Herbert A Simon

Annu. Rev. Psychol. 1990. 41:1–19

INVARIANTS OF HUMAN BEHAVIOR[1]

*Herbert A. Simon**

Department of Psychology, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213

CONTENTS

The fundamental goal of science is to find invariants, such as conservation of mass and energy and the speed of light in physics. In much of science the invariants are neither as general nor as "invariant" as these classical laws. For instance, the isotopes of the elements have atomic weights that are *nearly*

*This is the eleventh in a series of prefatory chapters written by eminent senior psychologists.

integral multiples of the weight of hydrogen. *Some* inheritable traits of plants and animals observe the classical 1–2–1 ratio of Mendel. The number of familiar information chunks that can be held in short-term memory is *approximately* seven. It takes about 30 seconds to memorize an unpronouncable three-consonant nonsense syllable, but only *about* nine seconds to memorize a three-letter word.

Much biological knowledge is extremely specific, for biology rests on the diversity of millions of species of plants and animals, and most of its invariants apply only to single species. Because of inter-species molecular differences, even the important general laws (e.g. the laws of photosynthesis) vary in detail from one species to another (and sometimes among different individuals in a single species). Only at the most abstract and qualitative level can one find many general strict invariants in biology.

Moreover, some of the most important invariants in science are not quantitative at all, but are what Allen Newell and I (1976) have called "laws of qualitative structure." For example, the germ theory of disease, surely one of Pasteur's major contributions to biology, says only something like: "If you observe pathology, look for a microorganism—it might be causing the symptoms." Similarly, modern molecular genetics stems from the approximately correct generalization that inheritance of traits is governed by the arrangement of long helical sequences of the four DNA nucleotides.

Finally, in biological (including human) realms, systems change adaptively over time. Simple change is not the problem, for Newton showed how we can write invariant laws as differential equations that describe the eternal movements of the heavens. But with adaptative change, which is as much governed by a system's environment as by its internal constitution, it becomes more difficult to identify true invariants. As a result, evolutionary biology has a rather different flavor from physics, chemistry, or even molecular biology.

In establishing aspirations for psychology it is useful to keep all of these models of science in mind. Psychology does not much resemble classical mechanics, nor should it aim to do so. Its laws are, and will be, limited in range and generality and will be mainly qualitative. Its invariants are and will be of the kinds that are appropriate to adaptive systems. Its success must be measured not by how closely it resembles physics but by how well it describes and explains human behavior.

On another occasion (Simon 1979a) I have considered the form a science must take in order to explain the behavior of an adaptive, hence of an artificial, system. By "artificial" I mean a system that is what it is only because it has responded to the shaping forces of an environment to which it must adapt in order to survive. Adaptation may be quite unconscious and unintended, as in Darwinian evolution, or it may contain large components of conscious intention, as in much human learning and problem solving.

Taking the artificiality of human behavior as my central theme, I should like to consider its implications for psychology. Moreover, since *Homo sapiens* shares some important psychological invariants with certain nonbiological systems—the computers—I shall want to make frequent reference to them also. One could even say that my account will cover the topic of human and computer psychology.

PHYSICAL SYMBOL SYSTEMS

An important law of qualitative structure underlies the information processing paradigm in psychology. The Physical Symbol System Hypothesis (Newell & Simon 1976) *states that a system will be capable of intelligent behavior if and only if it is a physical symbol system.* A physical symbol system is a system capable of inputting, outputting, storing, and modifying symbol structures, and of carrying out some of these actions in response to the symbols themselves. "Symbols" are any kinds of patterns on which these operations can be performed, where some of the patterns denote actions (that is, serve as commands or instructions).

We are all familiar with the physical symbol systems called computers. Computers store symbols in the form of electro-magnetic patterns of some kind (quite different kinds in different computers); some of these patterns serve to instruct the computer what to do next (the stored program), while others contain numerical or nonnumerical information.

Information processing psychology claims that intelligence is achievable by physical symbol systems and only such systems. From that claim follow two empirically testable hypotheses: 1. that computers can be programmed to think, and 2. that the human brain is (at least) a physical symbol system. These hypotheses are tested by programming computers to perform the same tasks that we use to judge how well people are thinking, and then by showing that the processes used by the computer programs are the same as those used by people performing these tasks. In making the comparison we use thinking-aloud protocols, records of eye movements, reaction times, and many other kinds of data as evidence.

The physical symbol system hypothesis has been tested so extensively over the past 30 years that it can now be regarded as fully established, although over less than the whole gamut of activities that are called "thinking." For starters in reviewing the evidence, I would recommend Newell & Simon (1972), Simon (1979a, b, 1989a), and Anderson (1983). Readers can continue the survey with numerous references they will find in those sources. The exact boundaries of our present knowledge need not concern us: The territory in which the hypothesis has been confirmed is broad, encompassing many of the kinds of activities that define human professional and scholarly work.

Some skeptics continue to regard thinking as something to be explained at some unknown future date. Their imperviousness to the empirical evidence, which shows that the main processes of thinking have already been accounted for quite specifically, perhaps stems from the reluctance of human beings to view themselves as "mere machines." Even some biologists who have long since rejected vitalism where bodily functions are concerned remain vitalists when it comes to the mind.

It is still incorrectly thought by some that contemporary information processing psychology leaves unexplained such "holistic" phenomena—treasured by humanistic, existentialist, Marxist, and Gestalt psychologists—as intuition, insight, understanding, and creativity. A brief guide to the literature that deals with these phenomena in terms of the physical symbol system hypothesis will be found in Simon (1986). I will say no more about these matters in this paper, but will simply use the present rather than the future tense in describing the psychology of thinking.

What is the unfinished business? There is plenty of it, but I will mention just two important research targets that remain. First, each kind of task to which the human mind addresses itself may be regarded as defining a different "species" of thought. A certain number of these species have already been described in greater or lesser detail (e.g. solving puzzles like the Tower of Hanoi or Missionaries and Cannibals, playing chess like a master or a novice, making medical diagnoses, solving problems in elementary physics and mathematics, making certain kinds of scientific discoveries, learning language, using diagrams to solve problems, and understanding problem instructions). But since many other species of thought remain undescribed, a vast work of taxonomy and empirical exploration lies ahead. We should avoid thinking of this work as "mere" taxonomy, for it will unearth multitudes of interesting and important phenomena and extend our repertory of explanatory laws and invariants accordingly.

Second, in stark contrast to our complete understanding of the physical underpinnings of the operation of computers, we have only the vaguest knowledge today of how the symbol processing capabilities of the human brain are realized physiologically. Information processing psychology explains the software of thinking, but says only a little about its "hardware" (or "wetware"?). Information processing psychology and neural science are still miles apart, with only slight indications of how a bridge will be built between them—as it certainly will.

This situation is not without precedent. Organismic and cell biology made extensive progress long before biochemistry could explain their structures and processes. Nineteenth-century chemistry achieved substantial understanding of the reactions among molecules long before physics supplied any picture of atomic structure that could account for the observed chemical regularities.

Science suspended from skyhooks is not new, nor is it limited to particular disciplines. Contemporary physics provides a prime example of skyhook science in its continual movement downward to ever more fundamental and "elementary" particles, its greatest uncertainties lying always at the foundations.

The separation of information processing from neural science presents an important challenge to research, and a second great direction for exploration in psychology. What arrangement of neurons or neuronal circuits corresponds to a symbol? What is the physiological basis for the magical number seven? By what mechanism does the presence of a symbol in short-term memory initiate or guide a mental action? The agenda, containing these and many other items, provides work for both neural scientists and information processing psychologists, for the bridge will have to be built out from both banks before it can link in the middle.

ADAPTIVITY

Let me now put aside biological questions and return to human adaptivity and its implications for the laws of psychology. A look at computer adaptivity may cast some light on the human kind. A computer, it is said, can only do what it is programmed to do (which may be quite different from what the programmer *intended* it to do). Generally, it is not instructed to do specific things at all (e.g. to solve a particular linear programming problem), but to adapt its behavior to the requirements of a given task chosen from a whole population of tasks (e.g. to solve *any* linear programming problem lying within given size limits). Then its behavior in response to each task is adapted to the requirements of the task, and it behaves differently, in appropriate ways, with each task it is given. In short, it is an adaptive system.

The adaptiveness of computers leads to a question that is the converse of the one raised above. Can a computer be programmed to do *anything?* Of course not. Upper limits are set by the famous theorems of Gödel, which prove that every symbol processing system must be, in a certain fundamental sense, incomplete. It is a truth of mathematics and logic that any program (including those stored in human heads) must be unable to solve certain problems.

Computational Limits on Adaptivity

Far more important than the Gödel limits are the limits imposed by the speed and organization of a system's computations and sizes of its memories. It is easy to pose problems that are far too large, require far too much computation, to be solved by present or prospective computers. Playing a perfect game of chess by using the game-theoretic minimaxing algorithm is one such infeasible computation, for it calls for the examination of more chess positions than there are molecules in the universe. If the game of chess, limited to its 64 squares and six kind of pieces, is beyond exact computation, then we may expect the same of almost any real-world problem, including almost any problem of everyday life.

From this simple fact, we derive one of the most important laws of qualitative structure applying to physical symbol systems, computers and the human brain included: *Because of the limits on their computing speeds and power, intelligent systems must use approximate methods to handle most tasks. Their rationality is bounded.*

Reasoning Under the Optimality Principle

Historically, human adaptiveness (that is to say, rationality) has preoccupied economists even more than psychologists. Modern mainstream economic theory bravely assumes that people make their decisions in such a way as to maximize their utility (Simon 1979a). Accepting this assumption enables economics to predict a great deal of behavior (correctly or incorrectly) without ever making empirical studies of human actors.

If we wish to know what form gelatin will take when it solidifies, we do not study the gelatin; we study the shape of the mold in which we are going to pour it. In the same way, the economist who wishes to predict behavior studies the environment in which the behavior takes place, for the rational economic actor will behave in whatever way is appropriate to maximize utility in that environment. Hence (assuming the utility function to be given in advance), this maximizing behavior is purely a function of the environment, and quite independent of the actor.

The same strategy can be used to construct a psychology of thinking. If we wish to know how an intelligent person will behave in the face of a particular problem, we can investigate the requirements of the problem. Intelligence consists precisely in responding to these requirements. This strategy has, in fact, been pursued occasionally in psychology; the theories of perception of J. J. Gibson (1966) and John Marr (1982) exemplify it, as do some of the recent rational models of my colleague John R. Anderson (1989).

Why don't we, then, close up the laboratory, frequently a place of vexing labors and unwelcome surprises, and build a psychology of intelligence by rational analysis, as the economists have done? The answer, already suggested, lies in the law that I have called the Principle of Bounded Rationality (Simon 1989b). Since we can rarely solve our problems exactly, the optimizing strategy suggested by rational analysis is seldom available. We must find techniques for solving our problems approximately, and we arrive at different solutions depending on what approximations we hit upon. Hence, to describe, predict and explain the behavior of a system of bounded rationality, we must both construct a theory of the system's processes and describe the environments to which it is adapting.

Computational Feasibility: Bounded Rationality

Human rational behavior (and the rational behavior of all physical symbol systems) is shaped by a scissors whose two blades are the structure of task environments and the computational capabilities of the actor.

The study of cognitive psychology is the study of computational capabilities in the face of diverse tasks. It is not a trivial detail but a fundamental limit upon computation that human short-term memory can hold only a half dozen chunks, that an act of recognition takes nearly a second, and that the simplest human reactions are measured in tens and hundreds of milliseconds rather than microseconds or picoseconds. These basic physiological constants determine what kinds of computations are feasible in a given kind of task situation and how rapidly they can be carried out (Newell & Simon 1972; Simon 1979a). They are among the most important invariants that cognitive psychology has discovered, accounting for many phenomena observed in thinking and learning.

Noting that computational limits must be a central preoccupation of cognitive psychology does not exhaust the complications of the subject. We have also to take into account that thinking capacities are a function of skill and knowledge, stored neural structures in the brain. The expert can reach solutions that are unattainable by the novice, using computations and knowledge that are simply not available to the latter.

A lightning calculator carries out elementary symbolic processes no more rapidly than a person with ordinary skills in arithmetic; empirical studies reveal little or no difference in the speeds of their basic processes. Superiority in computation derives almost entirely from superior knowledge of arithmetic facts (e.g. knowledge of the multiplication table up to relatively large numbers, or of the table of squares, or of prime factors), combined with a superior repertory of computational strategies that save steps and conserve short-term and long-term memory capacity. In Chi et al (1988) the reader will find recent papers on expert performance in a variety of tasks, including memory and computational feats.

A major way to relax the limits of bounded rationality is to store in long-term memory knowledge and strategies that reduce the computational requirements of tasks. This would seem to add new plausibility to the argument for studying the requirements of the task rather than the properties of the actor. But the argument still fails. In tasks of any complexity, knowledge and strategies do not allow the expert to find an optimal solution, but only to find approximations that are far better than those available to "native" (or naive) intelligence. A knowledge of the calculus allows its possessor to solve many problems that could not be solved without it, but the domain of differential equations that cannot be exactly integrated in closed form vastly exceeds the domain of those that can be.

Rationality Without Optimization

The wide-ranging attempts since the Second World War to apply the optimizing tools of operations research (linear programming, integer and dynamic programming, queuing theory, and so on) to the decision problems of management have underlined the computational complexity of real-world problems, even relatively well-structured problems that are easily quantified. Using queuing theory, an optimum production schedule can be found for a factory that manufactures one or two products, using one or two different pieces of equipment. Adding even one more product or piece of equipment puts the problem beyond computational bounds for the fastest supercomputer. (An optimal class schedule for a university lies even further beyond the limits of practical computation.)

Yet factories (and universities) are scheduled every day. We are forced to conclude that methods other than optimization are used—methods that respect the limits of human and computer rationality. Perhaps the feasible methods are specific to each specific situation, in which case it is hard to see what cognitive psychology should say about them.

On the other hand, it is possible that some common properties, deriving from human bounded rationality, are shared by the approximating procedures people use in many kinds of complex situations. If so, it is the task of cognitive psychology to characterize these procedures, to show how they are acquired, and to account for their compatibility with the known computational limitations of the human brain.

MECHANISMS FOR RATIONALITY

Let me illustrate some of the mechanisms used by human bounded rationality to cope with real-life complexity. I will give just three examples from a much larger number that could be cited: processes used in problem solving by recognition, processes of heuristic search, and processes for inducing sequential patterns.

Recognition Processes

We now know that experts make extensive use of recognition processes, based on stored knowledge, to handle their everyday tasks. This recognition capability, based (by rough estimate) on 50,000 or more stored cues and associated knowledge, allows them to solve many problems "intuitively"—that is, in a few seconds, and without conscious analysis. Recognizing key cues allows experts to retrieve directly from memory information for dealing with the situations that the cues identify. Recognition processes have been shown to play a major role, perhaps *the* major role, in such diverse tasks as grandmaster chessplaying, medical diagnosis, and reading. Introductions to the evidence will be found in de Groot (1978), Simon (1979a) and Chi et al (1988).

Computer simulation models like EPAM (Feigenbaum & Simon 1984) provide explanatory mechanisms for recognition-based expertise, including a learning mechanism for acquiring the stored chunks on which it is based. Alternative models are being developed in the form of parallel, connectionist systems (EPAM is a basically serial system). The theoretical explanations and computer models assume processing speeds that are well within the known human physiological limits, and EPAM, as least, predicts a wide range of the phenomena that have been reported in the verbal learning literature (including the times reported by Ebbinghaus for the learning of nonsense syllables). We can regard intuition as a phenomenon that has been rather thoroughly explained: It is achieved through acts of recognition.

Heuristic Search

What about problems whose solutions are not provided by immediate recognition, but which require analysis? Here also, a number of the principal processes have been identified and simulated. Collectively, they are usually called heuristic (or selective) search. When a great space of possibilities is to be explored (and humans commonly balk at searching spaces when the possibilities number even in the hundreds), search becomes very selective. It is then guided by various rules of thumb, or heuristics, some of which are specific to particular tasks, but some of which are more general (Newell & Simon 1972).

If the task domain is highly structured, the task-specific heuristics may be very powerful, drawing upon the structural information to guide search directly to the goal. For instance, most of us apply a systematic algorithm when we must solve a linear equation in algebra. We don't try out different possible solutions, but employ systematic steps that take us directly to the correct value of the unknown.

If the task domain has little structure or the structure is unknown to us, we apply so-called "weak methods," which experience has shown to be useful in many domains, but which may still require us to search a good deal. One weak method is *satisficing*—using experience to construct an expectation of how good a solution we might reasonably achieve, and halting search as soon as a solution is reached that meets the expectation.

Picking the first satisfactory alternative solves the problem of making a choice whenever (*a*) an enormous, or even potentially infinite, number of alternatives are to be compared and (*b*) the problem has so little known structure that all alternatives would have to be examined in order to determine which is optimal. Satisficing also solves the common problem of making choices when alternatives are incommensurable, either because (*a*) they have numerous dimensions of value that cannot be compared, (*b*) they have uncertain outcomes that may be more or less favorable or unfavorable, or (*c*) they affect the values of more than one person. Then a satisficing choice can still be made as soon as an alternative is found that (*a*) is satisfactory along all dimensions of value, (*b*) has satisfactory outcomes for all resolutions of the uncertainty, or (*c*) is satisfactory for all parties concerned, respectively.

Another weak mothod is *means-ends analysis*—noting differences between the current situation and the desired goal situation, and retrieving from memory operators that, experience has taught us, remove differences of these kinds.

A small collection of heuristics, of which satisficing and means-ends analysis are important examples, have been observed as central features of behavior in a wide range of problem-solving behaviors where recognition capabilities or systematic algorithms were not available for reaching solutions without search. The prevalance of heuristic search is a basic law of qualitative structure for human problem solving.

Beginning with the General Problem Solver (GPS) in about 1958, a sizeable number of computer programs have been built to simulate heuristic search in various task domains. With their help, a rather detailed account has been given of human heuristic search, particularly in relatively well-structured domains that call upon only limited amounts of domain-specific knowledge (Newell & Simon 1972). With these programs as foundation, other investigations have built processes that can create problem representations for simple situations, using natural language inputs to supply information about the problem and task domain.

Serial Pattern Recognition

Ability to find patterns in sequences of numbers, letters, or geometric figures is an important component of human intelligence (Simon & Kotovsky 1963; Kotovsky & Simon 1973). The Thurstone Letter Series Completion Test and the Ravens tests are examples of tasks aimed at measuring this component. Laboratory studies of these tasks and computer simulations show that successful human performance depends on a few basic pattern-recognizing and pattern-organizing capabilities. In extrapolating sequential patterns, for example, subjects notice when identical symbols recur, or when there are subsequences of symbols that are successive items in a familiar list or "alphabet."

When subsequences repeat in a sequence, subjects can notice this fact and treat the repetitive subsequences as unitary components in a higher-level pattern. Thus, recursive or hierarchical patterns can be detected and extrapolated. These capabilities can be shown to be adequate for detecting pattern in complex pieces of music, and their sufficiency has been demonstrated by simulation programs capable of carrying out nontrivial musical analysis (Simon & Sumner 1968).

Procedural Rationality

Problem solving by recognition, by heuristic search, and by pattern recognition and extrapolation are examples of rational adaptation to complex task environments that take appropriate account of computational limitations—of bounded rationality. They are not optimizing techniques, but methods for arriving at satisfactory solutions with modest amounts of computation. They do not exhaust, but they typify, what we have been learning about human cognition, and they go a long way toward explaining how an organism with rather modest computational capabilities can adapt to a world that is very complex indeed.

The study of human behavior in the face of difficult tasks shows why we need a theory of processes (procedural rationality) as well as a theory of the requirements of the task (substantive rationality). A theory based only on task requirements could not tell us how behavior depends on knowledge of relevant cues or strategies. It could not explain why we satisfice instead of optimizing, or how we solve most everyday problems by recognizing cues that evoke their solutions. It could not give us a grasp of the range of strategies that may be available for handling a particular task, or the differences between expert and novice performance on the task. All these phenomena become understandable as we explore, by laboratory experiments and computer simulations, actual human behavior in a variety of task environments.

THINKING AND REASONING

If we go back, say, to Woodworth's *Experimental Psychology* (1938), we find that accounts of the human "higher mental functions" flow along two quite different channels, representing different intellectual ties to the adjacent disciplines. Woodworth devotes two chapters to complex cognitive tasks: one to problem solving, the other to reasoning (but titled "Thinking"). Woodworth's own comment (1938, p. 746) is "Two chapters will not be too many for the large topic of thinking, and we may make the division according to the historical sources of two streams of experimentation, which do indeed merge in the more recent work. One stream arose in the study of animal behavior and went on to human problem solving; the other started with human thinking of the more verbal sort." In particular, research on problem solving had its origins in the disputes about trial-and-error versus insightful learning. Research on reasoning derived from attention to theories of language and logic as models of thought processes.

The problem-solving model, or metaphor, has generally been preferred by Gestalt psychologists, with their common belief that insightful thinking is nonverbal in nature, and by researchers in artificial intelligence, who have from the beginning described thinking as heuristic search. The reasoning model, or metaphor, has generally been preferred by linguists with an interest in cognitive science and by philosophers who stray into this domain. They describe thought processes in terms of propositions and logical manipulations of propositions.

A study of mutual citations would show that communication between these two streams of inquiry has been poor. This reveals itself also in the different programming languages the two groups adopt when they simulate thinking processes. The programming languages associated with heuristic search are list-processing languages like LISP and production-system languages like OPS5. The programming languages associated with reasoning are logic languages (languages adapted to theorem proving in the predicate calculus) like PROLOG.

The division is further reinforced by disagreement about the respective roles in thinking of sentences (or the propositions they denote) and imagery of one or another kind. The reasoning metaphor views goals as described by sentences, derived from other sentences by processes similar to the processes of logic. The problem-solving metaphor views goals as achieved by sequences of moves through a problem space. (The very phrase "problem space" suggests the importance that is attached to a visual or spatial metaphor.)

When the reasoning metaphor is used, information is expressed mainly in declarative sentences. A small number of rules of inference (like the rule of syllogism in formal logic) are used to derive new sentences from old. The research tasks most commonly employed to study human reasoning are tasks of concept formation or tasks of judging the validity or invalidity of formal syllogisms, the presumption being that human thinking consists in drawing valid inferences from given premises or data.

When the problem-solving metaphor is used, information is expressed in schemas, which may resemble interrelated sets of sentences or may resemble diagrams or pictures of the problem situation. The problem situation is modified by applying "move operators," which are processes that change a situation into a new one. Nowadays, the move operators usually take the form of productions—condition-action pairs, $C \rightarrow A$. Whenever the information in short-term memory matches the conditions of a production, the actions of the production are executed. The execution of a sequence of productions accomplishes a search through the problem space, moving from one situation to another until a situation satisfying the goal requirements is reached.

All of these differences can be seen by comparing the corresponding two chapters, mentioned above, of Woodworth (1938), and chapters 8 and 10 of Anderson (1985).

Determining to what extent human thinking fits the problem-solving metaphor and to what extent it fits the reasoning metaphor stands high on the agenda of cognitive psychology today. Of course, the answer may be "both of the above," the processes of thought varying with the task domain and with learned or innate differences among the thinkers. It is perhaps of interest that Johnson-Laird, one of the leaders among those who emphasize the ties between cognition and linguistics, has recently begun to describe thinking in terms of "mental models," an approach that lies much closer to the heuristic search paradigm than to the reasoning paradigm (Johnson-Laird 1983). But I would hesitate to predict what this particular defection from the linguistic camp portends for the future.

COGNITIVE ARCHITECTURE

The whole congeries of mechanisms of human rationality must somehow be organized in the human brain to work together in a coordinated fashion. Today, a good deal of effort of theorists in psychology is devoted to specifying the architectures that achieve this coordination. In this context, "architecture" refers to description of the cognitive system at an abstract, usually symbolic, level, and has little to say about the underlying biology of neurons.

Early information-processing architectures of cognition (e.g. Broadbent 1958) emphasized memory "boxes" and their interconnections. Today, architectures specify organizations of processes as well as storage. Among the proposals that enter prominently into current discussion are Anderson's (1983) Act*, Newell's (1989) SOAR (both symbolic), and the connectionist system of McClelland & Rumelhart (1986).

Confining our discussion to symbolic architectures, while there are significant differences among them, none of them are incompatible with the recognition, heuristic search, and pattern-induction mechanisms described in the last section. *For many purposes, it is sufficient to think of the "whole cognitive man" as comprised of the following components:*

- *Memories:* A short-term memory of limited capacity (working memory), an associative long-term memory, an EPAM discrimination net to index the long-term memory, and smaller short-term memories associated with various sensory and motor modalities.
- *Sensory processors* to extract features from stimuli.
- *Interpreters of motor signals.*
- *EPAM: A discrimination net* for sensory features that learns new discriminations and "chunks" familiar stimuli patterns.
- *GPS: A problem solver* that employs heuristic search, and that can be used by the learning subsystem.

- *A Pattern induction system* that searches for regular patterns in stimuli.
- *Systems for encoding natural language* input and producing natural language output.
- *Systems for encoding to and from image-like representations.*
- *An adaptive production system,* capable of creating new processes on the basis of information gained through instruction, through examining worked-out examples, and by solving problems.

This may appear to be a lot of baggage, but all of the processing systems listed are implementable as production systems that can be stored in the associative long-term memory. Moreover, examples of all of these components have been simulated with computer programs, and their mutual compatibility tested to some degree. For example, the UNDERSTAND system (Hayes & Simon 1974) can encode natural language descriptions of puzzles into internal representations that are suitable problem spaces for GPS. The ISAAC system (Novak 1976) can encode natural language statements of physics problems into internal images, and use these images to produce algebraic equations, which it then solves.

From this we may conclude that, while many issues about architecture are fluid at the present time, the knowledge we gain about architectures is unlikely to invalidate, or require major revision of, the knowledge we have already gained about component mechanisms like EPAM or GPS, or about their roles in cognition.

LINKAGES TO OTHER PARTS OF PSYCHOLOGY

Contemporary information-processing psychology holds forth significant possibilities for a greater unification among domains of psychology that are now quite separate. It may be possible to forge stronger links of cognitive psychology with the study of child development, with research on individual differences, with psycholinguistics, and with social psychology.

There is already vigorous research in cognitive developmental psychology that makes use of many of the constructs discussed here. Computational linguistics and psycholinguistics also have proceeded along parallel—and sometimes even intersecting—lines. Some researchers on individual differences [the names of Hunt (1975) and Sternberg (1977) come immediately to mind] work within an information-processing framework. The linkages to social psychology are somewhat more tenuous, but are beginning to form, as we shall see. Without attempting a systematic review, I would like to offer comments on several of these topics—in particular, individual differences, social psychology, and psycholinguistics.

Individual Differences

Traditionally, the study of individual differences has employed psychometric methods of research. It has been motivated by interest in the nature/nurture controversy as well as by more practical concerns of predicting and explaining school and job performance. L. L. Thurstone's *Vectors of the Mind* (1935), which characterized each individual by a vector of weights for individual traits and predicted individual performance on specific tasks from the correlations between these weights and the importance of the corresponding traits for the tasks, provides a template (or perhaps a caricature) for this view of individual differences.

Thinking-aloud protocols, by providing rich information about the behavior of individual subjects, have focused attention on the large differences in these behaviors (Ericsson & Simon 1984). While protocol analysis has proved much more receptive to the study of individual differences than experimental designs that take averages over sets of subjects, it has been incorporated only incompletely into the current literature on individual differences. For example, Carroll's (1988) chapter on "Individual Differences in Cognitive Functioning," in the new *Stevens' Handbook of Experimental Psychology,* Volume 2, makes almost no reference to the new methods or the research on expert-novice differences produced by them.

Attending to the processes that subjects use in performing complex tasks has enabled us to characterize the differences between expert and novice performance in many task domains. In all of these domains, differences in knowledge (which must include learned skills as well as factual knowledge) prove to be a dominant source of differences in performance (Chi et al 1988).

Of course, this finding should not be taken to deny the existence of "innate" differences, but rather to account for their relative (quantitative) insignificance in explaining differences in skilled adult performance. No one would argue that any randomly selected person could be trained to play world-class tennis; but one could argue, on the basis of the evidence now available, that most normal human beings could become reasonably good players with sufficient training and practice, and that none could become excellent players without extensive training and practice.

Knowledge includes knowledge of strategies. A good deal of the research on expert-novice differences has been aimed at understanding the strategies that experts acquire and apply, how these strategies can be learned, and to what extent they are transferable from one task domain to another. Hence, in the contemporary paradigms, the study of individual differences is closely tied to the study of learning and transfer of training. These ties, in turn, introduce a strong taxonomic aspect into the study of individual differences, making clear that a great many task domains will have to be analyzed before we can generalize safely about human skills.

The new connections between skills and processes affect not only our understanding of complex performances, but also our interpretations of the simple processes that underlie them. We are aware today of the centrality of short-term memory limits to performance on many, if not most, cognitive tasks, and we have long used George Miller's seven chunks to characterize those limits.

But a chunk is not an innate measure of storage capacity. A chunk is any stimulus that has become familiar, hence recognizable, through experience. Hence, the capacity of short-term memory is itself determined by learning, and can grow to vast size as individual acts of recognition access larger and richer stores of information in long-term memory. Two EPAM systems possessing the same basic structure can differ greatly in measured STM capacity simply because one has a more elaborate differentiation net and associated store of schemas than the other.

Social Psychology

Just as individual differences find a natural place in information processing psychology, so do social phenomena. To the extent that cognitive peformance rests on skill and knowledge, it is a social, rather than a purely individual, phenomenon. Language skills and skills in social interaction can be approached within the same theoretical framework as knowledge and skills for dealing with the physical environment.

The recent work of Voss et al (1983) illustrates how cognitive and social psychology can mutually reinforce each other. He has studied how people who have different professional backgrounds and information approach the same problem-solving situation. When asked to write an essay on agricultural reform in the USSR, subjects who are experts in agronomy address themselves to entirely different variables and strategies than subjects who are experts on Russian political affairs. And both of these groups of subjects respond quite differently from novices. When we study expert behavior, we cannot help studying the structure of professional disciplines in our society.

Cognitive psychology still has an important task of studying the domain-independent components of cognitive performance. But since the performance depends heavily on socially structured and socially acquired knowledge, it must pay constant attention to the social environment of cognition. Many of the invariants we see in behavior are social invariants. And since they are social invariants, many are invariant only over a particular society or a particular era, or even over a particular social or professional group within a society. Social variables must be introduced to set the boundaries of our generalizations.

CONCLUSION

Let me summarize briefly this account of the invariants of human behavior as they are disclosed by contemporary cognitive psychology. The problem of identifying invariants is complicated by the fact that people are adaptive systems, whose behavior is highly flexible. The invariants must be sought in the mechanisms that allow them to solve problems and learn: the mechanisms of intelligence.

The Physical Symbol System Hypothesis, strongly supported by empirical evidence, asserts that a system will be capable of intelligent behavior if and only if it is a physical symbol system: if it can input, output, store, and manipulate symbols. The hypothesis, and consequently information-processing psychology, describes intelligence at a symbolic, "software," level, saying little about brain physiology. That fact need not impede our progress, nor has it, toward understanding at this symbolic level how a physical symbol system like the brain achieves intelligent behavior.

Because of the limits on their computing speeds and power, intelligent systems must use approximate methods. Optimality is beyond their capabilities; their rationality is bounded. To explain the behavior of a system of bounded rationality we must describe the system's processes and also the environments to which it is adapting. Human short-term memory can hold only a half dozen chunks, an act of recognition takes nearly a second, and the simplest human reactions are measured in tens and hundreds of milliseconds, rather than microseconds, nanoseconds, or picoseconds. These limits are among the most important invariants of intelligence.

A major strategy for achieving intelligent adaptation with bounded rationality is to store knowledge and search heuristics in a richly indexed long-term memory in order to reduce the computational requirements of problems. Experts use recognition processes, based on this stored, indexed knowledge, to handle their everyday tasks. When recognition does not suffice, because a great space of possibilities must be explored, they resort to highly selective search, guided by rich stores of heuristics.

When intelligence explores unfamiliar domains, it falls back on "weak methods," which are independent of domain knowledge. People satisfice—look for good-enough solutions—instead of hopelessly searching for the best. They use means-ends analysis to reduce progressively their distance from the desired goal. Paying attention to symmetries and orderly sequences, they seek patterns in their environments that they can exploit for prediction. Problem solving by recognition, by heuristic search, and by pattern recognition are adaptive techniques that are compatible with bounded rationality.

Several cognitive architectures have been proposed to account for the processes just described, but these architectures represent relatively modest variations on a basic pattern that is widely accepted today. This basic pattern involves some sensory processors that provide input into short-term and long-term memory, a recognition process that discriminates the features detected by the senses, a problem solver that employs heuristic search, a pattern induction system, systems for handling natural language, systems for handling image-like representations, and learning mechanisms that permit new processes and data structures to be constructed and stored in memory.

The picture I have drawn of cognitive psychology and the invariants of intelligence holds forth the promise of linking several parts of psychology that now mostly go their separate ways. Developmental psychology has already been strongly influenced by the cognitive revolution. The approach to intelligence and individual differences is beginning to be modified in an information-processing direction. But we are just beginning to see that, because of the strong dependence of intelligence on stored knowledge, cognitive and social psychology must be brought much closer together than they have been in the recent past. When we have made these new connections solid, the challenge will remain of bringing affect and emotion more centrally into the picture.

ACKNOWLEDGMENTS

This research was supported by the Personnel and Training Programs, Psychological Sciences Division, Office of Naval Research, under Contract No. N00014-86-K-0768; and by the Defense Advanced Research Projects Agency, Department of Defense, ARPA order 3597, monitored by the Air Force Avionics Laboratory under contract F33615-81-K-1539.

Literature Cited

Anderson, J. R. 1983. *The Architecture of Complexity*. Cambridge MA: Harvard Univ. Press

Anderson, J. R. 1985. *Cognitive Psychology and Its Implications*. New York: Freeman

Anderson, J. R. 1989. The place of cognitive architectures in a rational analysis. *22nd Annu. Symp. Cognit., Dept. Psychol.*, Carnegie-Mellon Univ.

Broadbent, D. E. 1958. *Perception and Communication*. New York: Pergamon

Carroll, J. B. 1988. Individual differences in cognitive functioning. In *Stevens' Handbook of Experimental Psychology*, ed. R. C. Atkinson, et al, Vol. 2:813–62. New York: Wiley. 2nd ed.

Chi, M. T. H., Glaser, R., Farr, M., eds. 1988. *The Nature of Expertise*. Hillsdale, NJ: Erlbaum

De Groot, A. 1978. *Chance and Choice in Chess*. The Hague: Mouton. 2nd ed.

Ericsson, K. A., Simon, H. A. 1984. *Protocol Analysis*. Cambridge, MA: MIT Press

Feigenbaum, E. A., Simon, H. A. 1984. EPAM-like models of recognition and learning. *Cogn. Sci.* 8:305–36

Gibson, J. J. 1966. *The Senses Considered as Perceptual Systems*. Boston: Houghton Mifflin

Hayes, J. R., Simon, H. A. 1974. Understanding written problem instructions. In *Knowledge and Cognition*, ed. L. W. Gregg, pp. 167–200. Hillsdale, NJ: Erlbaum

Hunt, E. B., Mansman, M. 1975. Cognitive theory applied to individual differences. In *Handbook of Learning and Cognitive Processes*, ed. W. K. Estes, Vol. 1. Hillsdale, NJ: Erlbaum

Johnson-Laird, P. N. 1983. *Mental Models*. Cambridge, MA: Harvard Univ.

Kotovsky, K., Simon, H. A. 1973. Empirical tests of a theory of human acquisition of concepts for sequential patterns. *Cogn. Psychol.* 4:399–424

Marr, D. 1982. *Vision*. San Francisco: Freeman

McClelland, J. L., Rumelhart, D. E. 1986. *Parallel Distributed Processing*, Vol. 2. Cambridge, MA: MIT Press

Newell, A. 1989. *Unified Theories of Cognition*. The William James Lectures. Cambridge, MA: Harvard Univ. Press. Forthcoming

Newell, A., Simon, H. A. 1972. *Human Problem Solving*. Englewood Cliffs, NJ: Prentice-Hall

Newell, A., Simon, H. A. 1976. Computer

science as empirical inquiry: symbols and search. *Commun. ACM* 19:111–26

Novak, G. S. Jr. 1976. *Computer understanding of physics problems stated in natural language*. Tech. Rep. No. NL-30. Austin, TX: Dept. Comput. Sci., Univ. Texas

Simon, H. A. 1979a. *The Sciences of the Artificial*. Cambridge, MA: MIT Press. 2nd ed.

Simon, H. A. 1979b. *Models of Thought*, Vol. 1. New Haven: Yale Univ. Press

Simon, H. A. 1986. The information processing explanation of Gestalt phenomena. *Comput. Hum. Behav*. 2:241–55

Simon, H. A. 1989a. Models of Thought, Vol. 2, New Haven: Yale Univ Press

Simon, H. A. 1989b. Cognitive architectures and rational analysis: comments. *21st Annu. Symp Cognit*., Dept. Psychol., Carnegie-Mellon Univ.

Simon, H. A., Kotovsky, K. 1963. Human acquisition of concepts for serial patterns. *Psychol. Rev*. 70:534–46

Simon, H. A., Sumner, R. K. 1968. Pattern in music. In *Formal Representation of Human Judgment*, ed. B. Kleinmuntz. New York: Wiley

Sternberg, R. J. 1977. *Intelligence, Information Processing, and Analogical Reasoning*. Hillsdale, NJ: Erlbaum

Thurstone, L. L. 1935. *Vectors of the Mind*. Chicago: Univ. Chicago Press

Voss, J. F., Greene, T. R., Post, T. A., Penner, B. C. 1983. Problem solving skill in the social sciences. In *The Psychology of Learning and Motivation*, ed. G. H. Bower, Vol. 17. New York: Academic

Woodworth, R. S. 1938. *Experimental Psychology*. New York: Holt