To use a letter-number string can be unwieldy as well as distorting. What this spatial representation should be—a fraction, lattice, or matrix in depth, or something better—cannot now be determined. The forms suggested seem too rigid. One needs a spatial “notation” that is more amorphous.

Such a spatial representation is needed for another reason. The human mind can cross subject boundary lines without much difficulty. Ideas in one field are applied or adapted to another with comparative ease. It is commonplace for new *area* studies to be set up in academic circles by drawing upon specialists from several fields to bring their different backgrounds to bear on a set of problems common to all of them. Thus we have new programs in space research, brain research, biomedical engineering, non-Western civilization and so on, all interdepartmental affairs. Old geographic divisions are replaced by new ones: Sub-Saharan Africa, Southwest Asia, Middle America. The late Robert C. Binkley, in teaching history, broke down the vertical stream of time applied to political-geographic regions (England since 1485) into horizontal units (all of Europe in December 1587). The next step logically is to realize that there is no reason for not treating space and time as a unit because that is exactly what they are.

Space-time is a dimension common to everything. Even the apparent “here and now” of contemporary science has retrospective features. For example, scientists never cease to remind us that, so far as they are concerned, the printed announcement of new research is already

out-of-date on the day it appears. This means that even as the announcement is read, it is describing something already in retrospect. This has significance—and not just in priority struggles. Even if a new generalized classification system is no more than an ephemeral conglomeration, it has to take into account the space-time factor.

Other dimensions are the *Zeitgeist* and *Weltanschauung* mentioned earlier. Just as no scholar or creative artist functions in a vacuum, so no output, be it a scientific discovery or a piece of poetry or a musical composition, comes from a void.⁵ All of them have antecedents, and sooner or later most also have descendants. There are branches of scholarship which specialize in locating and studying connections and relationships among things which apparently have sprung up *de novo*. A classification which only describes a limited number of qualities pertaining to entities and which deliberately stultifies itself because it must fit into a preconceived code of very limited proportions has small chance of lasting. First a classification must be free to cross subject boundary lines as freely as our minds cross them. Then the notation, if there is one, must be so free that the classification can expand into infinity in any direction.

Another point to be considered in the future development of generalized classification, in addition to taking a clear look at the nature of the material to be classified and making a flexible system while freeing this system from dependence upon notation, is the need to be able to show *relationships* between elements of a classification system in a completely unhampered way. This is extremely difficult. Up to now, few classification schemes, no matter how well designed or modified, have been able to show area relationships or cross-connec-

⁵ John Livingston Lowes, *The Road to Xanadu* (New York: Vintage Books, 1959) is an interesting example.

tions between subjects which pertain to two or more fields simultaneously. The only solutions attempted have been to enter such things in each related spot as a recognition of each separate emphasis, or to lock the codes for the individual related topics together in a linear chain or a series of more or less connected chains.

In this respect, a system which can build its own class descriptions, as with Uniterms, has advantages. With present trends in research tending more and more to break down existing barriers between subjects, and the content of individual fields being realigned in consequence of this, the barriers between subjects, as classically outlined, are a positive menace. If growth continues to be nonlinear, in the mathematical sense,⁶ and expansion continues in revised patterns, added freedom to cross boundaries in all directions is a vital necessity.

A further point is related to the preceding ones. In a classification, it is not only the basic categorical concept that is significant, but also the modification thereof. The multidimensional quality of knowledge is no less complex than the multidimensional descriptive possibilities for each individual item in a classification schedule. Even the descriptive terms themselves can be permuted to widen the scope of the system. A class is by nature an exclusive generalization. It is also inclusive, and in this respect covers considerable variety, according to its criteria. There are as many potential classification entities as there are terms applicable to the description of an element. A classification element, in turn, is only a temporary convenience, since it is potentially a class in itself, in the light of later knowledge. It may also be eliminated in the same light. A classification system, on the other hand, may never be fully described in the terse terms of a

classification schedule. Actually, the traditional classification systems have been made by arbitrarily selecting some descriptive factors and deliberately ignoring the rest. Indexes to the systems pick up some of the alternatives, but only in a hit-or-miss fashion. The faceted classifications, by their very nature, attempt to allow for a multitude of possible descriptive factors.

In sorting out the multidescriptive possibilities for classification elements, the *criteria* of the classification system are important. In the past, these criteria have been unwritten as a rule; in fact, it appears that in most cases they have been intuitively assigned. Sometimes the intuition has been quite realistic, judging by the state of knowledge at the time the systems were made. The science sections in the Dewey Classification, for instance, with the emphasis on paleontology, fit American science in the later 19th century like a shoe. Unfortunately, the creature wearing the shoe turned out to be a millipede—a situation with which Dewey's successors were never able to cope. Ranganathan has departed from intuitive practices to the extent of trying to spell out the criteria for each class. Others are carrying the process farther, but as yet there is no general agreement on what these criteria should be or even on how they should be reached.⁷

The criteria of classification might serve as mental diffraction gratings. Something is needed to split concepts systematically the way such gratings work with rays of light or other rays. It would seem that current research on descriptor language⁸ is a fine beginning, but only a beginning. The argumentative leap which produces a generalization from a collection of particulars

⁷ For a summary of the work of the English Classification Research Group, see D. J. Foskett, "The Classification Research Group, 1952-1962," *Libri*, XII (1962), 127-38.

⁸ B. C. Vickery, *On Retrieval System Theory* (London: Butterworths, 1961), pp. 23-55.

⁶ Ladis D. Kovach, "Life Can Be SO Nonlinear," *American Scientist*, XLVIII (June 1960), 218-25.

must be made and its nature noted. No attempt to do this has come at all, perhaps because of lack of awareness among classification makers that such a leap exists. Those who work with machines may be more alert to this possibility.

The advent of the giant calculating machine has inspired documentalists to see if a use cannot be made of it to organize the tremendous body of knowledge which threatens to overwhelm us with the products of our own brains. Success both in the reduction of mathematical calculations and in handling huge masses of factual data in business has led to considerable experimentation with such machines for literature searching and information retrieval. Three features, among others, stand out as a net result of this endeavor.

First, those concerned with the organization of knowledge, in attempt to make use of the machines, have tended to pay more attention to the capability of the machine than to the needs of the person using the knowledge or of the internal makeup of the knowledge itself. Early meetings between "hardware" men and librarians were unproductive because the librarians asked "What can you do?" and the engineers said "What do you want done?" Neither side understood the answers given by the other. Machines with positive talent in searching certain kinds of data were designed, but their effort was negated by the erroneous assumption that input was a simple matter, suitable for untrained clerks. After a succession of failures, the realization dawned that input is the major key to successful output. Here again, there was no rapport between the "hardware" men and librarians, who could have predicted this result. A productive combination of vast experience with classification and subject analysis on one side and brilliant inventiveness of the other has not been realized because of lack of communication between the two

groups. This gap has existed because there was no common frame of reference between the two until the "hardware" men tried out their systems and in the process made the same mistakes that the librarians had made fifty years ago. It is to the lasting shame of the catalog librarians that so few of them have made any attempt to mechanize the experienced approach which they alone possess.

The second feature is one that gives a false sense of capability to the machine. On the surface, finding a piece of factual data in a mass of other factual data appears not unlike finding a box of rotary switches in a warehouse full of electronic supplies. The joker here is that the warehouse is a selective situation, limited to a homogeneous collection of highly specialized items, and serving a small segment of the total buying public. If a collection of information has the equivalent qualities, homogeneity and limited access, then the machine solution to the problem of information retrieval looks pretty good. However, under such restricted circumstances, practically anything, including a hierarchical classification of the most rigid type, works well. The difficulty comes when the collection becomes heterogeneous.

Even given the same special situation, if the information to be retrieved is theoretical instead of factual—an idea, a line of thought, an argument for or against a point, a suggestion for future research—the machine falls even flatter on its transistorized face than the old classification. The classification at least has some way of indicating relationships and can give a hint to where to look for an answer. The machine is an all-or-nothing proposition.

The third factor is a temptation which came with the machine. The very success with mathematics, the relatively easy manner in which a computer can be

programmed to deal with a formula, the facility with which symbolic logic can be handled, all suggested that the thing to do was to find a mathematical model, or a series of mathematical models, to express the essence of organization of knowledge. Fit the data to the model and the model to the machine and the problem of information retrieval was as good as solved. Unfortunately this was equivalent to putting the cart before the horse, as even the most elementary grasp of the scientific method should have warned. Quantification—the formula for the law, the shorthand for an accepted generalization—is usually the last step, not the first, in the whole process of organization. Even when one begins with a great unifying idea, subject to verification by later observation, quantification comes slowly. For information retrieval, the mathematical model, to be satisfactory, has to wait until answers are found to questions that have not yet been asked.

There are many questions to be asked and answered before new generalized classification systems of any stature can be expected. A few of the basic ones have been raised in this study: How do we represent a multidimensional reality, presuming such exists, on paper? How do we cross departmental lines in subject fields with a classification system the way the human mind crosses them? How do we show relationships between parts of a system in a completely unhampered way? How do we indicate the multidimensional descriptive possibilities for each individual item being classified? How do we orient our criteria for classification in such a way that concepts may be diffracted into elements in a systematic fashion? Should we dispense with notation altogether? If not, can we use something more graphic than a notation made up of letters and numbers? Is it possible to get better spatial perception into our classification systems? It

would be convenient to replace the plane and linear by something with depth, perhaps analogous to proper motion and radial velocity in astronomy. Can we avoid the pitfall of adapting our classification systems to the capability of the computer instead of vice versa? Can we make a machine that will handle the infinite variety that composes the sum total of human knowledge? Will our final classification synthesis be capable of being represented by mathematical models?

The lines of development to be pursued in the future make it imperative that we also ask questions pertaining to methodology: Where can classification borrow ideas, techniques, and philosophic approaches from other disciplines? Do we have to stoop to some degree of fantasy in our classification composition before we can conquer with cold, hard facts? In other words, how much creative imagination can we put into a classification system without losing touch with reality? Is it possible that we have made a mistake in tending to view classification as a science rather than an art? Is some combination of the two, perhaps achieving the *rapprochement* between science and the humanities that has eluded us, the ultimate answer?

The future of generalized classification depends in large part upon man's ingenuity. So far, there has been no limit to the capabilities of the human mind, and there seems, therefore, to be no justification for the view that classification as a way of organizing knowledge is dead merely because the philosophic approaches used so far have led to blind alleys. It is time to look for new approaches. ■ ■

ACKNOWLEDGEMENT

The author is grateful to Mrs. Pauline Atherton, American Institute of Physics, and Gertrude Oellrich, Rutgers University, for reading and commenting upon this article in manuscript.