BEHAVIORAL AND BRAIN SCIENCES (1992) **15**, 425–492
Printed in the United States of America

Précis of *Unified theories of cognition*

Allen Newell
School of Computer Science, Carnegie Mellon University, Pittsburgh, PA 15213
Electronic mail: *newell@ cs.cmu.edu*

Abstract: The book presents the case that cognitive science should turn its attention to developing theories of human cognition that cover the full range of human perceptual, cognitive, and action phenomena. Cognitive science has now produced a massive number of high-quality regularities with many microtheories that reveal important mechanisms. The need for integration is pressing and will continue to increase. Equally important, cognitive science now has the theoretical concepts and tools to support serious attempts at unified theories. The argument is made entirely by presenting an exemplar unified theory of cognition both to show what a real unified theory would be like and to provide convincing evidence that such theories are feasible. The exemplar is SOAR, a cognitive architecture, which is realized as a software system. After a detailed discussion of the architecture and its properties, with its relation to the constraints on cognition in the real world and to existing ideas in cognitive science, SOAR is used as theory for a wide range of cognitive phenomena: immediate responses (stimulus-response compatibility and the Sternberg phenomena); discrete motor skills (transcription typing); memory and learning (episodic memory and the acquisition of skill through practice); problem solving (cryptarithmetic puzzles and syllogistic reasoning); language (sentence verification and taking instructions); and development (transitions in the balance beam task). The treatments vary in depth and adequacy, but they clearly reveal a single, highly specific, operational theory that works over the entire range of human cognition. SOAR is presented as an exemplar unified theory, not as the sole candidate. Cognitive science is not ready yet for a single theory – there must be multiple attempts. But cognitive science must begin to work toward such unified theories.

Keywords: artificial intelligence; chunking; cognition; cognitive science; computation; problem solving; production systems; SOAR; symbol systems

The book begins by urging on psychology unified theories of cognition:

> Psychology has arrived at the possibility of unified theories of cognition – theories that gain their power by positing a single system of mechanisms that operate together to produce the full range of human cognition.
>
> I do not say they are here, but they are within reach and we should strive to attain them.

My goal is to convince the reader that unified theories of cognition are really worth striving for – now, as we move into the nineties. This cannot be done just by talking about it. An exemplar candidate is put forth to illustrate concretely what a unified theory of cognition means and why it should be a goal for cognitive science. The candidate is a theory (and system) called SOAR (Laird et al. 1987).

The book is the written version of the *William James Lectures*, delivered at Harvard University in spring 1987. Its stance is personal, reflecting the author's thirty years of research in cognitive science, although this précis will be unable to convey much of this flavor.

Chapter 1: Introduction

The first chapter describes the enterprise. It grounds the concerns for how cognitive science should proceed by reflecting on a well-known earlier paper entitled "You can't play 20 questions with nature and win" (Newell 1973a), which even then fretted about the gap between the empirical and theoretical progress in cognitive psychology and called for more integrative theories. This book may be seen as a step toward answering that call.

The nature of theories. Chapter 1 discusses the notion of theory, to ground communication, building on some concrete examples: Fitts's Law, the power law of practice, and a theory of search in problem spaces. There is nothing special about a theory just because it deals with the human mind. It is important, however, that the theory make predictions, not the theorist. Theories are always approximate, often deliberately so, in order to deliver useful answers. Theories cumulate, being refined and reformulated, corrected and expanded. This view is Lakatosian, rather than Popperian: A science has investments in its theories and it is better to correct one than to discard it.

What are unified theories of cognition? Unified theories of cognition are single sets of mechanisms that cover all of cognition – problem solving, decision making, routine action, memory, learning, skill, perception, motor activity, language, motivation, emotion, imagining, dreaming, daydreaming, and so on. Cognition must be taken broadly to include perception and motor activity. No unified theory of cognition will deal with the full list above

 0140-525X/92 $5.00+.00

all at once. What can be asked is a significant advance in its coverage.

As the title indicates, the book is focused on the plural, on many unified theories of cognition. This is not eclecticism, but a recognition of the state of the art. Cognitive science does not have a unified theory yet. Many candidates will arise, given the current practice of theorizing in cognitive science, where every scientist of note believes himself a major theorist. This point is important, since the book works with a single exemplar (SOAR). An exemplar is not *the* unified theory, and not necessarily even a candidate.

Why strive for unified theories, beyond the apple-pie desire of all sciences to be unified? The biggest reason is that a single system (the mind) produces behavior. There are other reasons, however. Cognitive theory is radically underdetermined by data, hence as many constraints as possible are needed and unification makes this possible. A unified theory is a vehicle of cumulation simply as a theoretically motivated repository. A unified theory increases identifiability and allows theoretical constructs to be amortized over a wide base of phenomena.

The human mind can be viewed as the solution to a set of multiple constraints. Exhibiting flexible behavior, exhibiting adaptive (goal-oriented) behavior, operating in real time, operating in terms of the four-dimensional environment of perceptual detail and a body with many degrees of freedom, operating in a world requiring immense knowledge to characterize, using symbols and abstractions, using language, learning from experience about the environment, acquiring abilities through development, operating autonomously but also within a social community, being self-aware with a sense of self are all essential functionalities of the mind. A system must satisfy these constraints to be mind-like. Humans also have known constraints on construction: a neural system, grown by embryological processes, and arising through evolution. How necessary these constructive processes are, so that only systems built that way can be minds, is currently an open question, but the major point is that the embodied minds we see satisfy all these constraints and any theory that ignores any appreciable number of them loses important sources of direction.

Is psychology ready for unified theories? Cognitive science is well into its fourth decade; it is no longer a young child of a science. Indeed, behaviorism reached its own peak in fewer years. Cognitive science must take itself in hand and move forward. This exhortatory point is not made to suggest that cognitive science has made little progress. The strongest reason cognitive science should attempt unified theories now is that it has accumulated a vast and elegant body of regularities, highly robust and often parametric. This is especially the product of cognitive psychology and psycholinguistics, which have developed an amazing experimental engine for discovering, exploring, and confirming new regularities. Other sciences (e.g., biochemistry) have many more regularities, but they all fit within a theory that is integrated enough so that they never pose the challenge cognitive science now faces. If we do not begin integration now, we will find ourselves with an increasingly intractable task as the years go by while the engine of regularities works ever more industriously.

Though cognitive science does not yet have unified theories, there are harbingers: Many local theories make evident what cognitive mechanisms must be operating. But important attempts at unified theories have also been made. John Anderson's work on ACT* (Anderson 1983) must be taken to have pride of place among such attempts. [See also Anderson: "Is Human Cognition Adaptive" *BBS* 14(3) 1991.] Other examples are the Model Human Processor (Card et al. 1983), the CAPS theory (Just & Carpenter 1987), and a collection of efforts in perceptual decisions (Ratcliff 1985).

The task of the book. The book endeavors to make the case for serious work on unified theories of cognition. It adopts a specific strategy, presenting an exemplar theory. Any other way seems to involve just talk and exhortation, guaranteed to have little effect. There are lots of risks to such a course – it will seem presumptuous and people will insist on subjecting the exemplar to a Popperian criticism to falsify it. But, on the positive side, one can hope the reader will follow a frequent plea of Warren McCulloch's, issued in similar circumstances: "Don't bite my finger, look where I'm pointing" (McCulloch 1965).

Chapter 2: Foundations of cognitive science

Chapter 2 works through some basic cognitive-science concepts to provide a foundation for the remainder of the book. This is cast as a review, although some novel points arise.

Knowledge systems. A particularly important way of describing the human is as a knowledge system. The human is viewed as having a body of knowledge and a set of goals, so that it takes actions in the environment that its knowledge indicates will attain its goals. The term *knowledge* is used, as it is throughout computer science and AI, as *belief* (it can be wrong and often is), not as the philosopher's *justified true belief.* Knowledge systems are one level in the hierarchy of systems that make up an intelligent agent. For current computers, this is physical devices, continuous circuits, logic circuits, register-transfer systems, symbol (or programming) systems, and knowledge-level systems, all of which are simply alternative descriptions of the same physical system. Knowledge-level systems do not give a set of mechanisms that determine behavior, the hallmark of all other descriptive levels. Rather, behavior is determined by a principle of rationality that knowledge is used in the service of the agent's goals. This is analogous to other teleological principles, such as Fermat's principle of least time for optics. Lower-level descriptions (the symbol level) describe how a knowledge-level system is realized in mechanism. The knowledge level is useful to capture the notion of a goal-oriented system and abstract away from all details of processing and representation. However, humans can only be described approximately as knowledge-level systems, and the departure can be striking.

Representation. Knowledge must be represented in order to be used. The concept of representation is captured by the representation law. In an external world, entity (X) is transformed (T) into entity (Y). A representation of X-T-

Y occurs in a medium within some system when an encoding from X to an entity in the medium (x) and an encoding of T into an internal transformation in the medium (t) produces an internal entity (y), which can be decoded to the external world to correspond to Y. Actual representations are comprised of myriad instances of the representational law to cover all of the specific representational connections that actually occur.

Obtaining a representation for a given external situation seems to require discovering an internal medium with the appropriate natural transformations – this is the essence of analog representation. But as external situations become more diverse, complex, and abstract, discovering adequate analogs becomes increasingly difficult, and at last impossible. A radically different solution exists (the great move), however, where the internal medium becomes freely manipulable with combinatorially many states and all the representational work is done by being able to compose internal transformations to satisfy representational laws. Sufficiently composable schemes of transformations allow the formation of highly general representational systems that simultaneously satisfy many of the requisite representational laws.

Computation. Computational systems are exactly those that provide composability of transformations. The prime question about computational systems is what functions they can produce. The great move to composable transformations for representations occurs precisely because most machines do not admit much variety in their selectable transformations. This leads to the familiar, but incredible, results from computer science about *universal* computational systems that can attain the ultimate in flexibility. They can produce, by being instructed, all the functions that can be produced by any class of machines, however diverse. Thus, systems (universal computers) exist that provide the universal composability of transformations needed to produce systems that can universally represent whatever needs to be represented. This also shows that computation does not in itself represent. It provides the wherewithal for a system to represent if the appropriate representational laws are satisfied.

Symbols. The book takes the term *symbol* to refer to the parts of expressions that represent, for example, the "cat" in "The cat is on the mat." Symbols provide distal access to knowledge-bearing structures that are located physically elsewhere within the system. The requirement for distal access is a constraint on computing systems that arises from action always being physically local, coupled with only a finite amount of knowledge being encodable within a finite volume of space, coupled with the human mind's containing vast amounts of knowledge. Hence encoded knowledge must be spread out in space, whence it must be continually transported from where it is stored to where processing requires it (distribution does not gainsay this constraint). Symbols are the means that accomplish the required distal access.

Symbol systems are universal computational systems with the role of symbols made manifest. Symbol systems consist of (1) a *memory*, containing independently modifiable structures that contain symbols; (2) *symbols* (patterns in the structures), providing the distal access to other structures; (3) *operations*, taking symbol structures as input and producing symbol structures as output; and (4) *interpretation processes*, taking symbol structures as input and executing operations (the structures thereby representing these operations). There must be sufficient memory and symbols, complete composability of structures by the operators, and complete interpretability (any sequence of operations can be represented).

Within this cognitive-science framework, the great philosophical puzzle of intentionality (Brentano 1874) – how symbols can be *about* external things – has a straightforward solution. There are knowledge-level systems. The knowledge in them is *about* the external world. Symbol systems implement knowledge-level systems by using symbols, symbol structures, and so on. Therefore, these internal symbol structures are about (i.e., represent) the external world. They will only approximate such representation if the symbol system cannot realize the knowledge-level system adequately. Moreover, as the amount of knowledge and the diversity of goals increases, it is not possible, even theoretically, to realize faithfully the knowledge-level description of a system. How a given system comes to have its knowledge is a matter of the system's history, including the knowledge available to the processes that created the system. This appears to be a satisfactory resolution to the vexed question of intentionality.

Architectures. Unified theories of cognition will be formulated as architectures. The architecture of the mind is a major source of commonality of behavior, both within an individual and between individuals. The architecture is the fixed structure that realizes a symbol system. In the computer hierarchy this is the description at the register-transfer level; in biological systems it is the level of neural structure that is organized to provide symbols.

The important question about the architecture concerns what functions it provides. The architecture provides the boundary that separates structure from content, but all external tasks require both structure and content for their performance. So the division of function is what in the architecture *enables* the content to determine task performance. An obvious part of the answer is that the architecture provides the mechanisms for realizing a symbol system, but two additional types exist. One is the mechanisms to exploit implementation technology for power, memory, and reliability – such as caches and parallelism. The other is the mechanisms to obtain autonomy of operation – interrupts, dynamic-resource allocation, and protection. What is understood about the functions of the architecture comes entirely from engineered computers. Additional functions are surely involved in natural architectures for autonomous, intelligent creatures.

Architectures exhibit an immense variety. Universal computation might seem to require highly specialized systems for its realization. On the contrary, any specific symbol system can be realized in an indefinite variety of architectures, and any specific architecture can be implemented in an indefinite variety of technologies. Any technology that can implement one architecture can implement an indefinite variety of them. All these systems must perform the key functions of symbol systems, but these can be realized in an indefinite variety of ways. This potential for variety means that strong inferences are not possible from the structure of engineered digital computers to how architectures are realized in the brain.

Intelligence. The concept of intelligence is crucial for cognitive science. Unfortunately, its long, variegated history produced a multiplicity of notions that bear a family resemblance but serve different masters – often designed to block any unified concept of intelligence. Still, cognitive science (and any unified theory of cognition) must conceptualize the potential of a given task to cause difficulty to a person who attempts it and the potential of a given person for solving difficult tasks. *A system is intelligent to the degree that it approximates a knowledge-level system.* This is what emerges from the concept in the second chapter. The distinction between knowledge and intelligence is key. If a system does not have some knowledge, failure to use it cannot be a failure of intelligence, which can work only with the knowledge the system has. If a system uses *all* the knowledge it has and nothing improves its performance, then there is no role left for intelligence. Thus intelligence is the ability to use the knowledge the system has in the service of the system's goals. This notion answers many requirements of a concept of intelligence, but it does not lead directly to a quantitative measure of intelligence, because knowledge per se is not quantifiable.

Search and problem spaces. What processing is required to obtain intelligent behavior? How does a system bring its knowledge to bear to attain its goals? For difficult tasks the general answer is that the system will *search*. Search is not just another cognitive process, occurring alongside other processes (the view prior to the cognitive revolution), but the fundamental process for attaining tasks that require intelligence. There are two fundamental reasons for this. First, a difficult task is one in which the system does not always know how to behave. But to make progress means to generate some behavior, and when an error arises and is detected, to attempt to correct it – a de facto search step. When errors occur within errors, combinatorial search emerges. Second, search provides a method of last resort. If no other methods are available to a system, it can always posit a space within which goal attainment lies, and then search that space. No matter how little it knows, it can always posit a bigger space, so this method of "generate and test" can always be formulated.

An intelligent system is always operating in a *problem space*, the space of the system's own creation that attempts to restrict the arena of action to what is relevant. The agent is at some current state in this space with a set of available operators. The system searches within this space to reach a desired state that represents task attainment. This search is combinatorial in character, just as all the experience in AI attests. Solving problems in problem spaces is not just an arbitrary search. Knowledge can be brought to bear to guide the search. Given enough knowledge, no search at all will occur: The appropriate operator will be selected at each state and the desired state will be reached forthwith. For general intelligent systems (and humans), life is a sequence of highly diverse tasks and the system has available a correspondingly large body of knowledge. Thus, besides the problem search in the problem space there is also at every current state a knowledge search to discover what knowledge can be brought to bear to guide the search fruitfully. Knowledge search is a major activity in general intelligent systems.

Summary. The concepts in this chapter constitute the cumulated yield of thirty years of attempting to understand the nature of computation, representation, and symbols. As cast here, all the concepts are not equally familiar. The knowledge level is still not common in theoretical treatments, although it permeates the practice of cognitive and computer sciences. The separation of representation from computation is not sufficiently appreciated. The concept of intelligence may even seem strange. Despite these traces of novelty, this chapter should be like a refresher course to the practicing cognitive scientist.

Chapter 3: Human cognitive architecture

The concepts of Chapter 2 apply to humans and computers alike, but a unified theory of human cognition will be expressed in a theory of human cognitive architecture. This chapter attempts to discover some generally applicable constraints on the human architecture. Any proposed specific theory of the architecture would take such constraints as given and not as part of its specific architectural proposal. The chapter is necessarily speculative, since general arguments are notoriously fragile.

The human is a symbol system. This chapter argues that human minds are symbol systems. The strongest argument is from the flexibility and diversity of human response functions (i.e., responses as a function of the environment) – the immense variety of ways that humans generate new response functions, from writing books to reading them, to creating recipes for cooking food, to going to school, to rapping, to dancing. Other organisms are also adaptive, and in fascinating ways, but the diversity and range of human adaptations exceeds these by all bounds, indeed it is beyond enumeration. Focusing on diversity of response functions links up directly with the defining property of symbol systems as systems that admit the extreme of flexible response functions. Any system that is sufficiently flexible in its response functions must be a symbol system (i.e., capable of universal computation). Actually, the argument holds only asymptotically: No one has the foggiest notion what a class of systems might be like that showed human-scale flexibility but weren't universal. In addition, the simplicity of the functional requirements for symbol systems makes it most unlikely that such systems exist. Thus, the human mind is taken to be a symbol system, establishing a high-level constraint on the human cognitive architecture. It must support a symbol system.

Systems levels and the time scale of human action. Intelligent systems are built up in a hierarchy of system levels. Each system level consists of a more abstract way of describing the same physical system and its behavior, where the laws of behavior are a function only of the states as described at that level. In computers, engineers work hard to make the levels perfect, so that nothing from a lower level ever disturbs the given level. Nature is not so compulsive and levels are stronger or weaker depending on how complete is the sealing off from effects from lower levels. Higher system levels are spatially larger and run more slowly than do lower ones, because the higher levels are composed of multiple systems at the next lower level and their operation at a higher level comes from the operation of multiple interactive systems at the next lower level. Increase in size and slow-down in speed are geometric, although the factor between each level need not be constant. The concern in this chapter is with time, not space. In particular, the temporal factor for a minimal system level is about a factor of 10, that is, an order of magnitude. It could be somewhat less, but for convenience we will take $\times 10$ as the minimal factor.

Ranging up the time scale of action for human systems, a new systems level appears just about every factor of 10, that is just about as soon as possible. Starting at organelles, they operate at time scales of about 100 μsecs. Neurons are definitely a distinct system level from organelles, and they operate at about 1 msec, $\times 10$ slower. Neural circuits operate at about 10 msec, yet another $\times 10$ slower. These three systems can be taken to constitute the *biological band.* Continuing upward reaches what can be called the *cognitive band* – the fastest deliberate acts (whether external or internal) take on the order of 100 msec, genuine cognitive operations take 1 sec, and above that, at the order of 10 sec is a region with no standard name, but consisting of the small sequences of action that humans compose to accomplish smallish tasks. Above the cognitive band lies the *rational band* where humans carry out long sequences of actions directed toward their goals. In time scale this ranges from minutes to hours. No fixed characteristic systems level occurs here, because the organization of human activity now depends on the task being attempted and not on the inner mechanisms. Above the rational band is the social band, dominated by the distributed activities of multiple individuals. As the scale proceeds upward, the boundaries become less distinct, due to the flexibility of human cognition and the dominance of task organization. The time scale of human action reflects both a theoretical view about minimal systems levels and an empirical fact that human activities, when ranged along such a scale, provide distinguishable system levels about every minimal factor.

The real-time constraint on cognition. That neurons are ~1 msec devices and elementary neural circuits are ~10 msec devices implies that human cognition is built up from ~10 msec components. But elementary cognitive behavior patently occurs by 1 sec. Fast arcs from stimulus to response occur five times faster (~200 msec), but their simplicity and degree of preparation make them suspect as cognition. Yet creative discourse happens in about one second. These two limits create the real-time constraint on cognition: Only about 100 operation times are available to attain cognitive behavior out of neural-circuit technology. This constraint is extremely binding. It provides almost no time at all for the cognitive system to operate. The constraint may also be expressed as follows: Elementary but genuine cognition must be produced in just two system levels. Neural circuits (at ~10 msec) can be assembled into some sorts of macrocircuits (one factor of 10) and these macrocircuits must then be assembled to produce cognitive behavior (the second factor of 10). This constraint is familiar (Feldman & Ballard 1982; Fodor 1983) and has been deployed mostly to deny the relevance of the algorithms developed in AI for vision and natural language processing because they take too long. But the constraint is much more binding than that and can be used to make a number of inferences about the human cognitive architecture.

The cognitive band. The human cognitive architecture must now be shaped to satisfy the real-time constraint. A particular style of argument is used to infer the system levels of the cognitive band. Functions are allocated to the lowest (fastest) possible system level by arguments that they could not be accomplished any faster, given other allocations (and starting at the bottom of ~10 msec). Whether they could be slower is undetermined. But as they stack up, the upper limit of cognitive behavior at ~1 sec is reached, clamping the system from the top, thereby determining absolutely the location of cognitive functions at specific system levels.

The upshot is that the distal accessing associated with symbols must occur at the level of neural circuits, about 10 msec. Above this, hence at ~100 msec, comes the level of elementary deliberations, the fastest level at which (coded) knowledge can be assembled and be brought to bear on a choice between operations. This level marks the distinction in cognition between automatic and controlled processing. What happens within an act of deliberation is automatic, and the level itself permits control over action.

A level up from elementary deliberations brings simple operations, composed of a sequence of deliberations with their associated microactions, hence taking of the order of 1 sec. This brings the system up against the real-time constraint. It must be able to generate genuine, if elementary, cognitive activity in the external world. Simple operations provide this: enough composition to permit a sequence of realizations of a situation and mental reactions to that realization, to produce a response adaptive to the situation. Thus, the real-time constraint is met.

With time, cognition can be indefinitely composed, though a processing organization is required to control it. Above the level of simple operations is the first level of composed operations, at ~10 sec, characterized by its operations being decomposed into sequences of simple operations. An important bridge has been crossed with this level, namely, simple operations are a fixed repertoire of actions and now the operations themselves can be composed.

The intendedly rational band. Composition is recursive and more complex operations can exist whose processing requires many sublevels of suboperations. What prevents the cognitive band from simply climbing into the sky? Cognition begins to succeed; as the seconds grow into minutes and hours, enough time exists for cognition to extract whatever knowledge exists and bring it to bear. The system can be described increasingly in knowledge-level terms and the internal cognitive mechanism need not be specified. This becomes the band of rational – goal and knowledge driven – behavior. It is better labeled *intendedly* rational behavior, since the shift toward the knowledge level takes hold only gradually and can never be complete.

Summary. This chapter has produced some general constraints on the nature of the human cognitive architecture. These must hold for all proposed architectures, becoming something an architecture satisfies rather than an architectural hypothesis per se. The gain to theorizing is substantial.

The different bands – biological, cognitive, and (intendedly) rational – correspond to different realms of law. The biological band is solidly the realm of natural law. The cognitive band, on the other hand, is the realm of representational law and computational mechanisms. The computational mechanisms are described by natural law, just as are biological mechanisms. But simultaneously, the computations are arranged to satisfy representational laws, so that the realm becomes about the external world. The rational band is the realm of reason. Causal mechanisms have disappeared and what determines behavior is goals and knowledge (within the physical constraints of the environment).

Chapter 4: Symbolic processing for intelligence

The chapter deals with the symbolic processing required for intelligence and introduces the SOAR architecture. The shift from general considerations to full details of an architecture and its performance reflects the cardinal principle that the only way a cognitive theory predicts intelligence is if the system designed according to that theory exhibits intelligent behavior. Intelligence is a functional capability.

The central architecture for performance. In SOAR all tasks, both difficult and routine, are cast as *problem spaces*. All long-term memory is realized as a production system in which the productions form a recognition memory, the conditions providing the access path, and the actions providing the memory contents. Unlike standard production systems, there is no conflict resolution, all satisfied productions put their contents into working memory. Thus SOAR is entirely problem-space structured, and the recognition of which productions fire constitutes the knowledge search.

Control over behavior in the problem space is exercised by the *decision cycle*. First, information flows freely from the long-term memory into working memory. New elements may trigger other productions to fire, adding more elements, until all the knowledge immediately available in long-term memory is retrieved. Included in this knowledge are preferences about which decisions are acceptable or better than others. Second, a decision procedure sorts through the preferences to determine the next step to take in the problem space: what operator to select, whether the task is accomplished, whether the problem space is to be abandoned, and so on. The step is taken, which initiates the next decision cycle.

The decision cycle suffices if the knowledge retrieved is sufficient to indicate what step to take next. But if not, an *impasse* occurs – the decision procedure cannot determine how to proceed given the preferences available to it. Impasses occur frequently, whenever knowledge cannot be found just by immediate pattern recognition. The architecture then sets up a subgoal to acquire the missing knowledge. Thus the architecture creates its own goals whenever it does not have what is needed to proceed. Within the subgoal, deciding what problem space to use and what operators to select occurs simply by continuing with decision cycles in the new context. Impasses can arise while working on a subgoal, giving rise to a hierarchy of goals and subgoals, in the manner familiar in complex intelligent systems.

Chunking. The organization of productions, problem spaces, decisions, and impasses produces performance, but it does not acquire new permanent knowledge. *Chunking* provides this function. This is a continuous, automatic, experience-based learning mechanism. It operates when impasses are resolved, preserving the knowledge that subgoals generated by creating productions that embody this knowledge. On later occasions this knowledge can be retrieved immediately, rather than again reaching an impasse and requiring problem solving. Chunking is a process that converts goal-based problem solving into long-term memory. Chunks are active processes, not declarative data structures to be interpreted. Chunking does not just reproduce past problem solving; it transfers to other analogous situations, and the transfer can be substantial. Chunking applies to all impasses, so learning can be of any kind whatever: what operators to select, how to implement an operator, how to create an operator, what test to use, what problem space to use, and so on. Chunking learns only what SOAR experiences (since it depends on the occurrence of impasses). Hence, what is learned depends not just on chunking but on SOAR's problem solving.

The total cognitive system. SOAR's cognitive system consists of the performance apparatus plus chunking. The *total cognitive system* adds to this mechanisms for perception and motor behavior. The working memory operates as a common bus and temporary store for perception, central cognition, and motor behavior. Perceptual systems generate elements in working memory, which are matched by the productions in long-term memory. Central cognition generates elements in working memory, which are interpreted as commands by the motor system. Perceptual processing occurs in two stages: the (sensory) mechanisms that deliver elements to working memory and the analysis and elaboration of these perceptual elements by encoding productions. Likewise on the motor side, decoding productions in long-term memory elaborate motor commands and produce whatever form is needed by the motor systems, followed by motor system proper that makes movements. The sensory and motor modules are cognitively impenetrable, but the encoding and decoding processes interact with other knowledge in working memory.

SOAR as an intelligent system. Intelligence is only as intelligence does. The chapter describes the range of different tasks, types of learning, and modes of external interaction that SOAR has exhibited. Two large SOAR systems are described in some detail. One, R1-SOAR (Rosenbloom et al. 1985), does the task of R1, a classical expert system (McDermott 1982), which configures VAX systems. R1-SOAR does the same task. It shows that a single system can mix general (knowledge-lean) problem solving and specialized (knowledge-intensive) operation as a function of what knowledge the system has available. R1-SOAR also shows that experiential learning can acquire the knowledge to move the system from knowledge-lean to knowledge-intensive operation. The second system, Designer-SOAR (Steier 1989), designs algorithms, a difficult intellectual task that contrasts with the expertise-based task of R1. Designer-SOAR starts with a specification of an algorithm and attempts to discover an algorithm in terms of general actions such as generate, test, store, and retrieve, using symbolic execution and execution on test cases. Designer-SOAR learns within the span of doing a single task (within-task transfer), and also between tasks of the same basic domain (across-task transfer), but it shows little transfer between tasks of different domains.

Mapping SOAR onto human cognition. SOAR is an architecture capable of intelligent action. Next, one must show that it is an architecture of human cognition. Given the results about the cognitive band, deriving from the real-time constraint, there is only one way to interpret SOAR as the human cognitive architecture. Moreover, since these results have established absolute, though approximate, time scales for cognitive operations, this interpretation leads to an order-of-magnitude absolute temporal identification of the operations in SOAR as a theory of cognition. SOAR productions correspond to the basic symbol access and retrieval of human long-term memory, hence they take ~10 msec. The SOAR decision cycle corresponds to the level of elementary deliberation and hence takes ~100 msec. The problem-space organization corresponds to higher organization of human cognition in terms of operations. Operators that do not reach an impasse correspond to simple operations, hence they take ~1 sec. SOAR problem spaces within which only simple (nonimpassing) operators occur correspond to the first level of composed operations. This is the first level at which goal attainments occur and the first at which learning (impasse resolution) occurs. Problem spaces of any degree of complexity of their operators are possible and this provides the hierarchy of operations that stretches up into the intendedly rational level.

Summary. This chapter has a strong AI flavor, because the emphasis is on how a system can function intelligently, which implies constructing operational systems. A prime prediction of a theory of cognition is that humans are intelligent and the only way to make that prediction is to demonstrate it operationally. The prediction is limited, however, by the degree to which the SOAR system itself is capable of intelligence. SOAR is state of the art AI, but it cannot deliver more than that.

Chapter 5: Immediate behavior

The book now turns to specific regions of human behavior to explore what a unified theory must provide. The first of these is behavior that occurs in a second or two in response to some evoking situation: immediate-response behavior. This includes most of the familiar chronometric experimentation that has played such a large role in creating modern cognitive psychology.

The scientific role of immediate-response data. When you're down close to the architecture, you can see it, when you're far away you can't. The appropriate scale is temporal and behavior that takes 200 msec to about 3 sec sits close to the architecture. Thus, immediate-response performance is not just another area of behavior to illustrate a unified theory, it is the area that can give direct experimental evidence of the mechanisms of the architecture. Furthermore, cognitive psychology has learned how to generate large numbers of regularities at this level, many of which are quantitative, parametric, and robust. Literally thousands of regularities have been discovered (the book estimates ~3000). Tim Salthouse (1986) provides an illustration by his listing of 29 regularities for just the tiny area of transcription typing (this and several other such listings are given and discussed throughout the remainder of the chapter and book). All of these regularities are constraints against the nature of the architecture. They provide the diverse data against which to identify the architecture. Thus, it is appropriate to start the consideration of SOAR as a unified cognitive theory by looking at immediate behavior.

Methodological preliminaries. SOAR is a theory just like any other. It must explain and predict the regularities and relate them to each other; however, it need not necessarily produce entirely novel predictions: An important role is to incorporate what we now understand about the mechanisms of cognition, as captured in the microtheories of specific experimental paradigms. A scientist should be able to think in terms of the architecture and then explanations should flow naturally. SOAR should not be treated as a programming language. It is surely programmable – its behavior is determined by the content of its memory and stocking its memory with knowledge is required to get SOAR to behave. But SOAR is this way because humans are this way, hence programmability is central to the theory. That SOAR is not only programmable but universal in its computational capabilities does not mean it can explain anything. Important additional constraints block this popular but oversimple characterization. First, SOAR must exhibit the correct time patterns of behavior and do so against a fixed set of temporal primitives (the absolute times associated with the levels of the cognitive band). Second, it must exhibit the correct error patterns. Third, the knowledge in its memory – its program and data – must be learned. It cannot simply be placed there arbitrarily by the theorist, although as a matter of necessity it must be mostly posited by the theorist because the learning history is too obscure. Finally, SOAR as a theory is underspecified. The architecture continues to evolve, and aspects of the current architecture (SOAR 4 in the book, now SOAR 5) are known to be wrong. In this respect, a unified theory is more like a Lakatosian research programme than a Popperian theory.

Functional analysis of immediate responses. The tasks of immediate responding comprise a family with many common characteristics, especially within the experimental paradigms used by cognitive psychologists. These common properties are extremely constraining, and make it possible to specialize SOAR to a theory that applies just to this class of tasks. Immediate responses occur in the base-level problem space, where the elements generated by perception arise and where commands are given to the motor system. This base-level space is also the one that

Johnson-Laird, P. (1983) *Mental models*. Harvard University Press. [aAN, NEW]
Johnson-Laird, P. N. & Bara, B. G. (1984) Syllogistic inference. *Cognition* 16:1–61. [NEW]
Johnson-Laird, P. N. & Byrne, R. M. J. (1991) *Deduction*. Erlbaum. [rAN]
Just, M. A. & Carpenter, P. A. (1987) *The psychology of reading and language comprehension*. Allyn & Bacon. [aAN]
Kaplan, S. (1973) Cognitive maps in perception and thought. In: *Image and environment*, ed. R. M. Downs & D. Stea. Aldine. [EC]
(1991) Beyond rationality: Clarity-based decision making. In: *Environment, cognition and action: An integrative multidisciplinary approach*, ed. T. Garling & G. Evans. Oxford University Press. [EC]
Kaplan, S. & Kaplan, R. (1982) *Cognition and environment: Functioning in an uncertain world*. Praeger. (republished by Ulrich's, 1989). [EC]
Kaplan, S., Sonntag, M. & Chown, E. (1991) Tracing recurrent activity in cognitive elements (TRACE): A model of temporal dynamics in a cell assembly. *Connection Science* 3:179–206. [EC]
Kaplan, S., Weaver, M. & French, R. (1990) Active symbols and internal models: Towards a cognitive connectionism. *Artificial Intelligence & Society* 4:51–71. [EC]
Katona, G. (1940) *Organizing and memorizing*. Columbia University Press. [CP]
Keele, S. W. (1986) Motor control. In: *Handbook of perception and human performance*, vol. II, ed. K. R. Boff, L. Kaufman & J. P. Thomas. Wiley. [rAN, EH]
Kieras, D. E. & Bovair, S. (1986) A production system analysis of transfer of training. *Journal of Memory and Language* 25:506–24. [RWR]
Kintsch, N. (1974) *The representation of meaning in memory*. Erlbaum. [SL, CP]
Klapp, S. T., Marshburn, E. A. & Lester, P. T. (1983) Short-term memory does not involve the "working memory" of information processing: The demise of a common assumption. *Journal of Experimental Psychology: General* 112:240–64. [RAC]
Kleinsmith, L. J. & Kaplan, S. (1964) Interaction of arousal and recall interval in nonsense syllable paired-associate learning. *Journal of Experimental Psychology* 67:124–26. [EC]
Koch, S., ed. (1959) Epilogue. In: *Psychology: A study of a science*, vol. 3. McGraw-Hill. [NHF]
Kolers, P. A. & Smythe, W. E. (1984) Symbol manipulation: Alternatives to the computational view of mind. *Journal of Verbal Learning and Verbal Behavior* 23:328–30. [PCD]
Kornblum, S., Hasbroucq, T. & Osman, A. (1990) Dimensional overlap, a cognitive basis for stimulus-response compatibility: A model and taxonomy. *Psychological Review* 97:253–70. [rAN, RWP]
LaBerge, D. (1980) Unitization and automaticity in perception. In: *Nebraska symposium on motivation*, ed. J. H. Flowers. University of Nebraska Press. [MVL]
LaBerge, D. & Brown, V. (1989) Theory of attentional operations in shape identification. *Psychological Review* 96:101–24. [rAN]
Lachman, J. L. & Lachman, R. (1979) Theories of memory organization and human evolution. In: *Memory organization and structure*, ed. C. R. Puff. Academic Press. [EC]
Laird, J. E., Congdon, C. B., Altmann, E. & Swerdlow, K. (1990) *Soar user's manual: Version 5.2*. University of Michigan; Department of Electrical Engineering and Computer Science. [JAM]
Laird, J. E., Newell, A. & Rosenbloom, P. S. (1987) Soar: An architecture for general intelligence. *Artificial Intelligence* 33:1–64. [aAN, JAM]
Laird, J. E. & Rosenbloom, P. S. (1990) Integrating execution, planning, and learning in Soar for external environments. *Proceedings of the National Conference on Artificial Intelligence*, July. [rAN]
Laird, J. E., Rosenbloom, P. S. & Newell, A. (1986) *Universal subgoaling and chunking*. Kluwer. [RWP]
Lakatos, I. M. (1978) *The methodology of scientific research programmes*. Cambridge University Press. [JMC]
Lehman, J. F., Lewis, R. L. & Newell, A. (1991a) Integrating knowledge sources in language comprehension. *Proceedings of the Cognitive Science Thirteenth Annual Conference*, August. Cognitive Science Society. [rAN]
(1991b) *Natural language comprehension in Soar: Spring 1991* (Technical Report CMU-CS-91-117). School of Computer Science, Carnegie-Mellon University. [rAN]
Lehman, J. F., Newell, A., Polk, T. A. & Lewis, R. (1992) The role of language in cognition: A computational inquiry. In: *Conceptions of the human mind*, ed. G. Harman. Erlbaum (in press). [rAN]
Levenick, J. R. (1991) NAPS: A connectionist implementation of cognitive maps. *Connection Science* 3:107–26. [EC]
Lewandowsky, S., Dunn, J. C. & Kirsner, K., eds. (1989) *Implicit memory: Theoretical issues*. Erlbaum. [SL]
Lewis, R. (1992) Recent developments in the NL-Soar garden path theory. Technical report, School of Computer Science, Carnegie Mellon University (forthcoming). [rAN]
Lewis, R., Huffman, S., John, B., Laird, J., Lehman, J. F., Newell, A., Rosenbloom, P., Simon, T. & Tessler, S. (1990) Soar as a unified theory of cognition: Spring 1990. *Proceedings of the Cognitive Science Twelfth Annual Conference* (July). Cognitive Science Society. [rAN]
Lewis, R. L., Newell, A. & Polk, T. A. (1989) Toward a Soar theory of taking instructions for immediate reasoning tasks. *Proceedings of the Cognitive Science Eleventh Annual Conference* (August). Cognitive Science Society. [rAN]
Lindsay, R. K. (1991) Symbol-processing theories and the Soar architecture. (Review of *Unified theories of cognition*). *Psychological Science* 2:294–302. [rAN, RWR]
Logan, G. D. (1985) Executive control of thought and action. *Acta Psychologica* 60:193–210. [RAC]
Mandler, G. & Mandler, J., eds. (1964) *Thinking: From association to Gestalt*. Wiley. [CP]
Mark, L. S. (1987) Eyeheight-scaled information about affordances: A study of sitting and stair climbing. *Journal of Experimental Psychology: Human Perception and Performance* 13:361–70. [KJV]
Markovitch, S. & Scott, P. (1988) The role of forgetting in learning. *Proceedings of the Fifth International Conference on Machine Learning*, Ann Arbor, Michigan. [EC]
Marr, D. (1982) *Vision*. W. H. Freeman. [rAN]
Maslov, S. Y. (1987) *Theory of deductive systems and its applications*. MIT Press. [LG]
Mayr, E. (1982) *The growth of biological thought: Diversity, evolution and inheritance*. Harvard University Press. [rAN]
McClelland, J. L. (1988) Parallel distributed processing: Implications for cognition and development. Technical report, Department of Psychology, Carnegie Mellon University. [rAN]
McClelland, J. L. & Rumelhart, D. E. (1986) *Parallel distributed processing*, vol. 2. MIT Press. [CP]
McCulloch, W. S. (1965) *Embodiments of mind*. MIT Press. [aAN]
McDermott, J. (1982) R1: A rule-based configurer of computer systems. *Artificial Intelligence* 19:39–88. [aAN]
Meyer, D. E., Abrams, R. A., Kornblum, S., Wright, C. E. & Smith, J. E. K. (1988) Optimality in human motor performance: Ideal control of rapid aimed movements. *Psychological Review* 95:340–70. [RWP]
Miller, C. S. & Laird, J. E. (1992) A simple symbolic algorithm for incremental concept acquisition. Artificial Intelligence Laboratory, University of Michigan, January, (unpublished). [rAN]
Miller, C. S. & Laird, J. E. (1991) A constraint-motivated model of concept formation. *Proceedings of the Cognitive Science Society*, August. [rAN]
Miller, G. A. (1986) Dismembering cognition. In: *G. Stanley Hall and the Johns Hopkins Tradition*, ed. S. H. Hulse & B. F. Green, Jr. Johns Hopkins University Press. [SFC]
Mitchell, T. M., Allen, J., Chalasani, P., Cheng, J., Etzioni, O., Rinquette, M. & Schlimmer, J. C. (1991) Theo: A framework for self-improving systems. In: *Architectures for intelligence*, ed. K. VanLehn. Erlbaum. [rAN]
Montague, R. (1974) *Formal philosophy*, ed. R. Thomason. Yale University Press. [PS]
Moore, E. F. (1956) Gedanken-experiments on sequential machines. *Automata studies, annals of mathematical studies, No. 34*. Princeton University Press. [JTT]
Moray, N. (1959) Attention in dichotic listening: Affective cues and the influence of instructions. *Quarterly Journal of Experimental Psychology* 11:56–60. [MVL]
(1969) *Attention: Selective processes in vision and hearing*. Hutchinson. [MVL]
Nagel, T. (1986) *The view from nowhere*. Oxford University Press (Oxford). [PCD]
Neisser, U. (1963) The multiplicity of thought. *British Journal of Psychology* 54:1–14. [CP]
(1987) From direct perception to conceptual structure. In: *Concepts and conceptual development: Ecological and intellectual factors in categorization*, ed. U. Neisser. Cambridge University Press. [KJV]
Nelson, T. O. (1976) Reinforcement and human memory. In: *Handbook of learning and cognitive processes*, vol. 3, ed. W. K. Estes. Erlbaum. [TRS]
Newell, A. (1969) Heuristic programming: Ill-structured problems. In: *Progress in operations research, III*, ed. J. Aronofsky. Wiley. [rAN]
(1973a) You can't play 20 questions with nature and win: Projective comments on the papers of this symposium. In: *Visual information processing*, ed. W. G. Chase. Academic Press. [arAN, SFC, SKC, JAM, JTT]

(1973b) Artificial intelligence and the concept of mind. In: *Computer models of thought and language*, ed. R. C. Schank & K. M. Colby. W. H. Freeman. [rAN]

(1973c) Production systems: Models of control structures. In: *Visual information processing*, ed. W. C. Chase. Academic Press. [rAN]

(1980) Physical symbol systems. *Cognitive Science* 4:135–83. [rAN]

(1983) Intellectual issues in the history of artificial intelligence. In: *The study of information: Interdisciplinary messages*, ed. F. Machlup & U. Mansfield. Wiley. [rAN]

(1989a) Putting it all together. In: *Complex information processing: The impact of Herbert A. Simon*, ed. D. Klahr & K. Kotovsky. Erlbaum. [JAM]

(1989b) How it all got put together. In: *Complex information processing: The impact of Herbert A. Simon*, ed. D. Klahr & K. Kotovsky. Erlbaum. [JAM]

(1992) Unified theories of cognition: The role of Soar. In: *Soar: A cognitive architecture in perspective*, ed. J. A. Michon & A. Akyürek. Kluwer. [rAN, JAM]

Newell, A. & Rosenbloom, P. (1981) Mechanisms of skill acquisition and the law of practice. In: *Cognitive skills and their acquisition*, ed. J. R. Anderson. Erlbaum. [aAN]

Newell, A., Rosenbloom, P. & Laird, J. (1989) Symbolic architectures for cognition. In *Foundations of cognitive science*, ed. M. Posner. MIT/Bradford Books. [rAN]

Newell, A., Shaw, J. C. & Simon, H. A. (1958) Elements of a theory of human problem solving. *Psychological Review* 65:151–66. [SKC, CP]

Newell, A. & Simon, H. A. (1972) *Human problem solving*. Prentice-Hall. [aAN, RAC, SKC, MVL, CP]

Nilsson, L.-G. & Gardiner, J. M. (1991) Memory theory and the boundary conditions of the Tulving-Wiseman law. In: *Relating theory and data: Essays on human memory in honor of Bennet B. Murdock*, ed. W. E. Hockley & S. Lewandowsky. Erlbaum. [SL]

Nilsson, L.-G., Law, J. & Tulving, E. (1988) Recognition failure of recallable unique names: Evidence for an empirical law of memory and learning. *Journal of Experimental Psychology: Learning, Memory and Cognition* 14:266–77. [SL]

Nilsson, N. J. (1983) Artificial intelligence prepares for 2001. *AI Magazine* 4/1:7–14. [SAV]

Norman, D. A. (1991) Approaches to the study of intelligence. *Artificial Intelligence* 47:327–46. [SL, RWR]

Norman, D. A. & Draper, S. W., eds. (1986) *User center system design*. Erlbaum. [JMC]

Norman, D. A. & Shallice, T. (1980) Attention and action: Willed and automatic behavior. Unpublished paper, Center for Human Information Processing, University of California San Diego. [MVL]

Parnas, L. (1985) Software aspects of strategic defence systems. *American Scientist* 73:432–30. [PCD]

Pashler, H. & Johnston, J. C. (1989) Chronometric evidence for central postponement in temporally overlapping tasks. *Quarterly Journal of Experimental Psychology* 41A:19–45. [RWR]

Payne, J. W., Bettman, J. R. & Johnson, E. J. (1992) Behavioral decision research: A constructive processing perspective. *Annual Review of Psychology* 43:87–131. [rAN]

Peck, V. A. & John, B. E. (1992) Browser-Soar: A computational model of a highly interactive task. *Proceedings of CHI'92 Conference*, May. [rAN]

Penrose, R. (1989) *The emperor's new mind*. Oxford University Press. [PCD]

Peters, R. S. (1958) *The concept of motivation*. Routledge & Kegan Paul. [PCD]

(1969) Motivation, emotion, and the conceptual schemes of common sense. In: *Human action*, ed. T. Mischel. Academic Press. [PCD]

Piaget, J. (1967) *Études sociologiques*. Droz. [CP]

Pitts, W. H. & McCulloch, W. S. (1947) How we know Universals: The perception of auditory and visual forms. *Bulletin of Mathematical Biophysics* 9:127–47. [PCD]

Polk, T. A. & Newell, A. (1988) Modeling human syllogistic reasoning in Soar. *Proceedings of the Cognitive Science Annual Conference, 1988* (August). Cognitive Science Society. [rAN]

(in preparation) A verbal reasoning theory of categorical syllogism. [rAN]

Polk, T. A., Newell, A. & Lewis, R. L. (1989) Toward a unified theory of immediate reasoning in Soar. *Proceedings of the Cognitive Science Annual Conference, 1989* (August). Cognitive Science Society. [rAN]

Polk, T. A., Newell, A. & VanLehn, K. (1991) Analysis of symbolic parameter models: A new technique for fitting models in cognitive science. Unpublished technical report, School of Computer Science, Carnegie Mellon University. [rAN]

Pomerantz, J. R., Kaplan, S. & Kaplan, R. (1969) Satiation effects in the perception of single letters. *Perception & Psychophysics* 6:129–32. [EC]

Premack, D. (1986) *Gavagai! or the future history of the animal language controversy*. MIT Press. [PCD, DQ]

Proctor, R. W. & Reeve, T. G., eds. (1990) *Stimulus-response compatibility: An integrated perspective*. North-Holland. [RWP]

Pylyshyn, Z. W. (1984) *Computation and cognition: Towards a foundation for cognitive science*. MIT Press. [JAM]

Quiatt, D. & Reynolds, V. (1992) *Primate behaviour: Information, social cognition, and the evolution of culture* (in press). Cambridge University Press (Cambridge). [DQ]

Quinlan, J. R. (1986) Induction of decision trees. *Machine Learning* 1:81–106. [rAN]

Rapaport, D., ed. (1951) *Organization and pathology of thought*. Columbia University Press. [CP]

Ratcliff, R. (1985) Theoretical interpretations of the speed and accuracy of positive and negative responses. *Psychological Review* 92:212–25. [aAN]

Reddy, R. & Newell, A. (1974) Knowledge and its representations in a speech understanding system. In: *Knowledge and cognition*, ed. L. W. Gregg. Erlbaum. [MVL]

Reisberg, D., Rappoport, I. & O'Shaughnessy, M. (1984) Limits of working memory: The digit span-digit. *Journal of Experimental Psychology: Learning, Memory and Cognition* 10:203–21. [RAC]

Rescorla, R. A. & Wagner, A. R. (1972) A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In: *Classical conditioning II: Current research and theory*. Appleton-Century-Crofts. [rAN]

Resnikoff, H. L. (1989) *The illusion of reality*. Springer-Verlag. [SKC]

Roitblat, H. L. & von Fersen, L. (1992) Comparitive cognition: Representations and processes in learning and memory. *Annual Review of Psychology* 43:671–710. [rAN]

Rolls, E. T., Baylis, G. C. & Leonard, C. M. (1985) Role of low and high spatial frequencies in the face-selective responses of neurons in the cortex in the superior temporal sulcus in the monkey. *Vision Research* 25(8):1021–35. [WMS]

Rosch, E. & Lloyd, B. B., eds. (1978) *Cognition and categorization*. Erlbaum. [rAN]

Rosenbloom, P. S., Laird, J. E., McDermott, J., Newell, A. & Orciuch, E. (1985) R1-Soar: An experiment in knowledge-intensive programming in a problem solving architecture. *Pattern Analysis and Machine Intelligence* 7:561–69. [aAN]

Rosenbloom, P. S., Laird, J. E. & Newell, A. (1987) Knowledge-level learning in Soar. *Proceedings of AAAI-87*. Morgan Kaufman. [rAN]

(1988) The chunking of skill and knowledge. In: *Working models of human perception*, ed. H. Bouma & B. A. G. Elsendoorn. Academic Press. [rAN]

Rosenbloom, P. S. & Newell, A. (1986) The chunking of goal hierarchies: A generalized model of practice. In: *Machine learning: An artificial intelligence approach II*, ed. R. S. Michalski, J. Carbonell & T. Mitchell. Morgan Kaufman. [MVL]

Rosenbloom, P. S., Newell, A. & Laird, J. E. (1991) Toward the knowledge level in Soar: The role of the architecture in the use of knowledge. In: *Architectures for intelligence*, ed. K. VanLehn. Erlbaum. [SAV]

Rumelhart, D. E., McClelland, J. L. & the PDP Research Group (1986) *Parallel distributed processing: Explorations in the microstructure of cognition*, vol. I, *Foundations*. Bradford/MIT Press. [MVL]

Salthouse, T. (1986) Perceptual, cognitive, and motoric aspects of transcription typing. *Psychological Bulletin* 99:303–19. [aAN, SKC]

Saunders, A. (1984) Ten symposia on attention and performance: Some issues and trends. In: *Attention and performance X, control of language processes*, ed. H. Bouma & D. G. Bouwhuis. Erlbaum. [SKC]

Savelsbergh, G. J. P., Whiting, H. T. A. & Bootsma, R. J. (1991) Grasping tau. *Journal of Experimental Psychology: Human Perception and Performance* 17:315–22. [KJV]

Schank, R. C. (1972) Conceptual dependency: A theory of natural language understanding. *Cognitive Psychology* 3:552–631. [CP]

(1982) *Dynamic memory: A theory of reminding and learning in computers and people*. Cambridge University Press (Cambridge). [RWR]

Schleidt, W. M. (1988) The eye illusion. *Archives of Psychology* 140(2):137–40. [WMS]

Schmidt, R. A., Zelaznik, H., Hawkins, B., Frank, J. S. & Quinn, J. T., Jr. (1979) Motor-output variability: A theory for the accuracy of rapid motor acts. *Psychological Review* 86:415–51. [RWP]

Schweickert, R. (1978) A critical path generalization of the additive factor method: Analysis of a Stroop task. *Journal of Mathematical Psychology* 18:105–39. [JTT]

Schweickert, R. & Townsend, J. T. (1989) A trichotomy: Interactions of factors prolonging sequential and concurrent mental processes in stochastic discrete mental (PERT) networks. *Journal of Mathematical Psychology* 33(3):328–47. [JTT]

Searle, J. R. (1980) Minds, brains, and programs. *Behavioral and Brain Sciences* 3:417–57. [rAN, LG]
(1990) Consciousness, explanatory inversion, and cognitive science. *Behavioral and Brain Sciences* 13:585–42. [PCD]
Seibel, R. (1963) Discrimination reaction time for a 1023-alternative task. *Journal of Experimental Psychology* 66:215–26. [MVL]
Seifert, C. M. & Hutchins, E. L. (in press) Error as opportunity: Learning in a situated task. *Journal of Human-Computer Interaction.* [RWR]
Selz, O. (1922) *Zur Psychologie des produktiven Denkens und des Irrtum.* Bonn: Cohen. [CP]
Shiffrin, R. M. & Schneider, W. (1977) Controlled and automatic human information processing, II: Perceptual learning, automatic attending, and a general theory. *Psychological Review* 84:127–90. [MVL]
Schultz, T. R. & Schmidt, W. C. (1991) A Cascade-Correlation model of balance scale phenomena. *Proceedings of the Thirteenth Annual Conference of the Cognitive Science Society.* Erlbaum. [rAN, TRS]
Siegler, R. S. (1976) Three aspects of cognitive development. *Cognitive Psychology* 8:481–520. [arAN]
(1981) Developmental sequences between and within concepts. *Monographs of the Society for Research in Child Development* 46:whole no. 189. [TRS]
Siegler, R. S. & Jenkins, E. (1989) *How children discover new strategies.* Erlbaum. [TRS]
Simon, H. A. (1991) *Models of my life.* Basic Books. [CP, KJV]
Simon, H. A. & Feigenbaum, E. A. (1964) An information-processing theory of some effects of similarity, familiarization, and meaningfulness in verbal learning. *Journal of Verbal Learning and Verbal Behavior* 3:385–96. [aAN]
Simon, H. A. & Kaplan, C. (1989) Foundation of cognitive science. In: *Foundations of cognitive science*, ed. M.-I. Posner. MIT Press. [NHF]
Simon, H. A. & Kotovsky, K. (1963) Human acquisition of concepts for sequential patterns. *Psychological Review* 70:534–46. [rAN]
Simon, T., Newell, A. & Klahr, D. (1991) Q-Soar: A computational account of children's learning about number conservation. In: *Computational approaches to concept formation*, ed. D. Fisher & M. Pazzani. Morgan Kauffman. [rAN]
Skinner, B. F. (1957) *Verbal behavior.* Appleton-Century-Crofts. [DB, CP]
Smolensky, P. (1988) On the proper treatment of connectionism. *Behavioral and Brain Sciences* 11:1–74. [rAN, EC, CP]
Snoddy, G. S. (1926) Learning and stability. *Journal of Applied Psychology* 10:1–36. [SKC]
Solomon, H. Y., Turvey, M. T. & Burton, G. (1989) Perceiving extents of rods by wielding: Haptic diagonalization and decomposition of the inertia tensor. *Journal of Experimental Psychology: Human Perception and Performance* 15:58–68. [KJV]
Sondheimer, N. K. (1989) The rate of progress in natural language processing. In: *Theoretical issues in natural language processing*, ed. Y. Wilkes. Erlbaum. [SAV]
Sowa, J. F. (1991) Towards a reintegration of AI research. In: *Future directions in artificial intelligence*, ed. P. A. Flach & R. A. Meersman. North-Holland. [SAV]
Sperber, D. (1984) Anthropology and psychology: Towards an epidemiology of representations. *Man* 20:73–89. [CP]
Sperling, G. (1960) The information available in brief visual presentations. *Psychological Monographs* 74(11): whole no. 498. [rAN]
Steier, D. M. (1989) Automating algorithm design within an architecture for general intelligence. Ph.D. dissertation, Computer Science, Carnegie-Mellon University. [aAN]
Sternberg, S. (1966) High-speed scanning in human memory. *Science* 153:652–54. [MVL, SL, JTT]
(1975) Memory scanning: New findings and current controversies. *Quarterly Journal of Experimental Psychology* 27:1–32. [aAN]
Sternberg, R. J. & Detterman, D. K., eds. (1979) *Human intelligence, Perspectives on its theory and measurement.* Ablex. [rAN]
Thorndike, E. L. (1903) *Educational Psychology.* Lemke & Buechner. [aAN]
Townsend, J. T. (1971) A note on the identifiability of parallel and serial processes. *Perception & Psychophysics* 10:161–63. [JTT]
(1975) The mind-body equation revisited. In: *Philosophical aspects of the mind-body problem*, ed. C.-Y. Cheng. [JTT]
(1990) Serial vs. parallel processing: Sometimes they look like Tweedledum and Tweedledee but they can (and should) be distinguished. *Psychological Science* 1:46–54. [rAN, SL, JTT]
Townsend, J. T. & Ashby, F. G. (1983) *Stochastic modeling of elementary psychological processes.* Cambridge University Press. (Cambridge). [JTT]
Townsend, J. T. & Schweickert, R. (1985) Interactive effects of factors affecting processes in stochastic PERT networks. In: *Cognition, information processing, and motivation*, ed. G. d'Ydewalle. North-Holland. [JTT]
Treisman, A. & Gelade, G. (1980) A feature integration theory of attention. *Cognitive Psychology* 12:97–136. [rAN]
Treisman, A. & Gormican, S. (1988) Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review* 95:15–48. [rAN]
Treisman, A. & Schmidt, H. (1982) Illusory conjunctions in the perception of objects. *Cognitive Psychology* 14:107–41. [rAN]
Tulving, E. (1983) *Elements of episodic memory.* Oxford University Press. [arAN]
Urmson, J. O. (1956) *Philosophical analysis: Its development between the two World Wars.* Clarendon Press. [PCD]
Uttal, W. R. (1990) On some two-way barriers between models and mechanisms. *Perception & Psychophysics* 48:188–203. [WRU]
Uttal, W. R., Bradshaw, G., Dayanand, S., Lovell, R., Shepherd, T., Kakarala, R., Skifsted, K. & Tupper, G. (1992) *The swimmer: An integrated computational model of a perceptual-motor system.* Erlbaum. [WRU]
VanLehn, K. (1991) Rule acquisition events in the discovery of problem-solving strategies. *Cognitive Science* 15:1–47. [TRS]
Vere, S. A. & Bickmore, T. W. (1990) A basic agent. *Computational Intelligence* 6:41–60. [SAV]
von Helmholtz, H. (1971) The origin and correct interpretation of our sense impressions. In: *Selected writings of Hermann von Helmholtz*, ed. R. Kahl. Wesleyan University Press. [LG]
von Neumann, J. (1966) *Theory of self-reproducing automata.* University of Illinois Press. [LG]
Wallechinsky, D., Wallace, I. & Wallace, A. (1977) *The book of lists.* William Morrow. [rAN]
Warren, W. H., Jr. & Hannon, D. J. (1988) Direction of self-motion is perceived from optical flow. *Nature* 336:162–63. [KJV]
Warren, W. H., Jr. & Whang, S. (1987) Visual guidance of walking through apertures: Body-scaled information about affordances. *Journal of Experimental psychology: Human Perception and Performance* 13:371–83. [KJV]
Watanabe, S. (1985) *Pattern recognition: Human and mechanical.* Wiley. [LG]
Wertheimer, M. (1945) *Productive thinking.* Wiley. [CP]
Wetherick, N. E. (1989) Psychology and syllogistic reasoning. *Philosophical Psychology* 2:111–24. [rAN, NEW]
Wetherick, N. E. & Gilhooly, K. J. (1990) Syllogistic reasoning: Effects of premise order. In: *Lines of thinking*, vol. 1, ed. K. J. Gilhooly, M. T. G. Keane, R. H. Logie & G. Erdos. Wiley. [rAN, NEW]
Whorf, B. L. (1956) A linguistic consideration of thinking in primitive communities. In: *Language, thought, and reality – selected writings of Benjamin Lee Whorf*, ed. J. B. Carroll. MIT Press. [WMS]
Wiesmeyer, M. D. (1992) *An operator-based model of human covert visual attention.* Ph.D. dissertation, University of Michigan, CSE-TR-123-92. [rAN]
Wiesmeyer, M. & Laird, J. (1990) A computer model of 2D visual attention. *Proceedings of the Annual Conference of the Cognitive Science Society*, July. [rAN]
Wilson, E. O. (1975) *Sociobiology: The new synthesis.* Harvard University Press. [rAN]
Winograd, S. (1964) Input-error-limiting automata. *Journal of the Association of Computing Machinery* 11:338–51. [JTT]
Witten, E. (1988) Interview with Edward Witten. In: *Superstrings: A theory of everything?* ed. P. C. W. Davies & J. Brown. Cambridge University Press. [LG]
Wolfe, J. M., Cave, K. R. & Franzel, S. L. (1989) Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance* 15:419–33. [rAN]