The Uses of Theory in the Simulation of Urban Phenomena

BRITTON HARRIS*, Professor of City and Regional Planning, University of Pennsylvania

A brief examination is made of the part played by theory in human development. Theory and practice are contrasted and a parallel contrast is drawn between science and engineering. An effort is made to demonstrate that in the modern world it is becoming increasingly difficult to draw a sharp line between these various opposed pairs of activities. It is also concluded that theory and practice have a well-known reciprocal relationship related to scientific induction and deduction. An examination is made of the creative jump from inductive generalization to the generation of new theories or new ways of looking at the world which will sustain extensive deduction and generate new hypotheses which are testable. Three criteria for useful theories are discussed: capacity for manipulation, fruitfulness, and economy. Other criteria such as realism, comprehensibility, and comprehensiveness are examined and found to be only partly applicable. The discussion is illustrated throughout with examples from land use and transportation simulation. Conclusions are drawn regarding future directions of research and some of the desirable characteristics of research establishments.

•MOST OF us who are engaged in one form or another of transportation and land use research have focused a very large proportion of our efforts on simulation. This means that we have devoted our efforts to reproducing in recognizable form certain aspects of human behavior and the performance of mechanical systems or a combination of the two. We have done this generally in order to make predictions, and we have been interested in the accuracy of predictions in order to assist our agencies or other policy-makers in making decisions. It is the aim of this paper to provide a brief review of some of the ways in which theory can be of assistance in improving the similitude of simulations, and consequently the accuracy of predictions and the wiseness of decisions.

Since there is a good deal of popular jargon which tends to imply that practical activities are useful while activities dealing with theory tend to be nonproductive, I intend to devote a part of this discussion to what might be called paradoxically a down-to-earth defense of these impractical activities—and to some extent I shall oppose what I would call crackpot realism with what might be termed realistic idealism. As Bertrand Russell has said, "Nothing is as practical as a good theory."

In very simple terms, theory is a general statement about the real world. In these simple terms, for example, the Pythagorean theorem is one of many consequences of the theory of Euclidean geometry. As such it makes its own general statement about the properties of right-angle triangles on plane surfaces, and has had tremendous

---

*Formerly Research Coordinator, Penn-Jersey Transportation Study.

Paper sponsored by Committee on Land Use Evaluation and presented at the 45th Annual Meeting.

practical influence in surveying and engineering. This theorem provides the basis for all the well-worn formulas of elementary trigonometry. There are two ways in which, however, we need to qualify this simple-minded definition of theory, and it is these qualifications which may tend to give the notion of theory some of its other-worldly character. First, when we say that theory concerns the real world, we have to include in that real world the minds and ideas of men. Thus, theory may deal to some extent with technology and concrete things on the one hand, and on the other hand with mental constructs which are seldom or never encountered in the physical world outside of men's minds until they have been written down. The real world of mental constructs is a very important one, and in the end has many practical applications. The extension of the trigonometry of measurement into trigonometric functions, for example, is the basis for other large parts of engineering. The second qualification is that a theoretic statement about the real world may not be, to the layman at least, a recognizable mapping of the real world, and the nature of the correspondence between the theory and the world and the consequences of the theory may not be readily expressible in everyday language. This sometimes makes it difficult for the layman to conclude at first glance that the theory is in any sense realistic or has any practical consequences.

There is of course an intimate relationship between theory and science or between the verification of theory and the scientific method. Since theories consist of statements about the real world, their degree of correspondence with this reality can be tested. Where the real world in question is one of mental constructs, as in logic and mathematics, the testing may be of a special and somewhat different nature, based on internal relations between constructs. It is not in general a requirement of the development of conceptual systems and their theory that any direct correspondence with material phenomena should be established, but it has frequently proved to be the case that after short or long periods of development, such concepts have found important and unforeseen applications to theories of phenomena. This course of events is analogous to, but not the same as, the laboratory development of methods and devices which for a long while remain mere curiosities, but which ultimately become technologically important.

We live in an age of rapidly expanding scientific endeavor. Science and the scientific method are being newly applied to old systems of human thought such as ethics and philosophy, and in these areas, the boundaries of untestable contention are constantly being narrowed. We now know that because of the atomic nature of matter only a finite number of angels can dance on the point of a pin; we also feel greater confidence in the rigor and cogency of philosophy. At the same time, new groups of phenomena are becoming the subject of science. Testable rather than speculative hypotheses are developed about these phenomena, and these hypotheses are organized in increasingly unified systems, frequently of a quantitative nature. Some of this movement towards new applications of the scientific method is occurring in the social sciences, and in this field the two tendencies to reduce the area of philosophical and ethical speculation and to systematize our understanding of objective phenomena go hand in hand.

It is hardly necessary to review the practical ways in which the advances of science during the last two centuries have greatly increased man's control over his natural environment through the application of science to technology. It is more useful to point out, first, that not only is science successful in an objective sense, but also that it is widely accepted publicly and politically, as may be judged by the governmental and private resources which are devoted to it, and second, that the growth and prestige of science have not simplified but have complicated the distinction between practical and theoretical endeavors. The customers of the scientific establishment are basically interested in results and frequently have shorter time horizons than the scientists. This dichotomy expresses itself in the distinction between science and technology as disciplines, and organizationally in the distinction between mission-oriented research and theoretically oriented research. As Alvin Weinberg (9) has recently emphasized, the objectives of mission-oriented research are externally imposed upon the scientific community by the real or imagined needs of society and by society's control over expenditures, while the objectives of theoretical research are largely generated within the scientific community.

It appears that these distinctions, while valid and useful for analyzing and discussing the problems of science and technology, can be unduly overemphasized. Many factors tend to blur the differences. Technology and engineering themselves are becoming more scientific in their basic methods, and consequently engineers are becoming scientists. Mobility between the professions tends to inject mission-orientation and social responsibility into the scientific community, which was in any case never detached from these values. The tremendously accelerated pace of science tends to shorten the scientist's time horizon and bring it more into accord with that of the decision-maker. Finally, the complexities of real life which face decision-makers are driving them away from simplistic common-sense judgments and in the direction of a more comprehensive and quantitative approach to the problems which they face.

It is in fact the magnitude of the problems of societal control in a period of rapid change and development which is providing the impetus for the truly scientific development of the social sciences. Problems such as maintaining peace, feeding billions of people, reducing racial discrimination, and organizing great cities require powerful instruments of control over men and machines. These problems of control can no longer be resolved by an engineering approach which is overwhelmingly oriented towards physical, inanimate, machine systems. Engineers working in transportation planning must pay increasing attention to problems of human behavior, and it is rapidly becoming evident that the relevant behaviors are not only in the fields of transportation demand and transportation system utilization, but also in the field of land use development and locational choice. In a sense, therefore, the planning-engineering professions find themselves working on a frontier of science. This is the area of social behavior and social control, in which the application of the scientific method has been unduly retarded. In order to explore what implications this situation will have for their work, we must therefore take a closer look at some of the elements of this method.

We are used to the idea that man and the other higher primates are endowed with an innate curiosity which leads them to explore their total environment in an apparently insatiable but not entirely purposive way. There is usually no a priori identifiable useful payoff in some of the exploratory activities of monkeys and children, and one is tempted to make an analogy with the data-collection propensities of social science research and transportation studies. It is also perfectly clear, however, that in man at least, curiosity extends beyond the accumulation of data about the environment. First, even the childlike exercise of curiosity involves the exploration of cause and effect. The experimenter will employ some of the simpler ploys of the scientific method to discover what worked when and where. And second, there is frequently an effort to generalize; there seems to be a tendency to seek out analogies and similar situations in which earlier findings and elementary theories can be tested. Thus we have in a primitive form the four main steps in some classic descriptions of scientific method:

1. Induction: the collection of information and its organization into patterns;
2. Generalization: a restatement of the cause and effect relations behind the patterns, or a redefinition of the patterns themselves in a more abstract form which includes the observations as a special case;
3. Deduction: the search for new special cases previously unstudied, as suggested by the more general statement, or theory; and
4. Testing: a check to see whether the new cases perform as predicted—if not, the theory must be revised.

This schema, while useful for analytic purposes, does not correspond in its rigid division of steps with the way in which scientific investigations actually proceed. We will use these categories as a basis for discussion, specifically maintaining, however, that the classification is artificial and if pressed too far, actually harmful.

The testing of theories about the world of phenomena raises special problems in the social sciences which should be generally understood before we take up other aspects of the scientific method. In this discussion I use the term testing in preference to the more usual verification because in principle no theory can be established, but only disestablished. A theory does not have verity, but verisimilitude. There are of course many theories which are outstandingly successful and for which there have been an almost unlimited number of successful tests, while the unsuccessful tests are nonexistent or occur only under well-defined special conditions. An especially significant case in which a theory may be regarded as firmly established because no counterexample exists is the correspondence between the counting numbers of everyday experience and the invented set of positive integers as defined in modern algebra by the use of the Peano postulates or otherwise. The theory states that these two systems, one from real life and one from mathematics, have the same form, and no exception to this theoretical statement is known or is likely to be discovered. A second example is the Newtonian statement based on the laws of motion, the law of universal gravitation, and Euclidean geometry. This theory relates the mechanics of the real world to a mathematical system of differential equations which Newton, in fact, was forced to invent. Up to the point where by relativistic considerations the Euclidean geometry no longer obtains in real space, there are no exceptions to this theoretic statement, and it too may be regarded as firmly established. In this latter case, it is important to note that the correctness of the theory has only been established by eliminating or controlling extraneous factors such as air resistance which affect the theoretically defined unimpeded motion of observed bodies.

The special problems of the social sciences arise, as is well known, out of the difficulties of pursuing this experimental method in which most variables are held constant (the ceteris paribus assumption of economics), while a limited subset of possible variables is manipulated. Where a large number of variables is involved, this difficulty can sometimes be overcome if a large number of diverse observations is available to the researcher, but this is unfortunately not the case with regard to the study of the development and the manipulation of large urban areas. Here the case material is limited in extent, and experimental manipulation is both extremely slow and vastly expensive with regard to the aggregate phenomena. Experimental cum statistical methods are only possible with respect to smaller elements of the total system. In these regards, science as related to total development of the function of the urban system is in most respects analogous with astronomy, which has a few cases of major interest, subject matter which is inaccessible to experimental manipulation, and the capacity for studying in the physics laboratory elements which do not aggregate by simple addition into the whole. The conclusions of the science of astronomy are not yet, however, directly useful in the guidance of societal action on a large scale—even though expenditures on the space program exceed expenditures on urban development.

It may however be argued that the disadvantages of the social sciences in establishing an experimental method have been greatly exaggerated. This argument is advanced on the grounds that the most important tests of theories concern their ability to predict new phenomena or phenomena not previously studied in detail, and to extrapolate the effects of causes beyond the ranges in which the causes were originally observed. It is curious to note that the literature of engineering and the social sciences abounds with warnings as to the dangers of extrapolation. If we wish to use, as we are almost forced to do, the power of extrapolation of a theory as a test of its credibility, then this cautionary advice is a frank confession that the relationships being extrapolated have no theoretical basis whatever. From the point of view of testing theories, the social scientist should welcome rather than shun opportunities for extrapolation, since this will provide his main basis for justifying a theory or for designing improvements in it.

Any acceptance of this criterion for testing theory tends to indicate the ultimate futility of a complete reliance on induction for generating theories. Even in the physical sciences, a fairly thorough knowledge of a particular range of joint variation of phenomena does not guarantee any adequate knowledge of cause and effect or even any complete description of relationships outside the range of observation, and this is even more true of the social sciences. Most of us are thoroughly familar with the situation which arises when we get a good fit of a polynomial to a set of observed data, only to find that it behaves extremely erratically outside the range of observation. Poincaré (6) has pointed out that if we had a complete knowledge of a portion, however small, of a continuous function, we would have a knowledge of the behavior of the function over its entire range. We could attempt to reach this happy state by developing the function as an infinite series and fitting all its coefficients statistically. Unfortunately, this procedure requires an infinity of observations (and a larger infinity than N_0). Even more important, our observations must be free of errors of measurement, and our function must include all relevant variables, each of which must be measured. It is quite clear that even the process of induction from observed phenomena must be guided by theoretical concepts based on previous experience which will suggest the ranges and objects of observation and the character of the functions to be fitted. A simple engineering example might be found in the difference between the parabolic curve and the catenary. These curves arise under different circumstances in the construction of suspension bridges, and lie extremely close together over a certain range of the variables. To distinguish between them by induction would be a hopeless task, especially since the formula for a catenary would not intuitively occur to a statistician, yet in extreme cases the distinction is important for engineering design. The differences are well defined a priori on the basis of a theory which may have been suggested by observation, but which does not spring automatically from it. The difficulties of induction are conclusively delineated, in fact, by the difficulties which arise in social science research in selecting functions for curve fitting and interpreting the results. Linear models are most frequently used because of their simplicity, perhaps with the justification that the linear approximation to some unknown function is not unreasonable over the range of the observations. The function being unknown means that theory is out the window. Perhaps a polynomial is used as some sort of an approximation to a Taylor expansion of a function. In this case, the catenary is defined as a parabola. Where the choice of function is deduced from a priori considerations and not merely to satisfy goodness of fit, we are suddenly in the realm of deduction rather than induction.

Deductive thinking is of very high value in science. Examined closely, the antinomy between induction and deduction is somewhat artificial; on the one hand, induction is almost never initiated without some kind of prior theory, however naive, which suggests areas of investigation, relevant variables, and the form which functions might take—while on the other hand, if deduction is unsuccessful and does not result in a confirming instance of the theory on which it is based, then the contradictory evidence may be the basis for a new round of induction. But the importance of deduction as a part of the scientific method remains in spite of the partial artificiality of its separation from induction. At the start of the process of deduction, the investigator is forced to make a statement of a general nature about the real world; in other words, he must formulate a theory. The motive for this formulation frequently comes from psychological forces very closely related to induction and to the search for generality. To follow the processes of deduction suggested by the theory, the investigator must search out new areas or new modes of application of the theory. It is useful to him in defining supposed cause and effect or functional relationships and variables to be investigated. In considering, therefore, the nature and power of the deductive process, we are led naturally to the final and perhaps the essential part of our discussion of theory construction, that of generalization, or the actual formulation of the theory. We must consider this in the light of all of the processes and problems which have been discussed previously.

Generalization is the bridge by which the scientist or theoretician crosses over from induction, or the observation of reality, to deduction, or the testing of theories and their application to new phenomena. For this reason, I rather like the name transduction, which is sometimes applied to it. No matter how often this bridge is crossed in the course of a scientific investigation, the act of transduction always involves some invention on the part of the investigator. The psychology of invention in this field is intricate and fascinating, but a discussion of it is out of place here. The sources of this invention may, however, be better understood through a consideration of its inherent nature.

The construction or invention of a theory involves in essence a precise statement regarding formal relationships, usually including relationships of cause and effect. There is an infinity of possible formal statements of relationships which may be made in their most abstract form in the language of mathematics or logic. Such statements regarding relationships are purely formal and have no reference to the real world. Within the sciences dealing with concepts, they may in fact be developed quite independently of the real world. The problem of theory construction or invention is, then, to make the correct identification between a real phenomenon and a mathematical or logical statement regarding relationships. There are three possible ways in which this may be done, two of which are merely suggestive and one of which tends to satisfy rigorous scientific requirements.

1. An analogy may be recognized at the level of phenomena. Thus, for example, a city may be compared with an organism—say a jellyfish. This analogy is scientifically useless unless two conditions are met: (a) the comparative object (the jellyfish) must have a form which has been clearly and logically defined; and (b) the object compared (the city) must be unequivocally said to be theoretically identical. In this case, we have identified a correspondence of the third type below, but otherwise we have merely made a statement which is useful for heuristic purposes.

2. An analogy may be recognized as between a phenomenon and a mathematical or logical construct, but may indeed be very loosely defined. Thus, for example, the gravity formula recognizes an analogy between the decay of trip frequency with distance and the power function X^{-a}. This analogy is extremely crude, seizing as it does on the most obvious and easily manipulated of a host of monotonically decreasing non-negative functions. No statement of the gravity model, to my knowledge, states any causal relationships which would generate this particular function in preference to others. I think that we may designate an analogy between phenomena and a logical function as a homomorphism, meaning a similarity of form.

3. An important qualitative change is introduced if a scientist identifies a particular phenomenon as having a clearly defined logical form. The definition of form may have already been made either in the development of logic and mathematics and unrelated to phenomena, or in connection with the development of theory dealing with some other and perhaps completely unrelated phenomena. The use of formal statements pertaining to other phenomena is indeed often suggested by analogies between the phenomena themselves. On occasion, the study of phenomena and the formulation of ideas about cause and effect necessitates the invention of a new relational calculus. This has been the case in Newtonian mechanics and quantum mechanics. In any event, the essence of a theoretical statement is to identify an isomorphism (identity of form) between a set of phenomena and a logical or mathematical relational system. Thus, the Schneider (7) model for trip distribution, in contradistinction to the gravity model, makes a rigorous statement that the decay of trip frequency is isomorphic to the negative exponential function and consequently also to the radioactive decay of fissionable elements; and Schneider identifies the precise cause and effect relationships on which the isomorphism is based.

It should of course be clear that the borderline between homomorphisms and isomorphisms is blurred, partly because it refers to the motivations and psychology of the scientists. A theory which is in fact generated as an analogy must be presented as an isomorphism, and a badly conceived isomorphism may turn out to be only an analogy. The appropriate distinction can be made only upon close examination of the theory and of its results.

In formulating a theory to serve as a bridge between induction and deduction, the analyst has a number of guides as to desirable features of his formulation. Some of the most significant criteria tend to conflict with one another while others reinforce each other, depending on circumstances. All arise out of the general characteristics of the scientific process as we have outlined it.

The outstanding criterion, of course, is that the theory should be correct, that is, that the theory if testable should pass its tests successfully. This is the basis for the essentially practical nature of science, that is, that it says true things about the real world. We have seen that such truth is impermanent, always awaiting contradictions. If these arise, it frequently is retained by circumscribing the generality of the theory, limiting the circumstances in which it applies, and creating new and more general theories to apply to other combinations of circumstances. This criterion of correctness is in general overriding. However, practical considerations frequently lead to the generally indefensible practice of applying theories which have been inadequately tested or which have known errors.

Thus a theory, if testable, must pass its tests. In fact, the untestable theory is a nontheory, and testability is therefore an important criterion in theory construction since, like Milton, science has "no use for a fugitive and cloistered virtue. . . . " There are of course examples of very important scientific theories, such as Einstein's theory of relativity, which when published appeared very difficult if not impossible to test. These difficulties in relation to a reputable and exciting proposal often serve as a spur to experimental work. In the field of social sciences, however, this particular type of nontheory has two other manifestations. One of these is a normative monitory description of how things should be done; the other is a literary or pseudostatistical description of the real world. The fact that these things are merely masquerading as theory can easily be exposed by searching for critical tests which could deny their validity. If such tests do not exist, the so-called theory is in fact a nontheory. Occasionally tests acceptable to the authors of the theory will be so circumscribed with restrictive conditions and assumptions as to expose the fact that the theory is of extremely limited application and has dubious stature.

The testability of a theory is a special case of a more general property of useful theories—productiveness or fruitfulness. Important theories in the development of science not only answer the problems posed in the initial stages of induction, but are pregnant with consequences which are only dimly seen by their inventors and which lay the basis for a wide variety and a great number of deductive experiments. Axiomatic systems in geometry, algebra, and logic exhibit this property. The tremendous accomplishments of modern mathematics follow (although not effortlessly) from a very limited set of carefully considered initial assumptions. Similar examples exist outside of conceptual systems. The quantum theory, which was invented to explain anomalies in black-body radiation, has found innumerable applications to phenomena as diverse as photoelectricity and solar spectrography, and is in fact a key element in all modern physics. The social sciences and the planning-engineering professions are somewhat lacking in such key theories, but some nominations could be made. These might include, for example, marginal substitution and general equilibrium concepts from economics, and applications of general systems theory. In any event, while it is somewhat difficult to define the process by which a theorist comes upon a fruitful theory with many applications while attempting to solve a more limited and more particular problem, it is apparent that solutions of this type are unusually desirable. At the least, theory builders should have this objective in view, especially since this state of mind leads to stripping any problem to its most essential elements, and thus may simplify as well as lead to greater generalities.

Simplicity is in fact an ancient and honored criterion for choosing between otherwise equipotent theories. Occam's Razor, named after a fourteenth century English philosopher, dictates that theories should contain the minimum possible number of hypotheses, and many of the more durable theories elegantly exhibit this characteristic. Because of the large number of conditions, relations, and variables which occur in social science research, this condition is difficult to meet here and frequently conflicts with requirements of realism and testability. It is nevertheless a desirable characteristic, not only for reasons of elegance, economy, and generality, but also for practical reasons which will be discussed later. Here there is a special pitfall which social science researchers can dig for themselves by the use of modern computational techniques. It has been suggested that, had computers been available at the time of Copernicus, the ease of computation of epicycles might have removed the practical difficulties which led to the construction of the elegant and economical heliocentric theory and the Newtonian theory of celestial mechanics. By the use of computers in the descriptive system of Ptolemy, navigational tables could have been constructed to any required degree of accuracy, and the practical impetus for the Copernican and subsequent revolutions would have been removed. I feel that we fall into the same trap when, as with the introduction of K-factors into the gravity model, we constantly patch up a nonexistent or inadequate theory with computational amendments.

A requirement which follows from testability and which is necessary to it is the requirement of manipulability. The experimental method in the social sciences is, as we have said, forced to rely on paper experiments, and for these our professions commonly talk about the use of models. To quote Harmer Davis, "A model is a smaller copy of the real thing, as the woman said about a model husband." This pointed definition does not permit us, however, to distinguish between a mathematical model and a simulation model on the one hand, nor between a simulation model and a theory on the other hand.

As we have emphasized, there are in principle distinct sources of an understanding of cause and effect in the real world and of formal representations of relationships in the world of mathematics and logic. Science is in many respects an effort to establish isomorphisms between these distinct realms. If we refer to a linear programming model of warehouse location, we are referring to just such an assertion about an isomorphism. We might then be correct in speaking of a mathematical model of warehouse location. Frequently, however, people speak of the linear programming model, and more generally of mathematical models in the abstract without relation to any particular real world phenomena. I would submit that this application of the word model is incorrect, though lamentably ineradicable, because the mathematical linear programming model is not a model or a smaller copy of anything.

The distinction between a theory and a simulation model is somewhat more subtle and difficult. A theory in fact could also be said to be a logical or mathematicl model of the phenomena to which it refers. It is smaller, it is a copy, and it is of the real thing. Yet this identification of a theory with a model somehow goes against the grain. On the basis of very serious consideration, I have redefined models as they are used in the simulation of social and economic events in a way which tends to provoke outraged reactions, but which I believe withstands serious examination and criticism: a model is an experimental design based on a theory. Let us examine the implications of this definition somewhat more carefully.

As is well known to workers in our fields, there are many theories which are testable in the sense that a critical experiment can be designed—but which remain untestable in the sense that the data requirements are for practical purposes excessive and involve presently unobservable variables, or perhaps most important, they cannot be cast in a form which will fit into a computer and run economically. These practical considerations do indeed provide a spur to all kinds of experimental ingenuity, and they should by no means dominate the process of theory construction.

In the process of developing a theory, there are many applications of experimental design in which the theorist must invoke models. First, in exploratory or inductive investigations, he is quite apt to use a severely truncated or patently inadequate experimental design such as a multiple regression model to explore relationships and to provide information as to the direction of his next transductive steps.

1. In a more developed form he will use a model more closely corresponding to theory inductively to establish the parameters of relationships.
2. He will use a model for testing in the deductive sense in order to determine the applicability of his theory under a wider range of conditions.
3. Used scientifically in a context of projections, the model will provide experimental evidence as to the consistency of the theory and possible inductive evidence as to the sensitivity of the real world to changes in conditions.

It may be a matter of scientific but not practical indifference to the scientist that the projective use of models also is important to decision-makers.

One may choose to make a distinction between the value of theory building and the value of experimental work with models, imputing a higher value to the first of these activities. However, in the tradition of British and American experimental science, the theorist usually has some responsibility for making feasible the tests of his ideas, and it is only the boldest and most brilliant innovator in pure theory who can expect others to accept a division of labor in which they will devise feasible tests for his impractical fomulations. It is this experimental difficulty which often leads to emphasis on the false dichotomy between theory and practice, which can only be overcome by a long-run view of the value of theory and by a nice sense of the potential contributions of new theories whose testing and application may appear outrageously difficult.

There are other criteria which may or may not be useful for the construction and selection of theories but which are frequently in the minds both of scientists and practical people. We have mentioned that for a variety of reasons a theory may not correspond directly with intuitive and popular ideas about the nature of reality. In this case, the theorist or scientist may be accused of being unrealistic and may feel a social obligation to change or tone down his theory in the direction of greater realism. Such a compulsion grossly distorts the role of the scientist, which is to identify a genuine isomorphism between the behavior of the real world and a set of mental constructs. Frequently he has to invent the mental constructs in order to disclose the isomorphism. Many of the most pregnant ideas of the physical and biological sciences, such as the quantum theory, the theory of relativity, or the independence of heredity from environment, run counter to widely held and deeply rooted popular ideas. Yet, without the discovery of these theories and their application to everyday life, the world would have given up a great deal of progress. A search for naive realism is counterproductive in science.

Frequently even though a naive demand for realism may be abandoned, the critics of science will continue to take refuge in an unthinking insistence on comprehensibility. In the field of social sciences, this insistence is based on two circumstances. First, every critic is a member of society, a user of cities, and a participant in the political process. Hence he feels intuitively that by virtue of this special status he and most other informed people ought to be able to understand directly all of the theories which purport to define the operations of society, of cities, and of politics. In my view, it would be equally ridiculous to say that because we are all made of protein, we should understand at a glance the theories of molecular biology. A second circumstance resides in the fact that a great deal of social science research is conducted in such a way that the scientists are close to the administrators, the administrators are close to the decision-makers, and the decision-makers are close to the voters, with no clear separation of function. Because of the personal and normative nature of the communication between these groups, each link in the chain feels that he ought to know all about what the adjacent link is doing. We may contrast the somewhat more impersonal relationships which govern research and development in industry. The laboratory scientist may understand solid-state physics in detail. The corporation executive will understand the main directions of this research and its potentialities. The sales department understands the capability of the resultant product, and the customer chooses in the market place between the products of competing technologies and competing companies. The man in the street could not care less about the crucial role of, say, quantum mechanics in the production of his transistor radio. Probably when social science theories produce as effective results as quantum mechanics, the administrators, policy makers, and voters will be less inclined to ask questions and more inclined to judge by results.

A possible requirement for theory which requires brief mention is more likely to be generated by the scientist than by the layman. As a result of the complexity of social phenomena which requires holding other things constant, and as a result of the drive for generalization which is inherent in theory building, there is a considerable drive to create theories which are comprehensive. This drive encounters resistance on two fronts. A comprehensive theory may in certain cases become so general as to say nothing about everything. Even if this is not the case, the more comprehensive theories may be the most difficult to manipulate for purposes of testing and application. An important part of theory building is therefore a nice sense of discrimination as to when comprehensive theories are necessary and when they may be appropriately avoided by discretion in the subdivision of the problem into manageable parts. In policy-related sciences improper subdivision of the policy-making problems may result in suboptimization, but a subdivision of the problems of the real world and its functioning for purposes of study need not entail this danger.

In the preceding sections of this discussion, I have developed my ideas with regard to the scientific construction of theory, mainly with respect to the problems of simulating events in the real world of mass behavior, in the use of transportation facilities, and in the choice of locations, even though this concern has been in the main implicit rather than explicit. There are two other areas related to public decision-making in which theories of a different kind will have to be developed. Transportation and planning literature already recognizes the need for the development of more general theories of decision-making. In crude terms, the questions to be answered by such theories are: what are we planning for; what trade-offs are involved in the public decision process; and what values will our plans satisfy? In more sophisticated terms we are dealing with difficult problems of public discount rates, collective consumption, spillovers and externalities, the aggregation of utilities, and the reconciliation of conflicting interests. It is hoped that the improvement of theories and models in this general area may be expected to reflect backward into the planmaking process so that sketch planning procedures are replaced by optimizing procedures, and optimizing is not limited to narrow engineering criteria but is extended to the most general of social objectives. I think it is also predictable that as we explore the problems of decision-making, planning, and optimizing more thoroughly, we will discover that there are ferocious computational problems which arise in the design process as a result of the huge combinatorial variety which exists in the possible combinations of policies and future conditions. Our fraternal theorists in the field of mathematical programming may be able to make contributions of a theoretical nature with practical applications which are related to the needs of decision-making. It is also probable that a clearer formulation of these needs will influence this development of what are essentially design models.

We have now reached the vantage point of a somewhat shaky and perhaps imperfect understanding of some of the processes of science, from which we may view the needs and accomplishments of experimental simulation of transportation systems and land use systems and the behavior of their users. I will not here belabor the point which is now becoming widely accepted in principle—that in many policy-making contexts we are dealing with these systems not independently, but as a part of the larger urban metropolitan system or regional system. I will emphasize the fact that most theories of locational behavior contain ideas about transportation costs and convenience, and consequently that locational models must contain as submodels some replication of the transportation aspects of the system. It will also prove useful in the discussion which follows to consider the salient features of all these problems together from the point of view of theory construction, drawing freely upon examples from any field wherever they may be appropriate.

The range of our interest in these phenomena covers a wide span from very large and complex total systems through subsystems which may be defined in engineering terms, in social and economic terms, or in spatial terms, down to the smallest elements of the system. These last may be mechanical components, but the greater interest attaches to decision units—a man driving a car, a family looking for a home, or a corporation deciding to build a new establishment. At each of these levels, different problems arise regarding the appropriate form and content of research.

The broadest view, of the overall system as a whole, is probably not in itself highly productive, but it is a starting point for certain applications of general systems theory which later affect our view of the components and the elements. General systems theory with respect to the total urban, metropolitan, or regional system will ultimately play a direct role in decision models. Meanwhile, it can be particularly useful in defining the appropriate limits of a system and in guiding the structuring of the problem in such a way that its decomposition into subproblems dealing with subsystems will entail a minimum of distortion. Up to now in both transportation and land use analysis these two problems have been approached largely by intuition and induction. I do not feel that the results have been seriously wrong, but a systematic and better informed approach might provide some surprises and prove a useful guide to research design.

With respect to subsystems properly defined and considered as systems in their own right, general systems theory may very well contribute powerful methods for dealing with system stability as a planning objective and with homeostatic or equilibrating tendencies within systems as handles for both planning and analysis. My own intuitive feeling is that concepts of equilibrium animate a great deal of research and theory in land use and transportation analysis, but that these concepts are inadequately explored and not sufficiently explicit. For example, transportation analysis and the assignment of traffic to networks with capacity restraints imply a whole pattern of equilibrating behavior on the part of individuals which may or may not lead to system equilibrium and may or may not be related to various forms of optimization. These questions have been very lightly explored by brute force iterative methods in modeling experiments, and their full implications remain to be seriously examined. In land use growth model simulations based on trend data, there is also a set of unexplored assumptions about tendencies to equilibrium. Whether such an equilibrium exists or ought to exist has in fact been very slightly examined in theory. Needless to say, one-shot sketch planning or design models and "instant cities" such as Ira S. Lowry's Model of a Metropolis (4) are constrained to use either simultaneous determination or optimizing, and it seems likely that the former method contains some optimizing assumptions in its behavioral parameters. More generally, I feel that land use behavior as well as land use system performance can hardly be explained without a consideration of land marked equilibirum and simultaneous determination—all of which pose major problems for system theory.

There are a number of interesting problems which arise out of the communication between subsystems and between elements and subsystems and out of the mechanisms by which equilibrating, disequilibrating, and determining forces are transmitted to and from decision units. The organs of the body communicate information leading to action by nerve impulses and those maintaining homeostasis by chemical messengers; what are the messengers in a large city or region? Many of these questions will arise again in the discussion of the behavior of decision units below, but there is some advantage in taking an overview in the context of systems. It is quite apparent that the generic name for these messengers will be information, and it seems quite likely that some gains for theoretical clarity will be achieved if a systematic application of communications theory can be made about the diffusion of information through and about the systems under study. The applicability of this concept is already apparent in the most elementary consideration of the stability of traffic flow systems, and these ideas can probably be extended to land use systems and larger transportation systems. Considered in the communications context, there is some merit in merging the study of decision units with a priori considerations from different disciplines as to what information is likely to be important and available. At one extreme this type of merger leads to a consideration of the individual's reaction to the visual environment as developed in studies by Lynch (5) and others. At a different extreme, economics suggests that prices are the messengers by which important economic information regarding, say, the housing market is transmitted. Between these extremes lie many combinations of phenomena which are observable, influential in behavior, and to some extent predictable as consequences of other developments.

The importance of prices as a messenger and of the allocation of money to different purposes (i.e., of economic behavior) in private decision-making is so great that it deserves special attention. It is a curious fact that in spite of this a priori importance of monetary phenomena, they have really received relatively little emphasis in transportation and land use planning and analysis. For somewhat understandable reasons, transportation planners have been reluctant to explore the importance of pricing policies in alternative transportation systems. Surely, however, this reluctance should not extend, as it frequently does, to the omission of cost factors and the exclusive emphasis on time-distance which is frequently found in network analysis, trip distribution, and even modal split. Fortunately, this default is not universal. In land use analysis, the problem is perhaps even more severe. Housing rents and values are the medium through which consumers communicate with each other their willingness or unwillingness to compete for space, and more commonly land prices are the medium of communication in the competition of residential, industrial, and public uses for land. Yet in the research field, housing value and land prices very seldom appear as variables. So pervasive is this omission that expensive and otherwise useful surveys of locational, social, economic, and housing variables by the Penn-Jersey Transportation Study and the Tri-State Transportation Committee are partly vitiated by the failure to inquire as to housing value or rent. It must be admitted that the collection of these data and especially of land value information in a research study is fraught with difficulty, but

I believe that there is a more serious reason why these values have been neglected in spite of strong theoretical reasons for their inclusion.

If values (prices) are made explanatory variables leading to changes in the behavior of decision units, then future applications of the same theory and its derivative models require that these values be projected under new circumstances. The theorist then faces an ugly dilemma. If he chooses to predict future prices by means of proxy variables, he must build a purely descriptive model for this purpose which contains no ideas about cause and effect; and this being the case, he might just as well have left prices out of the original analysis and included the same proxy variables, admitting from the outset that his theory was in part purely descriptive. If on the other hand he takes the importance of these economic variables seriously, he must face the difficulty of reconstructing a complete market through some form of simulation. This reconstruction is complicated by the existence of submarkets, institutional stickiness, imperfect dissemination of information, and probable lags in equilibrium. If economic considerations were largely peripheral to the theory of land use and transportation systems, there would be less objection to taking the easy way out of this dilemma. I believe, however, that these considerations are so central that economic models must in the future be added to the implementation of transportation and location theory at full scale. This approach will involve much deeper consideration of equilibrium tendencies than was previously suggested, and perhaps a much more serious look at some aspects of the behavior of decision units.

Before turning to a discussion of the theory of the behavior of decision units, I must emphasize a vital distinction between the study of that theory and its application. To a very large extent, the study of the behavior of decision units can be undertaken independently of the simulation of system and subsystem performance which has been the subject of the prior discussion. This is true because at the moment when we examine the actions of decision units, the systems in which they are embedded have already performed their functions, interacted with each other, and thereby generated the environmental conditions and information of which the decision unit has knowledge and on which it acts. In this analysis, the experimental approach consists of searching out instances in which the environment and its informational content differ significantly from other environments, or the decision units differ significantly from other decision units, so that the general application and fruitfulness of the theory may be examined. When, however, the behavior of decision units as understood on the basis of such an analysis is to be explored experimentally under changed assumptions as to policies and technology, an entirely new situation arises. We can no longer assume that various sets of decision-makers are independent of each other. Each reacts with the environment and creates changes which result in messages reaching other decision-makers. This interaction, which is irrelevant to some analyses, becomes critical in system simulation. I thus assume that system simulation and decision analysis interact strongly with each other and that each is necessary for the other. But as a matter of research emphasis, I would give short-term priority to system simulation on the grounds that relevant experiments to test our understanding of the behavior of decision units probably cannot be performed without it.

Engineers and planners are vitally concerned with the behavior of households and business establishments in making use of the transportation system and in making locational decisions. Such behavior is the source of transportation demand. Private decisions in respect of automobile ownership, location, and new construction in the aggregate greatly influence the development of cities and regions. Finally, I am sure that if we understood thoroughly the whole constellation of decisions made by individuals and firms, we could understand at the same time the extent to which various urban arrangements satisfy their needs. Such an understanding is a vital key to producing plans and policies which will best serve the public interest.

Some of the differences between practicing planners and engineers can be traced to their different approaches to decision-makers' needs and preferences. The planner typically approaches the problem from the viewpoint of normative standards of behavior and social welfare. This is in part based on notions of minimum socially acceptable levels of welfare and in part upon an emphasis on the externalities of individual behavior—i. e., on the effects of one's behavior on others. These notions are linked with strong ideas of social control. The engineering approach tends to be more adaptive. Individual behavior is regarded as being largely self-motivated and not widely amenable to control. In dealing with supposed patterns of behavior as necessary inputs to engineering estimates, the engineer approaches the problem with naive concepts of motivation and of measurement. Neither planners nor engineers are in general well trained in the intricate issues of choice behavior, and present-day economics, sociology, and psychology offer little which is of general applicability to the problems which they face. The following remarks are therefore observations on a dilemma which will ultimately be resolved only by training people and developing methods which embody a combination of all of these disciplines in a new format.

The basic theory of choices by individual decision units deals in terms of alternatives and trade-offs, yet if we examine transportation and locational theories, or models, we find that these trade-offs are very deeply buried, if indeed they may be presumed to have been considered seriously at all. Since the same thing is true of econometric models in related fields, this is not a particularly telling criticism in terms of past performance, but it is clear that it may constitute a barrier which will have to be removed before a great deal of progress can be made.

Much of the difficulty concerns observation and measurement, and perhaps this may best be illustrated with reference to the theory of industrial and commercial location. Industrial location in particular has long been very carefully studied by locational economists and regional scientists, and interregional locational theory is a particularly well developed field. In this location theory three factors are particularly important: internal economies of scale which depend on the size of establishment; external economies of scale or agglomeration economies which depend on the sizes of the geographical assemblages of activities in which the establishment is located; and locational costs which depend mainly on the cost of land and the costs of interaction. In the complicated urban metropolitan scene, these economic variables turn out to be very difficult to define, more difficult to measure, and still more difficult to value. While it may be well known, for example, that the garment industry has large agglomeration economies and is sensitive to its accessibility to a particular labor force and to the cost of industrial space, these variables and their relationships are not well defined. The interaction requirements of offices and the agglomeration economies of retail trade establishments are also imperfectly understood. While these ideas enlighten a good deal of research design, anyone who has tried to set up an industrial or commercial survey knows that it is very difficult to tie them down specifically. Because of this situation and for allied reasons, it is beginning to appear that in spite of the much more sophisticated work over many decades in industrial location, the problems of residential location are more tractable and amenable to sound solution.

In the area of consumer behavior, some difficulty is introduced by the fact that certain decisions are made by individuals, others by households, and still others by individuals in a household context. These difficulties must be faced in research design, but they are relatively minor compared with other more obvious problems. One which has been both recognized and ignored (often simultaneously) is that of aggregation. Some researchers, perhaps moved by data difficulties, are inclined to deal with the means and medians of areal aggregations of data. This method of work is almost enforced by the form of the availability of published census data in certain cases. There is clearly here a latent conception that area averages represent some sort of aggregation of behaviors, but the implicit rules of this aggregation are not explored, and frequently the assumed behaviors are not fully defined. The gravity model of trip distribution clearly takes this approach at a descriptive level, while multiple regression models of modal split may be but a step closer to postulating real cause and effect. The Schneider model of trip distribution postulates more explicit behavioral patterns and works with areal aggregate data. In practice, however, this model reveals unexplained variations in decision behavior because it requires an area-specific determination of the proportions of long and short trips. This specification amounts to a statement that the behavior in the model is incompletely defined.

Those who avoid the implicit assumptions of working with grouped area data by using individual or household observations encounter another level of difficulty which helps to elucidate the problem of aggregation. Behavioral models of individual and household choice invariably produce tests in which only a small part of choice behavior is adequately explained. Typical coefficients of determination are in the range of 0.15 to 0.30. While this may mean in some cases that the models and theories employed are inadequate, it is more likely to imply that behavior is influenced by more or less unobservable cultural and psychological factors which (at least at any one time) may be statistically distributed in the population. The delicate problem of research design is to know when to stop trying to identify these factors and when to introduce assumptions about the statistical nature of their distribution in the population. After the first recourse is exhausted or while it is being further developed, it is apparent that the nature of the assumptions about the statistical distributions of behavior around their observed statistical means may strongly affect the characteristics of their aggregation. As a simple example, I have demonstrated elsewhere that if Schneider's L parameter (7) is assumed to have a certain statistical distribution rather than being fixed, his model converts readily to a modified gravity model or to a combined model. Certainly considerable statistical expertise will be required to explore this problem further.

One of the more subtle and neglected aspects of the analysis of decision units is the role of the history of the unit in its behavior. To some extent, the history of certain units is implicit in their state description—a family head aged twenty is probably recently married. But other and more subtle historical aspects may be overlooked. It is quite clear, for instance, that the history of industrial establishments is related to their tendencies to relocate, and the ethnic background of many population groups is related to their choice of residence. It has even been reasonably suggested that consumer choice of mode of travel is related to the individual's history in learning to drive. These historical aspects of the behavior of decision units have two very important relationships with more general systems analysis. The historical aspects of decisions are closely tied to the extent of the lags in movement toward system equilibrium, and only systems in which the history of decision units is unimportant will rapidly achieve equilibrium. At the same time, the introduction of these histories is a means of dealing quite explicitly with trend data, without at the same time building into the theory an assumption that trends will indefinitely continue. It should be apparent that this historical approach does not lend itself to easy application to aggregate data, at least in analysis. And if the histories to be considered become very complex, then Monte Carlo methods are almost required for any projection simulations.

In the light of the foregoing incomplete review, we may justifiably conclude that a theoretically sound and scientific approach to systems simulation of transportation and land use will require a great deal of rethinking of our theory of decision-units' behavior.

Let us now take a brief final view of the workaday implications of the type of program that has been sketched. The essential elements of this approach are six in number.

1. Since sound theory has so much to offer for practical progress, the work should be organized on a scientific rather than a mission-oriented or technological basis. We would thus also avoid the dangers implicit in harnessing these activities to suboptimal policies, and rely on the social and policy motivations of the scientists to maintain a well-directed drive toward ultimate application.

2. We would view these problems as related to certain real world systems and would deepen our efforts to achieve successful theories of the operation of those systems.

3. We would give appropriate recognition of the need for the study of the behavior of decision units in the context of larger systems which create their environment.

4. We would give explicit recognition to the theoretical problem of communication between the systems, subsystems, and decision units.

5. We would recognize that the scope of these investigations will require the unification of parts of different disciplines in institutions and in individuals.

6. We would recognize that in a specially defined sense this work is experimental in the best traditions of experimental science and that the experimental method will require special conditions for success.

It would seem that the scale of these problems and their importance for long-term policy development tends to argue against scattered research in connection with specific projects. Such projects in any case tend to impose their mission orientation on individual researchers. The resulting tension between the desire of the researcher to satisfy his scientific conscience and the desire of the management to get the job done sometimes borders on the tragic, or the comic. In any event, the problems are of general and national significance, and if worthy of consideration should not be charged to local or to special-purpose studies. It is also apparent that the variety of ability and knowledge required for an assault on these problems can rarely be assembled even in a large study of an ad hoc nature. Consequently, many such studies are repeating the work and perpetuating the errors of other studies for lack of resources to go further and try new methods. Finally, there are serious difficulties of communication within this scientific community which result from the excessive fragmentation of effort.

Special attention should be devoted to the requirements of the experimental method in this field. Consider, for example, designing a laboratory for social, engineering, and planning research. Instead of white mice, our experimental material is extensive data about metropolitan areas and regions. These data must meet certain rigorous standards and be well organized and accessible. Our main experimental tool is probably the computer, but this will include the software or operating programs which embody many or most of the elementary processes of simulation and analysis which we have discovered. Our experimental design is a model or group of models based on theory and using experimental material (data) and experimental tools (computers and software). In any good experimental design, we are apt to discover that some special-purpose tools will have to be made—in this case new programs will have to be written and in some cases new data collected. The essential aim of an experiment will be to make critical tests of theories by good experimental design and thus to decide, for example, on a clear definition of the relative merits of the gravity model, the Schneider model, the Tomazinis model, and the Harris model of trip distribution (2, 7, 8). The essential ingredient for progress in addition to all the niceties so far discussed is quick turnaround so that experiments may be rapidly executed once they are designed. I would estimate that under current conditions, with practically no standing stock of data and widely diversified programs, the turnaround time on experimental work of this type is roughly three to five years. This time should be reduced by a factor of three or more, and the content of the experiments should be far more conclusive than it is today.

I believe that these standards, both of theoretical excellence and mechanical performance, are achievable and that if achieved they will have tremendous payoffs in improved planning.

REFERENCES

1. Carrothers, Gerald A. P. An Historical Review of the Gravity of Potential Concepts of Human Interaction. Jour. Amer. Inst. of Planners, Vol. 22, 1956.
2. Harris, Britton. A Note on the Probability of Interaction at a Distance. Jour. of Reg. Sci., Vol. 5, No. 2. In press.
3. Harris, Britton, ed. Urban Development Models: New Tools for Planning. Special Issue, Jour. Amer. Inst. of Planners., Vol. 31, No. 2, May 1965.
4. Lowry, I. S. A Model of Metropolis. Memorandum RM-4035-RC, The RAND Corp., Santa Monica, California, August 1964.
5. Lynch, Kevin. The Image of the City. Joint Center for Urban Studies, Cambridge, Massachusetts, 1960.
6. Poincaré, Henri. Science and Hypothesis. Dover Press, New York, 1952.
7. Schneider, Morton. Gravity Models and Trip Distribution Theory. Papers and Proceedings of the Reg. Sci. Assoc., Vol. 5, 1959.

8. Tomazinis, Anthony R. A New Trip Distribution Model. Paper #347, Highway Research Board, Washington, D. C., 1963.
9. Weinberg, Alvin. But Is the Teacher Also a Citizen? Science, Vol. 149, 601-6, August 6, 1965.