COGNITIVE SCIENCE 17, 87-116 (1993)

Situated Action: A Neuropsychological Interpretation Response to Vera and Simon

WILLIAM J. CLANCEY
Institute for Research on Learning

1. SYMBOLS INSIDE AND OUTSIDE: A CATEGORY ERROR

> The most fundamental contribution so far of artificial intelligence and computer science to the joint enterprise of cognitive science has been the notion of a physical symbol system, i.e.,. . .systems capable of having and manipulating symbols, yet realizable in the physical universe.
>
> **It becomes a hypothesis that this notion of symbols includes symbols that we humans use everyday in our lives.** (Newell, 1980)

Situated cognition research rejects the hypothesis that neurological structures and processes are similar in kind to the symbols we create and use in our everyday lives. The symbolic approach, as described by Vera and Simon (1993), conflates *neurological structures and processes* with *physical representations* that we perceive and manipulate in our environment (e.g., a journal article) and *experiences of representing in our imagination* (e.g., visualizing or talking to ourselves). This category error distorts the nature of perception, the nature of conceptual interpretation as we comprehend written plans, and in general the adaptive nature of every thought and action. At its heart, the symbolic approach confuses an agent's deliberated action—in sequences of behavior over time, as cycles of reperceiving and commenting (e.g., in speech, writing, drawing)—with what occurs *within* every ongoing coordination (Dewey, 1896/1981a).

Put another way, the symbolic approach conflates "first-person" representations in our environment (e.g., utterances and drawings) with "third-person" representations (e.g., mappings a neurobiologist finds between

Development of these ideas was greatly helped by discussions at the McDonnell Foundation conference on "The Science of Cognition," particularly conversations with Mark Bickhard, Ken Ford, John Haugeland, and Peter Slezak. Funding has been provided in part by grants from the Xerox Foundation and NSF.

Correspondence and requests for reprints should be sent to William J. Clancey, Institute for Research Learning, 2550 Hanover St., Palo Alto, CA 94304.

sensory surfaces and neural structures; Slezak, 1992). Saying that "the information in DNA and RNA is certainly represented symbolically" (Vera & Simon, 1993, p. 44) obscures what happens when an agent represents something or interprets something symbolically. DNA sequences aren't symbols in the first-person sense because they are not interpreted by the agent possessing them as meaning something. But a scientist looking inside can view DNA sequences as symbols in the third-person sense, by showing that they constitute a code that "contains information" about the agent's phenotype[1]. By ignoring the shift in frame of reference between the agent and the scientist looking inside, Vera and Simon ignore the role of perceiving and coordinating action in creating and using representations. Indeed, saying "the way in which symbols are represented in the brain is not known" (p. 9) oddly mixes the two points of view (whose symbols are represented?).

In effect, the symbolic approach obscures the nature of representations. A corollary is that the symbolic approach inadequately characterizes the nature, origin, and value of symbolic models like SOAR (Clancey, 1991b). Reformulating the relation of symbolic cognitive models to human behavior neither repudiates their value for understanding human behavior, nor the value of symbolic modeling techniques (Clancey, 1992a, 1992b). We must separate the content of cognitive models and meta-theoretic claims made about them from the modeling methods. We could, for example, use classification hierarchies or state transition networks to model neural interactions, without claiming that these models are also stored in the brain. More to the point, the claims of situated cognition lead us to reformulate knowledge acquisition, the process of constructing an expert system (Clancey, 1989). Rather that viewing it as transfer of expertise—that is, extracting "knowledge" already prerepresented, in stored structures in the expert's memory (e.g., stored rules or semantic nets)—we view it as a process of creating representations, inventing languages, and, in general, formulating models for the first time. In effect, reinterpreting the meaning of cognitive models relative to human memory leads us to reexamine the scientific process of model construction (Gregory, 1988), the relation of models and everyday activity (Lave, 1988), and the use of models in instruction (Schön, 1987).

Vera and Simon (1993) state several times that "complex human behavior . . . can be and has been described and simulated effectively in physical symbol systems" (p. 46). Citing an example of a human driving a car, they say, "the human part of this sequence of events. . .can be modeled by a symbolic pattern recognition cum production system" (p. 18). But situated cognition does not argue that "humans and their interactions with the world cannot be understood using symbol-system models and methodology" (p. 7). The

[1] To be clear, such a mapping (e.g., "a sequence of amino acids, CAGCTA corresponds to such and such") is a first-person representation to the scientist; that is, it is an expression the scientist creates and perceives as having meaning.

issue is not understanding or modeling, per se. The claim is that the model is merely an abstraction, a description and generator of *behavior patterns over time*, not a mechanism equivalent to human capacity.

Symbolic models have explanatory value as psychological descriptions. In stable environments, a symbolic model can serve to control a useful robot agent, like Navlab. Regardless of what we later understand about the brain or cognition, such models are unlikely to go away. For example, descriptions of patterns of behavior, in terms of goals, beliefs, and strategies, are especially valuable for instruction (e.g., Clancey, 1988). But arguing that all behavior can be *represented* as symbolic models (Vera & Simon, 1993, p. 46) misses the point: We can model anything symbolically. But what is the residue? What can people do that today's computer programs cannot? What remains to be replicated?

Our goal of understanding the brain and replicating human intelligence with computers will no longer be served well by using the term "symbol" (or "representation") to refer interchangeably to experiences of representing something to oneself, to neural structures/processes, and to forms in the world such as words in a computer program. Saying that "Brooks's [1991] creatures are very good examples of orthodox symbol systems" (Vera & Simon, 1993, p. 34) makes "symbol system" of little heuristic value for robot design. This meaning of "symbol system" lumps together vending machines, conventional expert systems, situated automata (Maes, 1990), SOAR, cats, and people. What then does "symbol system" explain? What value does it provide for better understanding the human brain? Perhaps Vera and Simon would say it moves us out of the behaviorist camp, but then they are still fighting the last battle. The distinction is too coarse for improving mechanisms today. Navlab's neural network is clearly not the kind of computational mechanism described in *Human problem solving* (Newell & Simon, 1972). Agre (1988), Brooks (1991), Rosenschein (1985) and the connectionists are moving us away from a flat "symbolic" architecture, leading us as designers to distinguish between our interpretations of structures inside the robot, and the robot's interpretations of structures it perceptually creates and manipulates (Clancey, 1991b).

Some of the interesting questions for advancing cognitive science today concern how children invent representations (Bamberger, 1991), how immediate behavior is adaptive (Laird & Rosenbloom, 1990), the role of metaphor in theory formation (Schön, 1979), the social organization of learning (Roschelle & Clancey, in press), the relation of affect and belief (Iran-Nejad, 1984), and so on. These psychological issues are raised by the progenitors of situated action, in work ranging from Dewey (1938, 1896/1981a, 1938/1981b), Collingwood (1938), and Bartlett (1932/1977) to Piaget (1960/1981), Vygotsky (1986), and Bateson (1972). But we needn't look so far afield: Recent work in neuroscience, language, and learning discusses the relation between situated cognition and symbolic theories (e.g., Bickhard & Richie, 1983; Edelman, 1992; Lakoff, 1987; Rosenfield, 1988; Sacks, 1987), providing a useful starting point for understanding social critiques (Lave, 1988; Suchman, 1987).

I agree with Vera and Simon that SA research needs to be reformulated in terms useful to psychologists and cognitive scientists. There is no single answer for what "situated" means to researchers in diverse fields such as robotics, psychology, education, or organizational theory. Each community will make their own interpretation. Necessarily, my assertions defining situated cognition in terms of neuropsychology may be uninteresting or incomprehensible to Lave or Suchman.[2] My only goal is to do justice to their insights in a way useful to cognitive science and AI.

Five central claims of SA are summarized here. In subsequent sections of the article, I show how these claims can be grounded in neuropsychology, that is, relating cognitive theories of problem solving, knowledge, memory, and so on, to neurological structures and processes.

1. The Representation Storehouse View of Memory Confuses Structures in the Brain with Physical Forms That Are Created and Used in Speaking, Drawing, Writing, etc. Assuming that the researcher's conjectured representational language and models preexist in the subject's brain has led cognitive science to deemphasize how people create representations everyday, reducing learning to syntactic modification of the modeler/teacher's presupplied ontology of standard notations (Iran-Nejad, 1984; Rosenfield, 1988). Representing, comprehending, meaning, speaking, conceiving, and so forth, are reified into acts of manipulating representations in a hidden way called "reasoning" (Ryle, 1949). Stored-schema models tend to view *meaning* as mapping between given information and stored conceptual primitives, facts, and rules; *activity* as executing rules or scripts; *problems* as given descriptions in an existing, shared notation (representation language); *information* as given, selected, or filtered from the environment; and *concepts* as stored descriptions, such as dictionary definitions (Lakoff, 1987; Reeke & Edelman, 1988).

2. Schema Models Wrongly View Learning as a Secondary Phenomenon, Necessarily Involving Representation (Reflection). No human behavior is strictly rote (Bartlett, 1932/1977). Learning occurs with every act of seeing and speaking. Categories are not stored things, but always adapted ways of talking, seeing, relating, or in general, ways of coordinating behavior. A

[2] Even today's "cognitive neuropsychologists" make little contact with symbolic theories. For simplicity, I will refer interchangeably to "situated cognition research" and "situated action theory" (abbreviated "SA"), not as a historical description of a uniform community or set of shared beliefs and theories, but as a useful synthesis.

person interpreting a recipe, diagram, or journal article is conceiving, not merely retrieving and assembling primitive meanings and definitions according to rules of grammer and discourse structure (Collingwood, 1938; Suchman, 1987; Tyler, 1978). Information is created by the observer, not given, because comprehending is conceiving, not retrieving and matching (Gregory, 1988; Reeke & Edelman, 1988).

3. Integration of Perceiving and Moving and Higher Order Serial Organizations Is Dialectic—Coherent Subprocesses Arise Together—Not Via Linear Causality or Parallelism. Perceiving, thinking, and moving always occur together, as coherent coordinations of activity (Dewey, 1896/1981a). Representing can occur within the brain (e.g., visualizing), but always involves sensorimotor aspects and is interactive, even though it may be private (i.e., representing without use of representational things in the environment). Representing is always an act of perceiving; we can't strictly separate perceiving and moving (e.g., when speaking I am also comprehending what I am saying), coordinations are circuits.

4. Practice Cannot Be Reduced to Theory. The source of cultural commonalities is not a set of common laws, grammars, and behavior schemas stored in individual brains (Lave, 1988). There is no locus of control in human activity, either neurobiological or social. Our ability to coordinate our activities without mediating theories is the foundation for our ability to agree on courses of action and theorize in similar ways (rather than the other way around; Lakoff, 1987; Lave & Wenger, 1991). Social processes involve organization of a larger system, transcending individual awareness and control, not merely transferring representations (as in speech). Models of social interactions are not required for their occurrences and reproduction [cf. Bartlett's (1932/1977) discussion of soccer, p. 277].

> Efficient practice precedes the theory of it; methodologies presuppose the application of the methods, of the critical investigation of which they are the products. It was because Aristotle found himself and others reasoning now intelligently and now stupidly and it was because Izaak Walton found himself and others angling sometimes effectively and sometimes ineffectively that both were able to give their pupils the maxims and prescriptions of their acts. It is therefore possible for people intelligently to perform some sorts of operations when they are not yet able to consider any propositions enjoining how they should be performed. (Ryle, 1949, p. 30)

> The difficulty with the identification of rule and intuition is that it implies that the knowledge in our heads is more explicit than the practices that might reveal it. If rules are discursive directions for action then they cannot be "nuances" or feelings of "subtle relations." (Tyler, 1978, p. 153).

Indeed, as we perceive patterns and articulate theories to explain them, we become increasingly alienated from the complexity of activity itself (Tyler, 1978). That is, as we perceptually abstract nature and human behavior in pattern descriptions (e.g., discourse patterns), and explain the occurrence of such regularities in theoretical laws (e.g., "facility conditions" explaining why discourse patterns exist), we become more distant from the phenomena we are studying. The apparent stability and completeness of our theories is always relative to our artifacts, ongoing purposes, and activities.

5. Situated Cognition Relates to Ideas in the Philosophy of Science, Concerning the Nature of Mechanisms and Pattern Descriptions. Several ideas about theories are intricately related in stored-schema models of cognition:

1. A simple view of science we learn in school is that regularities in nature are caused by laws (e.g., a falling object accelerates so as to obey the law F=MA);
2. Scientific theories are the epitome of knowledge (e.g., the emphasis on "deep models" in AI research);
3. Rational behavior obeys general laws of logic such as modus ponens; therefore,
4. Human knowledge consists of facts and laws stored in memory, causing observed regularities in behavior.

Many people have attacked these ideas, suggesting, in addition, that they are distorting human activity as diverse as architectural design (Alexander, 1979), organizational learning (Nonaka, 1991), professional training (H. Dreyfus & S. Dreyfus, 1986; Schön, 1987), musical invention (Bamberger, 1991), design of complex devices (Suchman, 1987), use of computers in business (Winograd & Flores, 1986), and scientific progress itself (Gleick, 1987).

To understand the SA argument, we should consider up front the common, but misleading misinterpretations of what "situated" means. Some of these interpretations arise because social scientists are trying to make contact with cognitive theories, but lack the background to speak in the language a cognitive scientist expects. Situated cognition research is:

1. Not merely about an agent "located in the environment" "strongly interactive," or "real time" (e.g., referring to Larkin's 1989 "display-based" system as situated, Vera & Simon, 1993, p. 32), rather a claim about the internal mechanism that coordinates sensory and motor systems.
2. Not rejecting the value of planning and representations in everyday life, rather seeking to explain how they are created and used in already coordinated activity.

3. Not claiming that *representing* does not occur internally, in the individual brain (e.g., imagining a scene or speaking silently to ourselves), rather seeking to explain how perceiving and comprehending are coorganized.
4. Not in itself a prescription for learning ("situated learning"), rather claiming that learning is occurring with every human interaction.
5. Not disputing the descriptive value of schema models for cognitive psychology, rather attempting to explain how such regularities develop in behavior, and what flexibility for improvisation and basis in previous, prelinguistic sensorimotor grounding is not captured by symbolic models.
6. Not claiming that current computer programs can't construct a problem space or action alternatives, rather revealing how programs can't step outside prestored ontologies.

To understand the implications of SA for psychology, we must first take away the memory of stored first-person representation structures. Second, we must replace the CPU view of processing, involving peripheral perception and motor systems, by a dialectic mechanism that simultaneously coordinates perception–action. Third, we must move deliberation out to the behavior of the agent, involving cycles of perception and action. Fourth, we must view learning not as a process of storing new programs, but as part of adaptive recoordination that occurs with every behavior, on top of which reflective representation (framing, story telling, and theorizing) is based (Schön, 1990).

In subsequent sections, I explain the dialectic view of neural architecture and its implications for theories of representation and learning. In Section 5 I summarize the separate lines of arguments from the external (social) and internal (cognitive) perspectives, using a table to summarize the sometimes dramatic shifts in our understanding of meaning, concepts, and memory. I conclude by summarizing the broader implications, which are already well underway in studies of organizational learning, business process modeling, and design of complex computer systems.

2. THE DIALECTIC VIEW: COORDINATING WITHOUT DELIBERATION

SA claims about the nature of deliberation and learning can be given a coherent neuropsychological interpretation, which both explains why symbolic cognitive models fit human behavior and how a nonsymbolic cognitive architecture could be the basis of intelligence (Edelman, 1992). I first contrast the symbolic view of coordination with a dialectic view. I then discuss

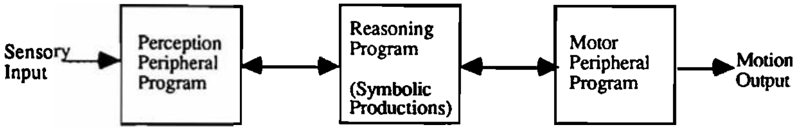

Figure 1. Symbolic view of sensorimotor coordination.

the properties of this architecture with respect to learning, control, immediate behavior, and deliberation.

The symbolic approach distinguishes between data and program, software and hardware, serial and parallel processing. A basic assumption is that perception and reasoning are possible without acting. Reasoning is contrasted with immediate, nondeliberated behavior. Even the parallel distributed view describes data as processed independently in different modules, with results passed to other modules downstream (Farah, 1992). In attempting to put the pieces of perception, cognition, and motor operations back together, it is unclear how to make reflexes sensitive to cognitive goals (Laird & Rosenbloom, 1990; Lewis et al., 1990). This dichotomy arises because deliberation is opposed with, and in the architecture disjoint from, immediate perception-action coordination. Rather than viewing immediate perception-action coordination as the basis of cognition, cognition is supposed to supplant and improve upon "merely reactive" behavior.

According to the neuropsychological interpretation of SA, *the neural structures and processes that coordinate perception and action are created during activity*, not retrieved and rotely applied, merely reconstructed, or calculated via stored rules and pattern descriptions (Bartlett, 1932/1977; Edelman, 1992; Freeman, 1991). That is, the physical components of the brain, at the level of neuronal groups of hundreds and thousands of neurons, are always new—not predetermined and causally interacting in the sense of most machines we know—but coming into being during the activity itself, through a process of reactivation, competitive selection, and composition (Edelman, 1987; Smoliar, 1989). In particular, this architecture coordinates perception and action without intermediate decoding and encoding into descriptions of the world or of how behavior will appear to an observer, thus avoiding combinatoric search and matching (as well as problems of symbol grounding and the frame axiom problem; Bickhard & Richie, 1983). With such a self-organizing mechanism—nonserial, nonparallel, but dialectic—coordination and learning are possible without deliberation (Dewey, 1896/1981a).

As Figure 2 illustrates, SA has indeed led Edelman (1992) "down novel paths, which cannot be followed along traditional information-processing lines" (Vera & Simon, 1993, p. 23). In contrast with the classical, symbolic architecture, *the processors coconfigure each other.* All action is embodied because perception and action arise together automatically: Learning is inherently "situated" because every new activation is part of an ongoing perception-action coordination. Situated activity is not a kind of action, but the nature of animal interaction at all times, in contrast with most machines we know. This is not merely a claim that context is important, but what constitutes the context, how you categorize the world, *arises together* with processes that are coordinating physical activity. To be perceiving the world is to be acting in it—not in a linear input-output relation (act-observe-change) —but dialectically, so that what I am perceiving and how I am moving co-determine each other. According to Edelman, this is accomplished by global maps:

> A global mapping is a dynamic structure containing multiple reentrant local maps (both motor and sensory) that are able to interact with nonmapped parts of the brain...allows selectional events occurring in its *local* maps to be connected to the animal's motor behavior, to new sensory samplings of the world, and to further successive reentry events.
>
> Categorization does not occur according to a computerlike program in a sensory area which then executes a program to give a particular motor output. Instead, sensorimotor activity over the whole mapping *selects* neuronal groups that give the appropriate output or behavior, resulting in categorization... always occurs in reference to internal criteria of value....(Edelman, 1992, p. 89)

We can view the cortex and neural structures in general as the hardware, but there are no fixed structural components by which symbols are stored, transferred, or manipulated as structures, in the sense of data structures in a computer memory. Rather, the software is the hardware being activated and configured "in line" as part of the living neural connection between sensory and motor systems. A *schema* is the hardware being chronologically and compositionally activated, by a process that is always adapted to other ongoing coordinations, and is always a generalization of previous coordinations (Bartlett, 1932/1977; Vygotsky, 1934/1986). (In contrast, by the symbolic view schemas are first-person *representations* of how the world or behavior appears, e.g., beliefs, scripts, strategies.)

Every activation reinforces physical connections, biasing active hardware to be reincorporated in future compositions, bearing the same temporal relations to perceptual and conceptual maps. Regularities of behavior, including an observer's construction of analogies, develop because every perceptual motor coordination—in both agent and observer—generalizes, includes, and correlates previous perceptions and coordinations. The nesting of boxes in Figure 2 illustrates how "reentrancy" or the feedforward aspect of past organizations (categories and sequential coordinations of multimodal sensory and motor systems) biases the present, ongoing activation.

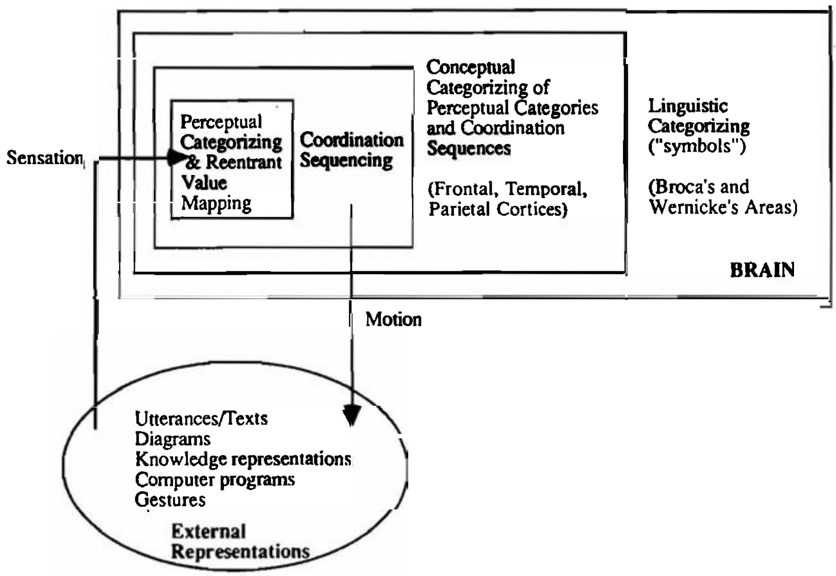

Figure 2. Situated action view of sensorimotor coordination (adapted from Edelman, 1992). Placement of "motion" line is arbitrary.

Other interesting aspects of this architecture, yet to be explained in any detail, include: Impasses in sequential processes feedback to produce awareness of failures; multimodal coordinations may occur simultaneously; acts of imagination bias future coordination; speaking occurs dialectically with self-comprehension, and so on. Edelman (1992) also emphasizes the correlation of internal homeostatic systems (via brain stem, hypothalamus, etc.) in the hippocampus, amygdal, and septum prior to incorporation in global mapping-coordination sequencing and conceptualization. Thus, categorizations of "value" are integrated with the lowest levels of perceptual categorizing, making emotion integral, not a secondary coloration or distortion.

Everything stated here is posed as a highly simplified set of interleaving hypotheses about neural processes and psychological effects pertaining to memory, learning, and language. Some PDP models in cognitive neuropsychology provide a step in this direction (Farah, 1992), but most assume that linguistic symbols are input at the lowest level, omitting the foundation of prelinguistic categorization and perception–action coordination (Rosenfield, 1988). Brooks's (1991) systems (see Vera & Simon, 1993, p. 33) are, at a cercertain level, mechanisms that coordinate behavior without intervening production and interpretation of first-person symbolic descriptions. The subsumption architecture bears some relation to Figure 2, with each layer constraining the others, and compositional ordering, so higher-order networks are only active when the lower sensorimotor organizations they bias become active. Because the repertoire of sensorimotor coordinations is fixed and indeed prewired, Brooks's robots model the self-organization of the human brain, but in lacking the ability to recreate previous sequential coordinations, they do not replicate human learning or the plasticity of sensorimotor recombination.

The idea of "incrementally...adding new control layers to simple systems that already work" (Vera & Simon, 1993, p. 33) adopts an evolutionary view of engineering design, espoused by both Braitenberg (1984) and Edelman (1992). The idea is to understand what nonsymbolic mechanisms could support higher order capabilities. This avoids the category error of building in higher order capabilities into the foundation from which they must be generated (Bickhard, 1992). The navigation, nest-building, hunting, and social relations of other animals strongly suggest that linguistic, first-person symbolizing is unnecessary for learning such behaviors. In contrast, the symbolically mediated view states:

> A symbol must be stored in memory by the perceptual encoding system to activate the production by satisfying the condition, "if the road curves to the left." Then, the motor response starts with a symbol ("turn to the left") that is transmitted to the motor system.... (Vera & Simon, 1993, p. 19)

This stored computer program terminology of storage and transmission suggests that all perceiving and acting is mediated by language, in the sense that first-person symbolic representations of the world are necessary to coordinate sensory and motor systems. The dialectic view is that symbolizing processes subsume and always involve prelinguistic processes (Figure 2). Sensorimotor recoordination occurs automatically; indeed, for the animal running through a forest, navigation occurs as a matter of course without symbolizing (i.e., without language, whether verbal or pictorial). Note that *conceptualizing* ["categorizing parts of past global mappings according to modality, the presence or absence of movement, and the presence or absence of relationships between perceptual categorizations...a mapping of types of maps" (Edelman, 1992, p. 109)] is still prelinguistic; that is, it is not constituted by syntactically ordered, conventional symbols like words or drawings (Lakoff, 1987). Rather than assuming that symbolic reasoning supplanted entirely prehuman emotional, reactive behavior, we ask instead how prelinguistic mechanisms are further organized by additional neural processes, which both depend on and bias perceptual categorization, coordination, and conceptualization.

This view of neuropsychology is inspired by and seeks to explain a variety of well-known cognitive phenomena that symbolic theories inadequately explain or ignore:

- Regularities develop in human behavior without requiring awareness of the patterns, that is, without first-person representations, such as grammars or strategy rules (Ryle, 1949).

- People speak idiomatically, in ways grammars indicate would be nonsensical (Collingwood, 1938).
- Both speech and writing require cycles of revision and rephrasing, rather that just outwardly expressing what has been prestated subconsciously (Tyler, 1978).
- We experience interest, a sense of similarity, and value before we create representations to rationalize what we see and feel (Dewey, 1896/1981a; Schön, 1979).
- Emotions provide an encompassing orientation for focusing interest and resolving an impasse (Bartlett, 1932/1977).
- Figure/ground perceptual reorganization and action reorientation are rapid and apparently coconstructed (Arbib, 1981; Freeman, 1991).
- Linguistic concepts are grounded in perceptual resemblance and function in our activity, constrained by, but not restrictively defined by other concepts (Lakoff, 1987; Wittgenstein, 1958).
- Know-how is at first inarticulate and disrupted by reflection (Zuboff, 1988).
- Development involves stage effects, levels of representationality (Bickhard, 1992).
- Remembering is aided by reexperiencing images and physical orientation (Bartlett, 1932/1977).
- Every thought is a generalization (Vygotsky, 1934/1986).
- Dysfunctions often involve inability to coordinate, not loss of modular, separable capabilities, or knowledge (Rosenfield, 1988; Sacks, 1987).
- Immediate behavior is adapted, not merely selected from prepared possibilities (Bartlett, 1932/1977).
- Modular specialization of the brain correlates with the dialectic relation between perceptual categorization and sequencing of behaviors (Edelman, 1992; Freeman, 1991).

From this perspective, situated cognition seeks to reintegrate psychological theories of physical and cognitive skills, uniting emotions, reasoning, and development, in a neurobiologically grounded way.

3. REPRESENTING AND COMPREHENDING: BREAKING OUT OF CONCEPTUAL SCHEMAS

Edelman describes language in terms of symbolic categorizing, using the term "symbol" in the first-person sense as a perceptual categorization used by an agent to refer to something. When SA researchers say that there are no internal representations or symbols, they mean in the first-person sense, as structures created, interpreted, and manipulated by the person. Symbolizing occurs in our behavior as we utter sentences (sometimes to ourselves), comment on what images mean, revise written paragraphs, amend architectural drawings, and so on. Symbolic categorizing incorporates phonological and syntactic (conceptual ordering) constraints. Crucially, this capability arises out of *prelinguistic coordinations—perceptual and conceptual categorization*. This grounding implies that the brain does not contain or require conceptual or meaning primitives (Edelman, 1992; Lakoff, 1987).

Articulating (symbolizing) one's behavior and beliefs provides a way to reorient what would otherwise be automatic responses and automatic ways of dealing with impasses. By putting things out into our environment (or imagination), we can break the loop of merely reflexive action. Inquiry is interactive, a coordinated process that goes on in our behavior over time, as we reperceive, reshape, and reinterpret, material forms by which we model the world and our actions (Bamberger & Schön, 1983).

> We reject the "no man's land" of words imagined to lie between the organism and its environmental objects in the fashion of most current logics, and require, instead, definite locations for all naming behaviors as organic-environmental transactions under observation. (Dewey & Bentley, 1949, p. 121).

The combination of new juxtapositions in physical space, with the possibility of new perceptual categorization, allows us to interactively, *deliberately*, form new symbolizations, and hence new conceptualizations, new ways of talking, new ways of seeing the world, new ways of coordinating our behavior. Indeed, Bartlett (1932/1977) wrote that this is the function of consciousness:

> An organism has somehow to acquire the capacity to turn around upon its own "schemata" and to construct them afresh. This is a crucial step in organic development. . . . It is where and why consciousness comes in. . . . (p. 206)

Vera and Simon (1993) write: "the difference is not consequential" that in GPS "the 'motor action' modified an internal problem representation rather than the external one" (p. 20). We can view manipulating a problem representation inside a computer as a motor action. But it is important to distinguish internal structure modifications within a perception–action coordination from *reshaping and reperceiving a representation in a sequence of activity.*

SA does not dispute that "a human needs symbolic reasoning to guide reactive processes" (Pomerleau, Gowdy, & Thorpe, 1991, quoted in Vera & Simon, 1993, p. 28). Quite the contrary: SA elevates the creation and use of representations to the primary object of study (e.g., see Suchman's 1987 study of people interpreting photocopier instructions). Reactive processes are strictly unguidable in themselves, but over a series of reflective reperceiving acts (themselves reactive), new organizations are composed. SA thus involves two kinds of learning: automatic recoordination (primary) and reflection involving representing (secondary).

Arguments about PDP versus symbolic internal processors miss this crucial distinction. "A gradual shift *from* SA to more planful action" (Vera & Simon, 1993, p. 32) is the wrong dichotomy. We are always situated because that is how our brains work. We are situated in an empty dark room, we are situated in bed when dreaming. People doing the Tower of Hanoi problem are always situated agents, regardless of how they solve the problem. SA is a characterization of the mechanism, of our embodiment, not a problem-solving strategy [cf. subjects' use of a generative rule "can hardly be regarded as SA" (p. 32)].

The Navlab network's "source of symbolic knowledge to plan and execute a mission" (Pomerleau et al., 1991, quoted in Vera & Simon, 1993, p. 28) must not be equated with either neural structures or a person interpreting plans (again, our three distinctions). First-person interpretation of a representation (e.g., reading a map or plan) is inventive, not a process of retrieving definitions or rotely reciting meaning. Human interpretation (and hence *use* of symbols, plans, productions, etc.) is not ontologically bound (Clancey, 1987; Winograd & Flores, 1986); comprehending isn't just manipulating symbolic categories. Every thought is a generalization, involving primary learning grounded in reconfiguration of prelinguistic categorizations and coordinations.[3]

Following Dewey (1938), SA views representations as tools to solve problems. For example, a knowledge engineer creates and stores representations in the knowledge base of an expert system. Beyond just conceptualizing—something other animals can do—humans create linguistic structures that objectify, name, classify, order, and justify relationships. Reflection includes acts of framing what we are doing, telling causal stories to link elements into a scene, project future events, and plan our future actions (Schön, 1990). Again, symbolizing inherently involves acting:

> Language is an activity; it is expressing oneself, or speaking. But this activity is not what the grammarian analyses. He analyses a product of this activity, "speech" or "discourse" not in the sense of a speaking or a discoursing, but in the sense of something brought into existence by that activity. (Collingwood, 1938, p. 254)

We may draw, write, speak, gesture, or play music. Within a sequence, each "act" occurs automatically, as part of the process by which new perception-action sequences are coordinated. In the case of natural and graphic lan-

[3] Indeed, the inadequacy of the retrieve and match model leads some researchers to say that human comprehension and speaking "cannot be understood or modeled" using the symbolic approach. They are calling attention to the residue, what remains to be explained. This is also what Winograd and Flores (1986) meant by the claim that "information-processing theory and methodology cannot account for behavior" (Vera & Simon, 1993, p. 21). "Skill and expert performance cannot be captured as a set of formal rules" (Greenbaum & Kyng, 1991, p. 13) means we must not equate the brain's mechanism with the model, the map with the territory.

guages, the sequences of behavior produce compositions that include and exclude elements, chain parts of a story, express causal relationships, and model a situation. As such, symbolizing builds on and relies on our capability to coordinate automatically multiple concerns in our behavior: multiple modes of perception, multiple values and goals on different levels, and, of course, syntactic and semantic relations. A second-order benefit is that our expressions, when placed in the external environment, can be preserved and reinterpreted later (e.g., using a trip diary to plan another visit to the same location) and, especially, shared with other people (e.g., maps, journal articles, texts). In this sense, referring to a box of camping gear in the closet, or a list of what to take on the next trip, SA theorists say that knowledge is not strictly inside the head, but a coupling between adapted neurological processes and how we have structured our environment.

Representing (e.g., saying to myself what I will do next) recoordinates behavior, not via symbolic "instructions" (coded tokens passed serially downstream to other subsystems), but via the process of representing itself. That is, articulating is a process of physically recoordinating how we see and act. Behavior is mediated temporally and causally by such acts, in the sense that my speaking now affects what I do later. Contrary to the symbolic view, behavior is not mediated in the sense of a plan or recipe or rule that must be later read, related to my current activity, and translated into motor commands to be causally efficacious: It already has that effect because of my comprehending. Models involving "compiling" symbolic instructions into symbolic rules for later more-automatic processing are no better, for they still view learning as strictly linguistic manipulation.

Deliberating and using plans goes on in my activity. If time allows and the occasion requires, I can construct or retrieve a plan (from my back pocket or memory) and reinterpret it, or remind myself of what I intended to do, again a sequence of interactions. I read and follow plans in my activity, not behind the scenes in hidden, subconscious deliberation:

> To put it quite generally, the absurd assumption made by the intellectualist legend is this, that a performance of any sort inherits all its title to intelligence from some anterior internal operation of planning what to do. "Intelligent" cannot be defined in terms of "intellectual" or "knowing how" in terms of "knowing that," "thinking what I am doing" does not connote "both thinking what to do and doing it." My performance has a special procedure or manner, not special antecedents. (Ryle, 1949, p. 31-32)

Using the terms "plan," "symbol," "representation," and "encoding" to simultaneously refer to what a person does over time in sequences of coordinated activity, as well as to internal neurological processes, is confusing at best (witness Vera & Simon's difficulty of applying their meaning of "plan" to Suchman's analysis) and scientifically bankrupt at worst (assuming what needs to be explained).

The symbolic approach reduces comprehending a representation to a matching process because memory is viewed as a body of stored descriptions and programs (Clancey, 1991a): "The memory is an indexed encyclopedia" (Vera & Simon, 1993, p. 10) and "...the storage of such strategies" (p. 17). Vera and Simon clearly identify internal neural structures and processes with the form and use of external representations (e.g., an encyclopedia). Saying "internal representations have all the properties of symbol structures" (p. 40) is clearly wrong. First-person symbol structures, such as sentences on a page, can be rearranged, *reperceived*, and reinterpreted (plus stored and brought out unchanged, handed over to other people directly, overlaid by graphics and other notations). Vera and Simon confuse the carpenter with his tools. Ryle (1949, p. 292 ff.) spoke clearly about this category error, which presumes that the behavior we see outside (inferring, forming hypotheses, posting alternatives, translating between representations, comparing cases, etc.) must also be going on inside. (Of course, it can go on inside in another sense when *the person* is talking to himself.)

The purely symbolic view of comprehending has also obscured why expert systems work. Crucially, a human user of an expert system is not just providing a link to a data base, but *reinterpreting what the queries mean in the current situation*. Providing an interpretation for the rule representations, what they mean, is prone to change and is not recorded in the program. We can store examples (e.g., a pregnant woman is a comprised host), but even these strings must be interpreted (e.g., is a woman who missed a period necessarily pregnant?). The judgment a person provides interpreting MYCIN's prompts for data, as adapted conceptualization, is radically different in kind from the judgment MYCIN provides in syntactically interpreting its rules. The person is recoordinating perceptual and conceptual categorizations; MYCIN is only manipulating a symbolic calculus.

Again, the inadequacy lies in the conflated meaning of "symbol" in the physical symbol-system hypothesis. Vera and Simon (1993) say, "We call patterns symbols when they can designate or denote" (p. 9). This doesn't make a distinction between a stand-in relational mapping (e.g., in sense of a name in a database designating a record on a remote file) and a person reconceptualizing when comprehending text (e.g., when MYCIN-team experts reinterpret what the symbols in a rule mean, with respect to how they fit new patient cases).

Brooks's (1991) insects and Pengi don't use first-person representations; they don't have symbolic expressions like texts or knowledge base rules that they reconceptualize. Vera and Simon's (1993) claim about Brooks's insects that there is "every reason to regard these impulses and signals as symbols" (p. 42) is, of course, a third-person view. These robots use symbols as "stand ins," like the relation between letters and Morse code (Bickhard, 1992).[4] But in contrast with third-person symbols in most other programs, Agre's (1988) and Chapman's (1989) stand-ins represent whole perception–action coordinations. It is no coincidence that Dewey used the same hyphenated notation to show a dialectic, sensorimotor relation. Dewey argued that the "reflex arc" is a circuit, a whole coordination that dialectically relates perception and action, indexically combining value categories with actions (e.g., "seeing-of-a-light-that-means-pain-when-contact-occurs" (Dewey, 1896/1981a, p. 138)).

"A person must carry out a complex encoding of the sensory stimuli that impinge on eye and ear" (Vera & Simon, 1993, p. 41) conflates stand-in perceptual categorization with encoding as *a person's* acts of modeling or sense making, which requires sequences of coordinated perception–action activity (e.g., when a nurse classifies a patient coming into a clinic by checking off symptoms in a chart). Using the term "encoding" for prelinguistic categorizing as well as for human problem solving (e.g., creating symbolic representations on a chart) conflates different kinds of representations, viewing not only what is inside as identical to what is outside, but losing the distinction of what different levels of representationality can do (Bickhard, 1992).

It should now be clear that "SA is not supposed to require a representation of the real-world situation being acted upon" (Vera & Simon, 1993, p. 28) because the agent has no names for prelinguistic neurological organizations; they cannot be strictly mapped onto names and rules. There is no body of behaviors that can be inventoried, no repertoire of mechanical coordinations. Each coordination is always a new composition, built by physical processes of activation, selection, and subsumption, but not ontologically grounded in some set of primitive objects, properties, or relations in the world; indeed, the grounding is the other way around. Regularities develop, but without requiring us to represent them as rules or graphic networks or pictures. The obvious example is of a child learning to speak before being taught an abstract grammar.

The term "third-person representation" is appropriate for describing a designer's relation to structures in a computer program. But are scientists likely to find stable neurological structures that can be given symbolic (linguistic) interpretation? Two hypotheses argue against a clear correspondence between the agent's words and neurological structures: Neurological structures are always new (not literally the same physical structure) and they are always activated as part of an ongoing coordination as circuits. But a

[4] Crucially, stand-ins may be expressed by the scientist in a language, but they do not constitute a language to the robot, in the sense that there is no conventional lexicon, no syntactic ordering conveying roles of terms, and no meaning other than a formal, relational mapping.

reasonable hypothesis is that stable maps involving Broca's and Wernicke's areas, coordinating linguistic activity, may have third-person interpretation. Our third-person definitions of these "symbols" might then be stated in terms of relations among perception–value–action relations (i.e., a mapping of a mapping of types of maps), rather than isolated perceptual or conceptual categories.

4. MODELS OF SITUATIONS AND PROBLEM FRAMING

Situated cognition research calls into question what constitutes a problem for a person (Lave, 1988; Reeke & Edelman, 1988; Schön, 1979). Problems in cryptarithmetic and the Tower of Hanoi, as "situations" given to a "problem solver," are contrasted with the manner in which a person experiences a problematic situation (Dewey, 1938, p. 108; 1938/1981b, p. 518; Wynn, 1991). Perceiving that a problem exists is simultaneously a new way of categorizing, as well as a new way of coordinating behavior. That is, *saying what the problem is* arises with new prelinguistic coordination of *a way of behaving*. Hence, the issue is not of planning versus immediate behavior, per se, but how linguistic behavior dialectically recoordinates sensorimotor systems.

A robot built on Phoenix/Navlab principles may indeed be a useful tool for modeling and hence improve our understanding of multiagent collaboration in uncertain environments. But we need to understand the residue: What must humans do in making sense of and adapting prescribed methods and policies that such models do not replicate? How can we account for development that steps outside of a given theory? In saying that the perceptual strategy model of the Tower of Hanoi problem solver "is clearly a symbolic information-processing system, demonstrating that such a system can carry out SA" (p. 31), Vera and Simon (1993) again equate a model with human capabilities.[5] But this program does not learn. Situated action inherently involves learning new coordinations in the course of every interaction; learning is not just a "second-order effect" (Newell & Simon, 1972; p. 7).

Vera and Simon (1993) state: "Arguing that the fire environment does not represent a legitimate testing ground for SA would be analogous to suggesting that a pilot in a flight simulation is not acting situatedly because the experience is being artifically created" (p. 25). This is the wrong comparison. A flight simulator is not equivalent to a formal model of a flight simulator. The control panels and their physical arrangements are not described, but

[5] Similarly, they say, "Brooks's [1991] robots. . . are good examples of such systems (p. 18)," quickly skipping from human coordinated behavior to models to Brooks's robots. Both the models and the robots are examples of symbolic systems, but not good examples of human adaptive capability.

actual; the simulator can be interacted with directly by a person in an embodied way (by coordinating sensorimotor activity while sitting in a seat, orienting to multiple displays, and pushing buttons). By viewing the firefighter's world *exclusively* in terms of linguistic representations of objects and events, Phoenix research (Cohen, Greenberg, Hart, & Howe, 1989) effectively omits the prelinguistic information and processes by which firefighters determine, for example, that a situation is problematic.[6] In particular, drawing analogies to interpret policy—and knowing that it must be reinterpreted—is grounded in previous sensorimotor experiences (Schön, 1979; Zuboff, 1988).

Chapman's (1989, quoted in Vera & Simon, 1993, p. 36) remark that "a problem. . .can be solved once and for all" is discussed by Lave (1988). The idea is that once we have a problem formulation, a problem description and operators in some ontology, we have a logical puzzle. What remains is purely calculus: manipulating expressions according to the predefined rules, modifying rules by other predefined rules, and so forth. In contrast, problem formulation for people arises in their coordinated activity, as perceptual categorizations, and linguistic creations that are always at some level new (relative to symbolic models of what is happening). That is, problems in part arise because of *discoordination*, an inability to automatically recoordinate perceptual, conceptual, and symbolic categorizations. Studying intelligence by supplying first-person problem formulations (as in the Tower of Hanoi) misses this underlying dialectic process (Schön, 1979), wrongly suggesting that action begins and proceeds only with linguistic descriptions (Reeke & Edelman, 1988; Rosenfield, 1988).

Pengo can "be solved once and for all" (Chapman, 1989, quoted in Vera & Simon, 1993, p. 36) in the sense that we can represent the Pengo world and write a program to play it. Obviously, the game will be played within changing situations and the program will be more or less successful given its available resources. It is true that "a new theoretical paradigm" (p. 38) is not required to play a video game. Vera and Simon miss the point that Pengo is a model, a tool, for helping us to articulate better the nature of coordinated activity.

Vera and Simon believe that learning by EPAM, adaptive production systems, UNDERSTAND, and so on, demonstrates flexibility and creativity equivalent to people: "Symbolic systems. . .continually revise their descriptions of the problem space and the alternatives available to them" (p. 13). But they do this in an ontologically closed space, with symbolic structures grounded only in other symbols. Any system with such a foundation will be limited relative to human creativity, regardless of how much is built in or

[6] If the flight simulator uses simulated images of runways, it is similarly ontologically bound; models and the world can be integrated for varying degrees of freedom.

how much it "learns." For example, AARON (McCormick, 1991) produces an infinite number of different drawings, but is categorically bound (i.e., its performance lies within prescribed classes and their relations, consisting of fewer than a dozen people, many trees, no crocodiles). Modifying the ontological structures after a program fails (Harold Cohen rewrote AARON four times in the past decade) merely begs the issue of how such mechanisms will achieve human competence without their human programmers and critics to tend them.

Saying that people have knowledge that programs lack is circular, suggesting again (in the knowledge-acquisition-as-transfer mistake) that such linguistic, first-person representations preexist before we create them. How do people create the representations they put into the program? By hypothesizing the structures inside and outside the head to be the same in kind, the symbolic approach is forced to reduce meaning to representations of meaning, speaking to what has already been predescribed inside, and hence problem formulation to searching stored or external representations (Tyler, 1978).[7]

A corollary is that the symbolic approach inadequately construes how cognitive models are themselves created, and their relation to the phenomena we sought to understand. For example, in our sense making and inquiry, we classify a physician's behavior in medical diagnosis as an instance of gathering data about the patient (e.g., the question, "When did the fever begin?"). We try to find rules that can generate a sequence of such classified behaviors (e.g., we determine that the physician is following the strategy of asking follow-up questions). But we ignore the tone of voice, the physician's posture, the wave of the hands, the glance over to the left corner of the room. Whenever we suppose that there are *acts* at all, as discrete units, we are adopting the eyeglasses of a functional analysis, which breaks behavior into *phases* of thinking and doing, stimulus and response, gathering information and making conclusions. In merely saying what the act is, we have classified. Abstracting the behavior, we are no longer dealing with neurological coordination processes themselves, but their product (indeed, a mixture of the agent's and the observer's first-person representations: words, plans, rules, and discourse patterns). Dewey told us that this is an old philosophical mistake: "They treated a use, function, and service rendered in conduct of inquiry as if it had ontological reference apart from inquiry" (Dewey, 1896/1981a, p. 320).

Again, we have a recursive problem: We fail to see the inadequacy of our models of problem solving because we judge their veracity in terms of our

[7] Bickhard (1992) asserts that "encodingism" is the asymptotic limit of interactivism (SA). That is, within a stable environment, a symbolic calculator can be supplied with an appropriate starting ontology and incrementally modified to approach human capability.

1st Order (External View)	2nd Order (Internal View)
1) **Social anthropology, sociolinguistics: Social** view of everyday practice.	4) **Neuropsychology:** More detailed understanding of cognitive processes in the brain.
Distinction between practice and theory; culture is not reducible to a list of beliefs, conventions, or general rules.	*Distinction between deliberation and coordination.*
Shift from viewing all behavior as generated from internally pre-stored facts and rules to inquiring about theories people actually do create and use. What is the role of a theory? What is the nature of skills, habits, & improvisation? Contrast with calculus: representation manipulation by prestated rules and procedures (e.g., subtraction, expert system inference, syntactic machine learning). ↓	How is complex, multi-modal sensorimotor coordination possible without symbolic instruction or reflection that describes patterns and reasons about alternative actions? ↑
2) **Situated cognition:** More detailed understanding of how individuals learn, how representations are created and used.	3) **Philosophy:** Foundational analyses of concepts, meaning, and models.
Distinction between representations we can perceive, representing to ourselves, and neural structures.	*Distinction between models and knowledge.*
What is the role of plans, models, maps? Why need a map and what do with it in your behavior (i.e., how actually use it)?	How do we act without having a theory (i.e., without manipulating representations and plans)? How do we "find our way about" (Wittgenstein, 1958)? How is familiarity embodied? (Dreyfus and Dreyfus, 1986)

Table 1. Line of Argument Relating Situated Cognition to Neuropsychology

models of problem *formulation*. To break out, to form a scientific theory of cognition that would enable us to build an intelligent machine, we must move to the social and neurological levels. This requires explaining how representation systems evolved in living systems and develop in each child from a nonsymbolic foundation (Edelman, 1992).

5. SUMMARY OF THE ARGUMENT: RELATING SOCIAL AND COGNITIVE VIEWS

Table 1 collects independent lines of inquiry in social anthropology, sociolinguistics, education, developmental psychology, philosophy, and neuropsychology into a single line of argument. The complicated move, involving many sources of information and lines of reasoning relates the second and third steps.[8] Relating social theories to cognitive theories is difficult because

[8] I am indebted to John Haugeland for summarizing the argument in terms of a first-order (social) and second-order (cognitive) lines of reasoning (personal communications, McDonnell Foundation Sante Fe Conference on "The Science of Cognition," June 1992).

researchers may be uncomfortable with philosophical analyses. The key claim is that psychological models in terms of beliefs, goals, schemas, inference, strategies, and so forth, describe and explain patterns of behavior of an-agent-in-an-environment, *the product* of neural and social interactions: not processes occurring subconsciously inside the brain, but *the agent's behavior in the world.*

The social view accepts the representational focus of cognitive science and AI. But applying the ethnomethodological stance (Berger & Luckman, 1967), rather than imputing representations as being hidden and manipulated in the head of agents, the social view looks to see how people create, modify, and use representations (e.g., slides for a talk, architectural drawings, texts, expert systems) in their visible behavior.

> We introduce no "faculties" or other operators (however disguised) of an organism's behaviors, but require for all investigation direct observation and usable reports of events.... In especial, we recognize no names that pretend to be expressions of "inner" thoughts, any more than we recognize names that pretend to be compulsions upon us by "outer" objects. (Dewey & Bentley, 1949, p. 120)

Thus, we have a fresh look at where representations come from, how they are learned, and what they do for us (Bamberger, 1991; Schön, 1987; Suchman, 1987; Wenger, 1990).

Unfortunately, this difficult attempt to bridge social and cognitive views has caused an number of misunderstandings. For example, Vera and Simon (1993) believe that the cottage cheese anecdote is meant to illustrate that "knowledge about interaction with real-world objects is not symbolically represented" (p. 23). Instead, the example demonstrates that inquiry is a complex coupling of physical materials, sensorimotor coordinations (including prelinguistic ways of seeing), and first-person representation and manipulation of constraints. Pedagogically, the cottage cheese story celebrates inventiveness and the importance of teaching representational conventions without destroying creativity, without leading kids to believe there is only one way to think, only one "correct" way of describing facts and working problems (Brown, Collins, & Duguid, 1988; Clancey, 1992b, in press).

The repeating theme is the relation of formal models to skills (know-how, unarticulated coordinations). Vera and Simon (1993) suggest that it doesn't matter what we call this problem, saying that Lave has merely "reformulated" it in different terms. Ironically, Lave (1988), Schön (1979), and others are calling attention to just this, *the process of problem formulation*, the effect of initial metaphors on the perceived problem space, how, in fact, naming and problem framing occurs. That is, the question at hand is "how shall we frame the problem of problem framing?" Labeling the learning problem as "transfer" ignores how framing a problem constrains a space of solutions (e.g., like calling people "homeless"; Schon, 1979). Saying that transfer is "frequently addressed in the learning literature" (p. 23) does not respond to Lave's concern that cognitive theories of learning in the literature are inadequate. The "capture and transfer" view of learning[9] suggests that all knowledge can be written down in symbolic models; that what people know can be inventoried (as facts and generative rules); that a symbolic program could be equivalent to a human being's intelligence. Lave rejects these assumptions. She grapples with the issue of what constitutes cultural knowledge and how it is reproduced and adapted without exhaustive first-person representation of its content.

Vera and Simon's (1993) phrase "so long as a system can be provided with a knowledge base" (p. 42) suggests that social knowledge can be inventoried, that we share representations of how to act (e.g., descriptions of conventions, grammars, scripts, belief networks). On the contrary, SA argues that our ability to coordinate our activity by gesturing, orienting our bodies, mimicking sounds and postures, and so on, interactively, in our behavior, couples our perception and action, providing the foundation for speaking a common language, constructing shared causal models, deliberating and planning complex activities (Lave, 1988). Again, first-person representations in our statements and formal models, exemplified by knowledge bases of expert systems, are not the substrate of intelligent behavior. Such social constructions are only possible because we have similar prelinguistic experiences, including similar coordinations of values, perceptual categories, and behaviors (Edelman, 1992). Lave's phrase, "transformational relations which are part of 'intentionless but knowledgeable inventions' " (quoted in Vera & Simon, 1993, p. 22), refers to the underlying prelinguistic, adaptive coordinating, and categorizing, what we commonly call "know-how." And indeed, at a higher level, the relation between increasing neural and social organizations is dialectic. In a group we are mutually constraining each other's perception and sequencing, so our capabilities to interact are both developed within, and manifestations of, social, multiagent interactions. Because the regularities in these interactions (e.g., turn taking) are not based on representations of the patterns themselves, the system of multiple situated agents constitutes a higher-order, self-organizing system:

> A group is maintained by its activity...the social group possesses a certain trend of development.... This trend need not be, and in the majority of cases it certainly is not, present in the mind, or fully represented in the behaviour, of any individual member of the group. (Bartlett, 1932/1977, p. 275).

> If we desire to explain or understand the mental aspect of any biological event, we must take into account the system—that is, the network of closed circuits within which that biological event is determined. But when we seek to explain

[9] Exemplified by my dissertation, *The transfer of rule-based expertise in a tutorial dialogue*. (Clancey, 1979).

> the behavior of a man or any other organism, this "system" will usually not have the same limits as the "self"—as this term is commonly (and variously) understood. . . mind is immanent in the larger system, man plus environment. (Bateson, 1972, p. 307)

Table 2 summarizes most of the distinctions I have drawn here.

6. CONCLUSION: DELIBERATING RECONCEIVED

The symbolic view suggests that human reasoning rescues us from irrational, immediate, emotional behavior. It posits intermediate deliberation *between* peripheral perception and motor systems. The idea has been to replace raw, reactive, reflexive behavior by deliberated behavior, the mark of reason. That is, we think before we act. But rather than viewing thinking, inquiry, as occurring in our behavior—as part of coordinated looking, manipulating of materials, and representing new possibilities—inquiry is asserted to be between perceiving and acting, so thinking goes on subconsciously, mediating every behavior. By this maneuver, we sought to distinguish human problem solving from uncontrolled, animal processes (Ryle, 1949).[10]

Vera and Simon's (1993) view of SA is that "planning and representation, central to symbolic theories, are claimed to be irrelevant in everyday human activity" (p. 7). Quite the contrary, Suchman is explaining how plans are used in everyday activity. Vera and Simon state that "Suchman takes the rather extreme position that plans play a role before and after action, but only minimally during it" (p. 16). This reveals their view that deliberation, referring to stored internal plans, is required for every action. Newell's (1990, p. 135) use of the term "elementary deliberation" for automatic behavior shows the same bias. When Suchman wrote, "Plans as such neither determine the course of situated action nor adequately reconstruct it" (quoted in Vera & Simon, 1993, p. 16), she meant that interpretation of plans (reading and comprehending a plan or creating a new one) goes on in our behavior as coordinated sequences of perceiving and acting, not by internally linking peripheral perceptual and motor systems.

By such a neuropsychological interpretation, we find that SA is not making obviously wrong statements, but rather profound, and at first difficult to understand, claims about cognition, essentially turning symbolic models inside out: Every act of deliberation occurs as an immediate behavior. That is, every act of speaking, every motion of the pen, each gesture, turn of head, or any idea at all is produced by the cognitive architecture as a matter

---

[10] Ryle was not a behaviorist, but he did believe that the behaviorist, cowering in the shadow of a fortress, is better off than the "intellectualist" hiding in a castle with broken walls and no moat (Ryle, 1949, p. 330). Ryle's genius was realizing that there is an alternative to both traditions.

TABLE 2
Theoretical Differences Between Symbolic Approach and Situated Cognition

	Symbolic Approach	Situated Cognition
memory	stored rules or schema structures in a representation language	neural nets reactivated and recomposed in-line via selection; not a place or body of descriptions of how the world or agent's behavior appears.
representation	meaningful forms internally manipulated subconsciously	created and interpreted in our activity (first person); external representations ≠ representing to self ≠ neural structures
internal processes	modularly independent; can perceive and reason without acting	codetermined, dialectic; always adapted (generalized from past coordinations), inherently chronological
immediate behavior	selected from prepared possibilities ("preexisting actions")	adapted, composed, coordinated, always new; always a sensorimotor circuit
reasoning	supplants immediate behavior; goes on subconsciously	occurs in sequences of behavior over time
speaking	meaning of the utterance is represented before speaking occurs	speaking and conceiving occur dialectically; representing meaning occurs as later commentary behavior
learning	secondary effect (chunking)	primary learning is always occurring with every thought, perception, and action; chunking occurs as categorization of sequences; secondary (reflective) learning occurs in sequences of behavior over time; requires perception.
knowledge representation (cognitive model)	corresponds to physical structures stored in human's brain	a model of some system in the world and operators for manipulating the model; abstracts agent's behavior, explaining interaction in some environment over time.
concepts	labeled structures, corresponding to linguistic terms, with associated descriptions of properties and relations to other concepts, i.e., meanings are symbolically represented and stored	prelinguistic categorizations of perceptual categorizations; ways of coordinating perception and action; has no inherent formal structure; cannot be inventoried; meaning and perception are inseparable.
analogy	feature mapping of concept representations	process of perceiving and acting by recomposing previous coordinations (e.g., "seeing as").

of course, as a new neurological coordination. The contrast is not between immediate behavior and deliberate behavior (cf. Newell, 1990, p. 136). The contrast is between (1) dialectic coupling of sensorimotor systems in an ongoing sequence of coordinations, and (2) *perceived sequences of behaviors,* which we name, classify, and rationalize.

For example, the diagnostic strategies of NEOMYCIN (Clancey, 1988) abstractly describe the sequence of physician requests for patient data in terms of goals, shifting attention, and the developing form of the model of the patient. If a physician refers to such a plan, it is in his or her activity, not in a hidden way, but as a person generating ideas and reflecting on them (Ryle, 1949). NEOMYCIN is inadequate as a cognitive model in the sense that it doesn't distinguish between first- and third-person representations, between the gearlike mapping of relational models and the adaptive interpretation involved in using NEOMYCIN's rules in problematic situations.

Deliberating, what Dewey (1938) called "inquiry," occurs in the course of our behavior, in the relation of our shifts of attention, representational acts, and reshaping of materials around us: not between perceiving and acting, but through it, over time, in cycles of reperceiving, reshaping, and reenacting what we have said and done before. This *remaking* involves both internal recomposing and recoordinating of perception-action sequences, as well as physical manipulation of materials in our environment (e.g., rewriting a paragraph). In this sense, Bamberber and Schön (1983), following Dewey, view sense making as integral with a process of *making things,* which puts primacy on observing what people do. Compare the sterile representation of disks and pegs in the Tower of Hanoi program to a journal article, an architectural drawing, or blocks arranged to describe a melody (Bamberger, 1991). How do people create such representations? How do interpretations and uses change within changing activities and values (Schön, 1987)?

The physical symbol-system hypothesis, which equates symbols inside and out, is bolstered by the idea that comprehension of text or interpretation of a plan or graphic involves only reconfiguring representations (e.g., dictionary definitions, grammars, scripts). Similarly, speaking is viewed as a process of prerepresenting inside what you intend to say, as well as its grammatical form and meaning. The implied claim of SA is that when today's computer programs read and "comprehend" text, they do so only in a mechanical, rotelike way, bound by symbolic structures and calculi.

Vera and Simon (1993) say that "the symbolic approach has never disagreed...that there is more to understanding behavior than describing internally generated symbolic goal-directed planning" (p. 23). Vera and Simon may have had a broader view, but this is a poor summary of the assumptions driving student modeling in intelligent tutoring systems, knowledge acquisition in expert systems, or the CYC project (Smith, 1991). As I have shown, SA is critically at odds with Vera and Simon's beliefs: Human knowledge is not equivalent to a body of knowledge representations. Symbolizing and comprehending are processes that occur in human behavior, not as internal linkages between perception and action. Learning is a primary phenomenon.

To make progress, to evaluate and refine the emerging neuropsychological model of SA (Figure 2), we should consider the broader analyses of the social sciences and philosophy. Sacks (1987), for example, showed how an ethnographic approach provides a broader view of how cognition develops within and is manifest in everyday life. As a heuristic for research, we must reexamine structures and processes in our cognitive models, and clearly state which correspond to internal third-person representations; which to representing to oneself (without creating structures that are perceived); and which to first-person representations in the world (which are reshaped and perceived in cycles of activity). SA research strongly suggests that philosophical arguments, such as Ryle's (1949) and Wittgenstein's (1958) critique of rule-based models of cognition, are understandable and provide relevant heuristic guidance (H. Dreyfus & S. Dreyfus, 1986; Edelman, 1992; Lakoff, 1987; Slezak, 1992; Tyler, 1978).

SA doesn't require "a whole new language," (Vera and Simon, 1993, p. 46) but it does require that we watch how we use our words, particularly, "memory," "knowledge," "information," "symbol," "representation," and "plan." SA *does* suggest a different research agenda, as is amply demonstrated by the work of all the people cited by Vera and Simon, who either are studying human behavior in a new way (Lave, 1988; Suchman, 1987), using models in a new way in the classroom (Bamberger, 1991; Schön, 1987), inventing new robot designs (Maes, 1990), using computers in new ways in business (Kukla, Clemens, Morse, & Cash, in press), and developing new models of the brain (Edelman, 1987, 1992; Freeman, 1991; Iran-Nejad, 1987; Rosenfield, 1988; Sacks, 1987). Certainly, it isn't necessary (or perhaps possible) to break "completely from traditional...theories" (p. 46) but instead to reconsider the relation of our models to the cognitive phenomena we sought to understand. Symbolic models, as tools, will always be with us. Yet, already the shift has begun from viewing them as intelligent beings, to but the shadow of what we must explain.

REFERENCES

Agre, P. (1988). *The dynamic structure of everyday life.* Unpublished doctoral dissertation, MIT, Cambridge, MA.

Alexander, C. (1979). *A timeless way of building.* New York: Oxford University Press.

Arbib, M.A. (1981). Visuomotor coordination: From neural nets to schema theory. *Cognition and Brain Theory, 4,* 23–40.

Bamberger, J. (1991). *The mind behind the musical ear.* Cambridge, MA: Harvard University Press.

Bamberger, J., & Schön, D.A. (1983, March). Learning as reflective conversation with materials: Notes from work in progress. *Art Education*, pp. 68-73.

Bartlett, F.C. (1977). *Remembering—a study in experimental and social psychology*. Cambridge: Cambridge University Press. (Original work published 1932).

Bateson, G. (1972). *Steps to an ecology of mind*. New York: Ballentine Books.

Berger, P.L., & Luckmann, T. (1967). *The social construction of reality: A treatise in the sociology of knowledge*. Garden City, NY: Anchor Books.

Bickhard, M.H. (1992). Levels of representationality. Paper presented at the McDonnell Foundation Conference on The Science of Cognition, Sante Fe, NM.

Bickhard, M.H., & Richie, D.M. (1983, June). *On the nature of representation: A case study of James Gibson's theory of perception*. New York: Praeger.

Braitenberg, V. (1984). *Vehicles: Experiments in synthetic psychology*. Cambridge: MIT Press.

Brooks, R.A. (1991). How to build complete creatures rather than isolated cognitive simulators. In K. vanLehn (Ed.), *Architectures for Intelligence: The 22nd Carnegie Symposium on Cognition*. Hillsdale, NJ: Erlbaum.

Brown, J.S., Collins, A., & Duguid, P. (1988). Situated cognition and the culture of learning (IRL Report NO. 88-0008, Institute for Research on Learning, Palo Alto, CA). (Shorter version appears in *Educational Researcher, 18*(1), February, 1989)

Chapman, D. (1989). Penguins can make cake. *AI Magazine, 10*, 45-50.

Clancey, W.J. (1979). *Transfer of rule-based expertise through a tutorial dialogue*. Doctoral dissertation, Department of Computer Science, Stanford University, CA (Appeared in revised form as *Knowledge-Based Tutoring: The GUIDON Program*, Cambridge, MA: MIT Press, 1987).

Clancey, W.J. (1987). Review of Winograd and Flores' *Understanding computers and cognition:* A favorable interpretation. *Artificial Intelligence, 31*, 232-250.

Clancey, W.J. (1988). Acquiring, representing, and evaluating a competence model of diagnosis. In M.T.H. Chi, R. Glaser, & M.J. Farr (Eds.), *The nature of expertise*. Hillsdale, NJ: Erlbaum.

Clancey, W.J. (1989). The knowledge level reconsidered: Modeling how systems interact. *Machine Learning, 4*, 285-292.

Clancey, W.J. (1991a). Review of Rosenfield's *The invention of memory. The Journal of Artificial Intelligence, 50*, 241-284.

Clancey, W.J. (1991b). The frame of reference problem in the design of intelligent machines. In K. vanLehn (Ed.), *Architectures for Intelligence: The 22nd Carnegie Symposium on Cognition*. Hillsdale, NJ: Erlbaum.

Clancey, W.J. (1992a). Model construction operators. *Artificial Intelligence, 53*, 1-115.

Clancey, W.J. (1992b). Representations of knowing: In defense of cognitive apprenticeship. *Journal of AI and Education, 3*, 139-168.

Clancey, W.J. (in press). *Guidon-Manage revisited: A socio-technical systems approach. Journal of AI and Education.*

Cohen, P.R., Greenberg, M.L., Hart, D.M., & Howe, A.E. (1989). Trial by fire: Understanding the design requirements for agents in complex environments. *The AI Magazine, 10*(3), 34-48.

Collingwood, R.G. (1938). *The principles of art*. London: Oxford University Press.

Dewey, J. (1938). *Logic: The theory of inquiry*. New York: Henry Holt & Co.

Dewey, J. (1981a). The reflex arc concept in psychology. In J.J. McDermott (Ed.), *The philosophy of John Dewey*. Chicago: University of Chicago Press. (Original work published 1896).

Dewey, J. (1981b). The criteria of experience. In J.J. McDermott (Ed.), *The philosophy of John Dewey*. Chicago: University of Chicago Press. (Original work published 1938).

Dewey, J., & Bentley, A.F. (1949). *Knowing and the known*. Boston: Beacon Press.

Dreyfus, H.L., & Dreyfus, S.E. (1986). *Mind over machine.* New York: Free Press.

Edelman, G.M., (1987). *Neural Darwinism: The theory of neuronal group selection.* New York: Basic Books.

Edelman, G.M. (1992). *Bright air, brilliant fire: On the matter of the mind.* New York: Basic Books.

Farah, M.J. (1992, June). Cognitive neuropsychology without the "transparency assumption." Paper presented at the McDonnell Foundation on The Science of Cognition, Sante Fe, NM.

Freeman, W.J. (1991, February). The physiology of perception. *Scientific American,* pp. 78–85.

Gleick, J. (1987). *Chaos: Making a new science.* New York: Viking.

Greenbaum, J., & Kyng, M. (1991). *Design at work: Cooperative design of computer systems.* Hillsdale, NJ: Erlbaum.

Gregory, B. (1988). *Inventing reality: Physics as language.* New York: Wiley.

Iran-Nejad, A. (1984). Affect: A functional perspective. *Mind and Behavior, 5,* 279–310.

Iran-Nejad, A. (1987). The schema: A long-term memory structure or a transient functional pattern. In R.J. Tierney, P.L. Anders, & J.N. Mitchell (Eds.), *Understanding readers' understanding: Theory and practice.* Hillsdale, NJ: Erlbaum.

Kukla, C.D., Clemens, E.A., Morse, R.S., & Cash, D. (in press). An approach to designing effective manufacturing systems. In *Technology and the future of work.*

Laird, J.E., & Rosenbloom, P.S. (1990). Integrating execution, planning, and learning in SOAR for external environments. *Proceedings of the Eighth National Conference on Artificial Intelligence.* Menlo Park, CA: AAAI Press.

Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind.* Chicago: University of Chicago Press.

Lave, J. (1988). *Cognition in practice: Mind, mathematics, and culture in everyday life.* Cambridge: Cambridge University Press.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation.* Cambridge: Cambridge University Press.

Lewis, R., Huffman, S.B., John, B.E., Laird, J.E., Lehman, J.F., Newell, A., Rosenbloom, P.S., Simon, T., & Tessler, S. G. (1990). Soar as a unified theory of cognition. Spring 1990 Symposium. *Proceedings of the 12th Annual Conference of the Cognitive Science Society.* Hillsdale, NJ: Erlbaum.

Maes, P. (Guest Ed.). (1990). Designing autonomous agents. *Robotics and Autonomous Systems, 6*(1,2).

McCormick, P. (1991). *AARON's code: Meta-art, artificial intelligence, and the work of Harold Cohen.* New York: W.H. Freeman.

Newell, A. (1980). Physical symbol systems. *Cognitive Science, 4,* 135–183.

Newell, A. (1990). *Unified theories of cognition.* Cambridge, MA: Harvard University Press.

Newell, A., & Simon, H.A. (1972). *Human problem solving.* Englewood Cliffs, NJ: Prentice-Hall.

Nonaka, I. (1991). The knowledge-creating company. *Harvard Business Review,* 96–104.

Piaget, J. (1981). *The psychology of intelligence.* Totowa, NJ: Littlefield, Adams, & Co. (Original work published 1960).

Reeke, G.N., & Edelman, G.M. (1988). Real brains and artificial intelligence. *Daedalus, 117* (1), 143–173.

Roschelle, J., & Clancey, W.J. (in press). Learning as social and neural. *Educational Psychologist.*

Rosenfield, I. (1988). *The invention of memory: A new view of the brain.* New York: Basic Books.

Rosenschein, S.J. (1985). *Formal theories of knowledge in AI and robotics* (SRI Tech. Note 362, Palo Alto, CA).

Ryle, G. (1949). *The concept of mind.* New York: Barnes & Noble.

Sacks, O. (1987). *The man who mistook his wife for a hat.* New York: Harper & Row.

Schön, D.A. (1979). Generative metaphor: A perspective on problem-setting in social policy. In A. Ortony (Ed.), *Metaphor and thought.* Cambridge: Cambridge University Press.

Schön, D.A. (1987). *Educating the reflective practitioner.* San Francisco: Jossey-Bass.

Schön, D.A. (1990). *The theory of inquiry: Dewey's legacy to education.* Paper presented at the Annual Meeting of the American Educational Association, Boston.

Slezak, P. (1992, June). *Situated cognition: Minds in machines or friendly photocopiers?* Paper presented at the McDonnell Foundation Conference on The Science of Cognition, Sante Fe, NM.

Smith, B.C. (1991). The owl and the electric encyclopedia. *Artificial Intelligence, 47,* 251–288.

Smoliar, S.W. (1989). Review of *Neural Darwinism: The theory of neuronal group selection* [by G.M. Edelman]. *Artificial Intelligence, 39,* 121–136.

Suchman, L.A. (1987). *Plans and situated actions: The problem of human-machine communication.* Cambridge, Cambridge University Press.

Tyler, S. (1978). *The said and the unsaid: Mind, meaning, and culture.* New York: Academic.

Vera, A.H., & Simon, H.A. (1993). Situated action: A symbolic interpretation. *Cognitive Science, 17,* 7–48.

Vygotsky, L. (1986). *Thought and language* (A. Kozulin, Ed.). Cambridge, MA: MIT Press. (Original work published 1934)

Wenger, E. (1990). *Toward a theory of cultural transparency: Elements of a social discourse of the visible and the invisible.* Unpublished doctoral dissertation, Department of Information and Computer Science, University of California, Irvine.

Winograd, T., & Flores, F. (1986). *Understanding computers and cognition: A new foundation for design.* Norwood, NJ: Ablex.

Wittgenstein, L. (1958). *Philosophical investigations.* New York: Macmillan.

Wynn, E. (1991). Taking practice seriously. In J. Greenbaum & M. Kyng (Eds.), *Design at work: Cooperative design of computer systems.* Hillsdale, NJ: Erlbaum.

Zuboff, S. (1988). *In the age of the smart machine: The future of work and power.* New York: Basic Books.