Mind and Brain Sciences in the 21st Century

edited by Robert L. Solso

Sperling, G. (1997). The goal of theory in experimental psychology. In R. L. Solso (Ed.,) *Mind and Brain Sciences in the 21st Century.* Cambridge, MA: MIT Press. **Pp. 253-264.**

A Bradford Book
The MIT Press
Cambridge, Massachusetts
London, England
1997

13

The Goal of Theory in Experimental Psychology

GEORGE SPERLING

Aesthetics and Utility in Science

When we evaluate a scientific theory in terms of its utility for a particular task, there is little problem in deciding whether a theory is good or bad. The question becomes simply: Is the theory useful or not? For example, when electromagnetic theory is used to design a generator for producing electricity, or a psychological testing theory is used to design a better test for predicting an individual's success in a particular training program, the theories are evaluated according to how well they meet their goals. There also are secondary practical goals for theories, such as serving as mnemonic aids to facilitate the recall of experimental facts. We do not consider these separately.

Much of science is not directly practical. It serves, rather, to satisfy our curiosity about ourselves and our universe. For example, while it is fascinating to have learned that the present universe originated in a big bang 17 or so billion years ago, such knowledge does not have any immediate practical applications. This is an example of the purely aesthetic pursuit of science. Of course, some of such knowledge may ultimately prove useful, so we can regard the aesthetic pursuit of science as being merely one extreme on the continuum that runs from immediate to infinitely deferred utility.

For brevity, I will designate as aesthetic the pursuit of problems where there is no immediate practical application by which to evaluate a theory. While it is difficult to estimate precisely what fraction of research in experimental psychology has been aesthetically motivated, I suspect that it is large—

This article is an elaboration of Sperling, G. (1978). The goal of theory in experimental psychology. *Bell Laboratories Technical Memorandum TM-78-1221-12*, July 26, 1978. Murray Hill, NJ: AT&T Bell Labs.

larger than most psychologists would care to admit. Therefore it is reasonable to explicitly consider aesthetic goals for theories, and also, incidentally, for theoretical enterprises in general, for example, choice of problems, style of presentation, and so on. What precisely is the goal of theory in the realms of deferred utility and pure aesthetics? I propose that the goal of theory is to provide the best description of a phenomenon (or class of phenomena) at a particular desired level of complexity. Experiments and careful observation provide the critical phenomena to be described.

Lest the reader's hopes be prematurely raised, be warned hereby that the terms "phenomenon," "theory," and "complexity" will not be defined here—they are assumed to have their ordinary meanings. The aim is not to make fine distinctions but to call attention to very general characteristics of psychological research that might be interesting and perhaps useful to readers regardless of the particular, idiosyncratic meanings these words may have for them. Nor is the problem of criteria to evaluate the "goodness" of theories treated here. The purpose of this chapter is not to rank existing theories but to expose implicit assumptions and to suggest profitable new directions for theorizing by psychologists.

Theories in Classical Physics

We begin by observing that psychology and biology are fundamentally different from a discipline such as classical physics which all too often has been inappropriately used as a model for science as a whole. In classical physics, complex properties are predicted from elementary principles because the principles are relatively exhaustive for a particular situation. Classical thermodynamics is a noted example. The various relations between the mass, temperature, pressure, volume, and specific heat of a gas are derivable as consequences from relatively simple assumptions about the masses and movements of the molecules of which the gas is composed. On the basis of these principles, precise prediction about the behavior of gases is possible. These marvelous predictions are restricted, however, to gases. Prediction about the change of state from gas to liquid requires radically new principles.

Gravity is perhaps the most famous example of classical physics. Newton's law states that the force F of gravimetric attraction between two objects is proportional to the product of their masses (m_1m_2) and inversely proportional to the square of the distance d between them: $F = m_1m_2/d^2$. It is simple and incredibly accurate. When there are several objects in motion and they interact via gravity, the motions can be very complex. By working out the consequences of Newton's equation in great depth, one can predict to the nearest second eclipses that will occur hundreds of years in the future. Complex properties derived from simple principles occur throughout classical physics: the simple relations between electricity and magnetism ultimately govern the design of complex electrical circuits; the simple principles of optics govern the design of exceedingly sophisticated multicomponent lenses; and there are many other examples. In these cases, physicists start with a few assumptions and soon are involved in incredibly complex consequences. Insofar as physicists persevere and are able to calculate the consequences, they find that their predictions are quite accurate because no important new principles come into play. The original assumptions were relatively exhaustive. In these cases, classical physics has achieved theoretical simplicity at the cost of computational complexity.

The Discovery vs. the Invention of Theories

Physicists say they are discovering fundamental laws of nature—as though there were laws, like Easter eggs, lying in hiding, just waiting to be discovered. Scientists invent elegant computations that accurately describe nature. The laws exist only in the physicists' minds and writings, not in nature. To illustrate the problem, consider the question: Is light best described as particles or waves? Then consider the analogous question about a newly discovered minianimal called a "mouse." We wish to describe the mouse in terms of previously known animals no smaller than dogs or rabbits. We observe that a mouse has some rabbitlike behavior. It is very fearful, it is momentarily motionless when threatened, and it lives in holes in the ground. On the other hand, when it comes to food, the mouse has doglike preferences. So for some behaviors, the rabbit model of the mouse works best; for others, the dog model. But the mouse is not a mixture of rabbit and dog—it is a mouse! The rabbitlike and doglike descriptions are invented to aid the human conceptualization of a new entity. And the same is true of the phenomena of particle physics. These phenomena are given names, and described by analogy to more familiar concepts, with equations developed in other contexts. But a phenomenon such as light is neither a wave nor a particle, it is *light*. The descriptions or laws that have been invented to describe it have been invented by, and for the convenience of, the human mind.

WHAT TO MEASURE?

Another aspect of theory—one that is overlooked in physics but is critical in psychology—is that the theory must tell us what measurements to make and how. For example, in classical thermodynamics, measurements are of volume, temperature, and pressure. In gravitation, we measure mass and position (as a function of time). In psychology, the measurement problem often is so critical and difficult that a measurement "solution" is considered to be an essential ingredient of the theory. For example, in a theory that proposes that success in school is some joint function of ability and motivation, the crux is the measurement of ability and motivation rather then the precise specification of the functional relation between them.

Theories in Psychology

Like physics, psychology, too, has powerful principles: for example, reinforcement. If a behavior produces a favorable outcome, the behavior is more likely than before to occur again. But as a student of cognitive psychology, I have observed over and over again that close examination of any particular phenomenon in cognitive psychology reveals that a simple principle accounts for only part of the, explanation. The attempt to improve the precision does not require better computation—it requires additional principles for every special case, and every case is special. As we gain more knowledge, the new knowledge does not yield an aesthetically pleasing theory. On the contrary, new knowledge usually causes psychologists to reject simpler, aesthetically more pleasing theories in favor of more complex, aesthetically less pleasing theories. For example, in aptitude testing it used to be thought that one factor, IQ (intelligence quotient) or *g* (general ability factor), was the main predictor of academic performance. Now we recognize that there are many independent and partially independent specific abilities that determine components of particular performances, not to mention other important, multidimensional factors such as ambition and prior training. In practical applications, the trend clearly is—as it should be—toward more complex, specialized tests rather than toward simpler, general tests. In fact, in studying almost any behavior, at first glance only a few variables seem to be important; later we discover that practically everything matters a little.

SUPERFICIAL DIFFERENCES BETWEEN PHYSICAL AND PSYCHOLOGICAL THEORIES

Despite the apparent difference in the structure of physical and of psychological theories, it has often been suggested that there is no essential difference between them. It is simply that physicists and engineers surround themselves only with those systems they understand; therefore, their theories appear to be more potent. If physicists had to deal with naturally occurring problems, such as smashing a particular martini cocktail on the kitchen floor and predicting precisely the shapes of the broken glass and of the puddles, and the location of the olive, physicists would be as unspecific as psychologists—claiming an understanding of the general mechanisms involved but unable to generate precise predictions for the individual case.

THE IMPORTANCE OF EVOLUTION AND LEARNING FOR PSYCHOLOGICAL THEORIES

While there undoubtedly is a kernel of truth to these "choice-of-problem" and "individual-case" hypotheses, I believe there is a profound difference between classical physics and the biological sciences that lies even deeper. I suggest that there is a twofold reason for the complexity of psychology and biology. Darwinian evolution has provided organisms with complex, interacting, innate mechanisms; and learning makes the current state of a complex organism (such as a human) an incredibly complex function of its history.

Consider first evolution: Organisms evolved to deal successfully with their environment. As a new environmental problem arose, a new biological adaptation evolved. But, while the first adaptation was solving part of the problem, it probably uncovered new, second-order unsolved problems. Thus adaptation follows upon adaptation. The result, ultimately, is a piling up of corrections, modifications, and adaptations upon each other. Moreover, an optimal solution balances many considerations related to survival of the individual and of the species. For example, there is only a limited amount of genetic code that can be devoted to solutions for any particular environmental problem, only a limited amount of space and energy within the organism for mechanisms subserving a single function, and only a limited amount of time to be allocated for but one purpose. The solution to one environmental problem cannot operate in isolation. We, and the other survivors of this evolutionary process, are the beneficiaries of an incredibly large number of interdependent, interacting mechanisms. And interacting mechanisms are inherently more difficult to analyze than the independent processes that classical physics concerns itself with.

Second, even if we were to understand all the mechanisms that evolved to facilitate the performance of some particular psychological function, through learning these mechanisms will have been modified and new ones acquired. Because of learning, any psychological theory with a pretense of precision must have a method for either (1) dealing with the entire history of the organism (so that its present state can be computed) or (2) precisely measuring its present state, where measurements of the present psychological state are quite a few orders of magnitude more complex than measurements of physical state, such as temperature or position.

THE COMPLEXITY OF A "SIMPLE" PSYCHOLOGICAL TASK

To illustrate what this complexity means for cognitive psychology, consider perhaps the simplest psychological experiment: the measurement of visual reaction time. The subject is told "Press the button as soon as you see the light. Get ready." The measured reaction time, from light flash (stimulus) to button press (response), depends on numerous factors, among which are:

- The intensity, size, shape, color, temporal wave form, and retinal location of the stimulus (to list just a few stimulus variables)
- The sequence of previously presented stimuli and knowledge of the set of alternative stimuli
- The particular foreperiod (from warning "get ready" to light onset)
- The distribution of prior foreperiods
- Previous reaction times
- Attention and expectancies
- Overall alertness, physiological state

- The explicit and implicit systems of payoffs and penalties for quick reactions and for false reactions
- Previous practice
- Phase of the stimulus relative to the brain's electrical alpha rhythm
- Individual differences (in age, body type, past training, motivation, interpretation of instructions, values of rewards, etc.)

A choice reaction time, in which the subject makes a different response depending on which of two or more stimuli occur, is enormously more complicated, as are tasks that involve the use of linguistic stimuli, such as words or sentences. Even in simplest reaction time experiments, many of the relevant factors interact, that is, the effect of one factor depends on the values of the others, so factors cannot be studied in isolation.

THE INDIRECT USE OF PRECISE THEORIES

In the past century, an enormous amount has been learned about many of the factors that control reaction times (and many other behavior tasks). There has been a gradual acceptance of the inherent complexity of even simple behavioral tasks, and the development of experimental paradigms and conceptual tools to deal with these complexities. It is fortunate that there is no urgent demand for a really precise theory of simple reaction times; we are satisfied to know the distribution of reaction times over a large number of trials, and we do not require prediction of the individual trial. Indeed, at this time, most of the interest in refinement of experimental theories is aesthetic. The thrust of research that attempts precision has been mainly to establish evidence for some particular proposed theory (relative to others) rather than to describe the observations themselves very precisely. To build a modern airplane or an electronic computer, numerous precise physical theories are required. We do not need the same precision to merely "further understanding," though precise prediction is taken by some as evidence that we do indeed "understand." For example, suppose it were to be demonstrated that the entire variation of reaction time with light intensity could be attributed entirely to the transducer properties of the retina. It would greatly further understanding but would not, in and of itself, improve the accuracy of a reaction-time theory one bit. That would require a theory of the retina.

THEORIES ARE INHERENTLY INCOMPLETE OR IMPRECISE

Not surprisingly, the quest for precision in psychological theories has thus far been futile. Every psychological theory that purports to deal with significant behavior is either wrong when examined in critical detail, or the theory is so imprecise that it makes no sense to examine it in detail.

For example, consider again the reinforcement theory. It seems so simple it could hardly be wrong, but the lack of precision is damaging. Reinforcement theory assumes we know what particular response has been reinforced. And that implicitly assumes there is only one internal, reinforceable motor program that can generate the response we observe. To choose an extreme counterexample, suppose the animal (or person) is testing a hypothesis and the hypothesis functions as a reinforceable response (e.g., "If I push button A once and button B twice I will get the reward"). If the theorist guesses wrongly about precisely which particular hypothesis is operable at the moment, his theory cannot predict individual responses, though it may be quite accurate in a statistical sense when averaged over many subjects and trials. A similar argument holds with respect to the discriminative stimulus that is provided. Perhaps the organism was paying attention to something else on that particular trial. These examples illustrate the problem if one does not have complete knowledge of the organism's internal state.

If the theorist's concern is just with the average behavior of a large number of animals or of one animal over a large number of trials, ignorance about precisely what aspect of the stimulus was conditioned to precisely which particular response may be unimportant—these fine details disappear in the averages. However, as soon as we examine behavior microscopically in individual cases, these fluctuations in conditioning will be revealed as significant departures from a theory that assumes a particular stimulus and a particular response to be conditioned and does not treat individual and trial-to-trial variations.

THE FUTILITY OF STATISTICAL TESTS

Because theories are inherently incomplete or imprecise, there is no ready solution for the evaluation of theories—not even that favorite of psychologists, comparing predictions of a theory with data from an experiment by means of a statistical test. The aim of statistical tests is to determine whether or not the predictions of the theory and the observed experimental data agree as well as would be expected if the theory were true. Unfortunately, as statisticians perpetually remind us, the "acceptance of a fit" does not mean the theory is either good or true—it only means the data are even "worse" than the theory. Better or more accurate data would always have rejected the theory because mechanisms (sources of variance) extraneous to the current theory would have been discriminated and revealed as significant departures. Statistics is most useful when it is used not to make absolute judgments of quality but to describe quantitatively the magnitude of the failure of a theory or to facilitate comparison of two theories that have different numbers of parameters.

Faced with the fact that every theory is only partially specified and only partially correct, we need other criteria to evaluate the goodness of a theory. Even when the number of parameters in two competitive theories is the same, accuracy of prediction as measured by a statistical test is only one of many criteria that might be used; other criteria are the generality of the theory, its consistency with other theories, and so forth.

A Criterion for Evaluating Theories

A general approach to the problem of "which theory" is to consider the best theory at each level of complexity. The sense of "best" and of "complexity" are matters of judgment, and will change in time as the scientific vocabulary and context change. We may improve the precision of any theory by making it more complex, but that does not necessarily make it preferable.

Consider the following example of theoretical difficulties that arise in answering the question, What is the function of the retina? This question also makes clear the essential confusion between yesterday's theory and today's fact (i.e., well-established theory). Hundreds of years ago, the avant-garde theorist might have proposed what today we regard as a fact, that the retina transduces light to provide neural input to the brain for further processing.

At a more complex level of theory, we may propose that the retina reduces the dynamic range of visual signals, that it emphasizes information near the fovea relative to the periphery, and that it reduces dimensionality of color information. Still further specification of retinal function might involve writing equations for different spatial-frequency channels of retinal output, and specifying interactions and connectivity within the retina.

Unfortunately for theory, but fortunately for vision, the retina is an exceedingly complex organ. It contains dozens of different kinds and sizes of neurons; their relative distribution varies from place to place; and they have evolved to satisfy many different purposes. The complete description of the retina is an enumeration of the properties of these neurons and their connections. Eventually we expect to have this ultimate description of the retina, that is, to be able to specify the properties of every class of neurons and to list their connections to all other neurons. Insofar as such a specification (presumably recorded in a computer larger and faster than today's models) would enable us to calculate the output of the retina for every conceivable input, it answers the question of What does the retina do? (It is an example of perfect reductionism—deriving the properties of the retina from more elementary knowledge.) But the level of complexity is too high. Studying the computer program is hardly better than studying the retina directly. This complete answer may not help us to discover elegant, useful, and partially correct answers at much simpler levels—the sort of answers readers of this chapter might desire.

The Necessity of Theory

In the United States, during the second quarter of the 20th century, in revulsion against the overblown theories of the turn-of the century mentalists, several movements arose that abjured formal theory. Most notable was *behaviorism* which originated with John B. Watson and persisted though B. F. Skinner. In parallel, there was a movement toward an eclectic style of research that has been dubbed "dustbowl empiricism." In both approaches, there was an implicit notion that one could simply enumerate and catalog the useful stimulus-response relationships that characterize human and animal behavior with requiring any formal theory.

The difficulty of a reliance on enumeration of stimulus-response relationships is illustrated by a simple, low-resolution computer screen. For example, a primitive screen consisting of 256 pixels, each with eight levels of intensity resolution, can produce 8^{256} different displays. This is a much larger number than the number of elementary particles in the universe. Without theories to reduce the input space, there would not be enough matter in the universe to record the stimulus-response relationships for even this static computer display. And this is an incredibly reduced situation compared to the dynamic real world in which experimental psychology operates. On the other hand, relatively simple theories of sensory processing can predict quite well which of these displays could be discriminated from a blank screen, from one another, and so on, although psychology is still a long way from a comprehensive theory of visual processing. The take-home message is that theories are necessary for progress. A second principle is that natural processes of Darwinian selection permit some theories to survive longer than others; the next section considers the bases of longevity among theories.

Looking Ahead

It is interesting to contrast the views of their discipline of physicists and psychologists. Classical physicists were able to simultaneously achieve theoretical simplicity, precision, and generality—at the price of derivational complexity. Some physicists still believe that new physical theories will encompass an increasingly greater range of phenomena and be even more precise than their predecessors. Possibly, the new theories may even have simple axioms, succeeding in maintaining simple axioms at the cost of ever more complex derivations and computations.

Today, in psychology, the most general theories also are simple, but they are of limited practical utility because either the theories are incomplete or imprecise or the quantities in the theory cannot be measured precisely. Increases in precision are wrought at a double cost: loss of generality and loss of simplicity. This I believe reflects the essential nature of psychology, and it represents an essential difference between psychology and classical physics.

What is a theoretical psychologist to do—theorize in greater detail about less and less until he achieves perfect mastery of almost nothing at all? A psychologist working on a practical problem works until he achieves cost-effective precision and the generality required to solve that problem. Better solutions are possible but not worth the effort. A psychologist seeking basic understanding also confronts the certainty that there is no final answer to any interesting question; there always will be new, unforeseen complications. But if a theorist can discover the best answer to a problem, at a given level of complexity, then this is a permanent contribution that cannot be improved.

My view is that science is like the Great Wall of China, and that a lasting scientific contribution is like the placement of one stone in that wall. The best theory at a given level of complexity is such a stone because, at that level of complexity, it cannot be improved. Of course it would be better if a scientist who proposed such a theory also provided us with a proof, or at least a compelling argument, that the proposed theory was indeed "best." But these arguments usually require a higher level of information, information that is not available when the theory is proposed. And notions of utility, of aesthetic desirability, and of complexity change. Methods that seemed complex yesterday are routine today.

Insofar as there is a central, unchanging core of meaning to the concepts of complexity and accuracy, a good, partially correct theory conceived today can look forward to immortality. This is the most favorable outlook of any scientific endeavor. Scientists, particularly psychologists, who seek the aesthetic experience of "understanding" should strive not only for more precise theories but also for simpler and more general ones at a given level of precision, and seek the best possible theory at a given level of complexity.

Acknowledgment

Supported by the Air Force Office of Scientific Research, Life Sciences Directorate, Visual Information Processing Program.