Embodied Cognitive Science: Concepts, Methods and Implications for Psychology

Wolfgang Tschacher, Christian Scheier

appeared in: Matthies M, Malchow H und Kriz J (eds.) (2001), Integrative Systems Approaches to Natural and Social Dynamics (pp 551-567). Berlin: Springer.

Since the "cognitive shift" of psychology, a close association between psychology and the advances in computer technology and artificial intelligence research has evolved. According to the 'computational' symbol processing approach, cognition consists of a series of sequentially ordered processing stages. Between perception and action, input is processed by higher cognitive functions, such as categorization, memory, and planning. These cognitive functions are conceived as independent modules lacking a direct interface with the environment. This approach is criticized due to its inherent fundamental problems. Alternative research programs, such as embodied cognitive science, primarily address the issues of embodied cognition, i.e., cognition is viewed as originating from the interaction of body and environment. The methods of the corresponding "new AI" encompass robotics and the use of autonomous agents. It is investigated here which implications for psychology may arise. A theoretical conceptualization of autonomous agents based on dynamical systems theory and synergetics is outlined. Within this context, the cognitive system is conceived as a complex system comprising numerous sensorimotor loops; coherent and adaptive perception-action processes emerge from the influence of affordances. Examples cited from the field of applied psychology indicate that these perspectives lead to the formulation of new research questions and reinterpretation of empirical findings.

Keywords. Affordance, Artificial intelligence, Cognition, Dynamical system, Embodied cognitive science, Self-organization

1 Introduction: Problems of 'Classical' Artificial Intelligence Research

Since the advent of the cognitive shift in psychology four decades ago (Miller et al. 1960), the symbolic information processing approach has dominated cognition research and other essential subdisciplines of psychology. Formulated as an alternative to behaviorism, this approach addressed the study of higher cognitive processes such as thinking, reasoning, planning, and memory. These cognitive functions were viewed as relatively independent modules lacking a direct interface with the person's environment. Analogous to the (von Neumann) computer, these modules process symbols the meaning of which are defined in relation to other symbols. The human symbol system is believed to exist independently of any biological substrate, and may by analogy be compared to software which itself is characterized independently of the hardware on which it is implemented (Newell 1980). The physical realization of symbols (i.e., their *embodiment*) is irrelevant as long as their syntactical relation to other symbols remains invariant. Since the computational or cognitive function of symbols is viewed independently of their physical realization, this symbolic information processing approach is occasionally termed computationalism or cognitivism.

The difficulties arising from this perspective on intelligence have become increasingly manifest in the course of the past two decades (Dreyfus 1972; Kolers and Smythe 1984; Winograd and Flores 1986). In this chapter, before moving on to discuss possible implications for cognitive psychology, we will first elucidate the problems of 'classical' symbolic information processing in the research on artificial intelligence (AI), for it is in this domain that these problems were most clearly evident.

A typical finding in AI has been that tasks which can be easily performed by humans – such as perceiving, walking, and playing football – are particularly difficult for computers. Inversely, tasks that are difficult for humans – such as logical reasoning or playing chess – are relatively easy for computers. It is striking that the majority of problems that are difficult to solve for a computer are associated with the interface to the real world[1]. The problems of classical AI became immediately evident when the systems were connected to the environment by means of devices such as cameras or grippers, in other words when information processing systems were supplied with "bodies". The fundamental reason for these problems lay in neglecting the interaction of cognition, body, and the world. The bulk of work in classical AI was related to abstract virtual worlds with clearly definable states and operations.

As a result of these shortcomings, several authors conceptualized these problems (Franklin 1996; Hendriks-Jansen 1996; Pfeifer and Scheier 1999). We will now outline three topics that are treated in these and other recent publications.

With regard to the first topic, Harnad (1990) isolated the so-called *symbol-grounding problem* as especially virulent in this respect. Symbol grounding addresses the association of a real object to the symbol representing this object, which can be stored inside a system. Harnad asked, "How can the meanings of the meaningless symbol tokens, manipulated solely on the basis of their (arbitrary) shapes, be grounded in anything but other meaningless symbols?" (Harnad 1990, p 335). In classical AI – and in computationalism in general – the meaning of symbols is defined purely syntactically, i.e., by the manner in which symbols relate to other symbols and how they are processed by an interpreter (Newell 1990). The relation between symbols and the real world or between symbols and a human observer is rarely explicated. Symbolic systems (such as chess computers) usually operate in closed virtual worlds. This need not pose a problem in information technology (e.g., in data base applications or expert systems) provided a human observer is there to interpret the symbols and thus guarantee the link between the symbol and the outside world. With the exception of real time applications, this link is seldom addressed in computer science;

[1] The "real world" is to be understood as the antonym to a virtual world (one that is realized inside a computer) of a symbol-processing program.

the observer is presumed to "know" what the symbols represent. Hence, the meaning of the symbols in the real world is grounded in the observer's experience and his or her interaction with the world. Symbols have meaning for the observer (and for the system designer), but not for the system itself. What happens in the absence of an interpreting observer, as is the case, for instance, in computer vision tasks? The system must then itself generate a link between the symbols and the external world, which in turn leads to symbol grounding problems. In point of fact, the machine recognition of real objects is a highly non-trivial problem which thus far has been only inadequately solved (e.g., Ullman 1996). Symbol grounding can be effected solely by the interaction of a system with the real world, which necessitates the system having a body. The symbol-grounding problem precludes a solution being found within the scope of computationalism alone (Bickhard and Terveen 1995; Barsalou 2001).

Concerning the second topic, the *frame of reference problem* (Clancey 1991) focuses on the system-environment link in a more general way by pointing out the importance of selecting the frame of reference used for describing and explaining behavior. From which perspective do we observe a system? A system (and its behavior) may be described by an external observer (third-person perspective) or may address the mechanism underlying the behavior (first-person perspective, i.e., the system's perspective) (cf. Atmanspacher and Dalenoort 1994). Description and mechanism must be clearly distinguished. A review of the literature, however, shows that this rule is not followed in many cases, notably in AI and cognitive psychology. One is tempted to describe behavior as goal-oriented even if no explicit goals have had an effect on this behavior. Quite simple cybernetic mechanisms implemented in Braitenberg vehicles (see below) can generate behavior that even trained observers tend to describe, or even erroneously 'explain', by using complex constructs such as goal, intention, or plan. Nor can behavior be reduced to internal mechanisms alone. Behavior is "situated" in that it develops from the interaction between the system and the environment (Greeno 1989; see the discussion of 'situated action' by Vera and Simon (1993) and the corresponding debate in the journal *Cognitive Science*). Categorization, for example, must not be restricted to a mapping of a stimulus to an internal representation, but must also encompass the agent[2] and its interaction with an object. If behavior emerges from the interaction between system and environment, an agent may generate and use categories that are not explicitly represented within the agent. Complex behavior does not necessarily depend on complex internal mechanisms since a greater portion of complexity of behavior results from the interaction of the system with its environment.

A third topic deals with stability. Symbol-based representation and information processing in AI applications are confronted with the *frame problem* (Dennett 1984; Pylyshyn 1987). Given the complex environment of a cognitive

[2] "Agent" was defined by Minsky (1985) as an isolable cognitive skill or process such as "put block A on block B". Complex behavior is built from the interplay of many simple agents in a "society of agents", Minsky's concept of the cognitive system. More generally, the term agent is also used to mean an animal or machine that can operate autonomously in an environment.

system, a symbolic representation of this environment must by necessity also be complex. What appears to be a simple quantitative problem of memory capacity has quite virulent consequences as soon as the environment is changed or the position of the system is altered. Each change, movement, or dynamics enforces real-time changes of the representation within the system. This rapidly results in a combinatorial explosion. The explosion of computational demands on the system is caused by the pressure to assess all (or almost all) of the implications for the world model of any change experienced "outside". The system cannot know a priori which of these implications are irrelevant and which are necessary for survival. In the former condition of an irrelevant change, the cognitive system must be kept stable (assimilation), whereas in the latter condition flexible adaptation to an altered world is needed (accommodation). In symbol systems (usually based on propositions), this stability-flexibility tradeoff cannot be achieved because of the exponentially growing computational load. However, the quick and seemingly effortless solving of such dilemmas in real time characterizes cognizing and learning animals – the rat in the maze *does not sit lost in thought*, while updating its knowledge base and wondering which turn to take (a cogent argument in the annals of psychology which was used to dispute Tolman's concept of a cognitive map).

This concludes our brief outline of the problems encountered in AI. We have not dealt with the unrealistic expectations that were generated between the 1960s and 1980s regarding the imminent evolution of human-like machine intelligence. Criticism voiced in the 1990s is in fact much more fundamental; it lead to the suggestion to limit first an approach to models of insect-level intelligence (Brooks 1991). Moreover, the problems we have touched upon are not only confined to symbolic systems (in contrast to subsymbolic, connectionist architectures). Connectionist systems (neural nets) may provide a partial solution to the stability problem (Strube 1990; Caspar et al. 1992). It should be noted, however, that connectionist systems are generally examined as cognitive modules (e.g., associative memory) that have no direct interface with the environment. The limitation of connectionist systems is, therefore, identical to those found in other architectures, in that the input of the system is (pre) processed by the designer and its output is not fed back to the system but is interpreted by the designer. In no application to date has a connectionist system (nor a classical rule-based system) been able to stand on its own. A mediating observer is indispensable in each case. If such systems are to be used as models for cognitive-psychological research, the aspect of an 'observer in the loop' must always be kept in mind, otherwise one risks proffering pseudo explanations by inserting a homunculus[3].

2
Autonomous Agents: A New Approach in Cognitive Science

The aforementioned critical issues that are inherent to the classical AI approach concern the interplay between mechanisms, their embodiment in the agent, and

[3] Interestingly, the problem of a homunculus in psychological explanation was extensively discussed in Gestalt psychology (Köhler 1947).

the ensuing interaction of the agent with the environment. The finding that the body is pivotal to an understanding of intelligence and cognition has led to a new discipline which transcends AI, namely *embodied cognitive science* (Pfeifer and Scheier 1999; Varela et al. 1991). An important method of this discipline is the examination of cognitive processes using autonomous robots. As the following demonstrates, this synthetic approach (based on material modeling) is reinforced by a series of recent empirical findings, which illustrate the significance of the body for intelligent behavior.

Accordingly, a fundamental prerequisite of the so-called "New AI" ("Behavior-Based AI", "Nouvelle AI") is to provide the system with an interface to the real world. This interface is represented by a body which is sensorially and behaviorally embedded in the real world, thereby permitting the system to build up autonomously a knowledge base without the aid of an interpreting observer. In this way the system is made 'complete'. The synthetic methodology of AI has been enhanced by this new approach, in that it now not only incorporates computer simulation – as is the case in symbolic AI and in the subsymbolic connectionist paradigm – but also physical systems that exist in real environments. These systems are *autonomous agents*.

However, the empirical paradigm of autonomous agents is not limited to specific modeling methodology and the construction of mobile robots, but additionally opens up theoretical perspectives on cognition. Many of the central ideas were formulated as early as 1961 by Masanao Toda, a professor of psychology in Harvard (Toda 1962). Toda's basic idea (as an alternative to experimental psychology) was to investigate complete – albeit simple – systems instead of focussing on isolated modules of cognition on the basis of a limited number of tasks. Among other attributes, these systems must possess the abilities to perceive, categorize, learn, navigate, memorize, and also be capable of free choice of action. Toda argued that the integration of these competencies into a system would furnish new insight into intelligence and cognition. Toda's 'Solitary Fungus Eater' is an autonomous agent that is sent on a mission to search for uranium on a distant planet. The agent's reward is in proportion to the amount of uranium it collects. It feeds on a specific fungus that grows on the planet and possesses the sensorimotor competence that enables it to gather uranium and deposit this at an installation. Obviously, the agent is *autonomous* because the planet is too far away to permit remote control. The agent must also be *situated*, in other words perceive the world from its own perspective because information is only accessible via its own sensors. In addition, the agent must possess a *body* in order to collect uranium and also be *adaptive* because the planet's landscape is only partially explored and it must be able to differentiate among uranium, fungus and obstacles. These concepts – autonomy, situatedness, embodiment, and adaptivity – are of central significance to embodied cognitive science.

A further book of historical importance to the new approach is Valentino Braitenberg's "Vehicles – experiments in synthetic psychology" (Braitenberg 1984). In a series of thought experiments, Braitenberg describes 14 "vehicles" (i.e., autonomous agents) of increasing complexity. It is shown that even very simple systems can generate highly complex behavior. The simplest of these vehicles possesses only one sensor and one motor. Depending on the wiring of the single components, the system exhibits qualitatively differing behavior. If, for example, a light sensor is linked by a positive weight to the motor, the vehicle will move with increasing speed towards the source of light. However, if the same sensor is then linked by a negative weight to the motor, the vehicle will move away from the source of light. An outside observer may describe the first behavior sequence as "aggressive", and the second as "anxious". Braitenberg's other vehicles are designed with progressively complex links between multiple sensors and one or two motors.

In contrast to traditional modeling approaches, such as connectionism, in autonomous agents the observers fall out of the loop between input and output. In other words, the input-output-input loop is "complete". This poses a number of fundamental *challenges in the design* of such agents (cf. Scheier and Pfeifer 1999):

1. Sensors channel input from the environment into the system. As opposed to connectionist models that indiscriminately process and learn *every* input pattern, a system must be capable of determining which parts of the (typically) continually changing and high-dimensional input are relevant. Thus, for example, it must differentiate between noise and signals, and separate the signals into those that are relevant for the present task from other, irrelevant signals (the previously mentioned stability problem). Generally speaking, the system must be able to perceive the figure-ground relationship and also possess attentional control.
2. There is no clear delineation between the learning and test phases, as is the case in connectionist systems. On the contrary, an autonomous agent must ceaselessly learn (so-called incremental learning) in order to be truly adaptive. This poses substantial problems in the fields of learning theory and modeling, in that the neural networks in autonomous agents must be capable of continually absorbing and learning new material without becoming saturated. Accordingly, forgetting takes on a central role, and inevitably the question of which things should be forgotten arises (a stability-flexibility dilemma). In this, memory models that provide this dynamics in autonomous agents are required. The learning process cannot be supervised, as is the case in the majority of connectionist applications, because the agent must learn independently in order to be adaptive and situated.
3. The system's output comprises concrete kinetic actions performed by devices such as grippers or fingers. The central requirement is therefore to design a complete system that encompasses the entire sensorimotor palette. Consequently, the intrinsic question of how perceptual and cognitive processes should be mapped to motor action arises. This question can be probed into by referring to connectionistic models which view the categorization process as completed if the neural network activates a unit of the output layer that represents the categories. The activation of the unit triggers the network's response to a stimulus; this activation must then be interpreted by the designer, for the network itself does not generate behavior. This procedure is, however, not practicable in the case of autonomous agents because these are obliged to function without the help of the designer. How then must categories and ac-

tion be represented so that coherent behavior results? Put in concrete terms, the categorization mechanism must be embedded in the overall architecture of the agents.
4. The system's output largely determines the next input. As a consequence, the system must maintain consistency in its own output because mistakes in output lead unavoidably to mistakes in input, which in turn generates meaningless behavior in the long run. It is this recursiveness that precludes any arbitrariness in terms of whether or what type of mistake an agent may make in its behavior. This again derives from the fact that autonomous agents are complete systems. In building such agents it has been shown that precisely this recursiveness can be profitably exploited in the learning process. Instead of just passively registering input data, an autonomous agent can generate its own sensorial input, which by using appropriate strategies can also lead to a dramatic simplification of the perceptual and learning process problem.

In summary, agent-based theory and modeling are the source of a new catalogue of questions that coincides with certain areas of interest in psychology, such as attentional processes, incremental learning and forgetting, new memory models, categorization, and sensorimotor interfaces. This list represents only a few of the possible references to psychology, and it is viable that autonomous agents – apart from robotics in industrial applications – can be used as models in psychological systems.

3
Action and Self-Organization: Conceptualization of a Cognitive System

We have examined in what way embodied cognitive science differs from the ubiquitous symbol information processing approach in psychology. The problems intrinsic to symbol grounding, the selected frame of reference, and the combinatorial explosion are indicative of the shortcomings of the classical approach. How can the gap between cognition and environment be closed and, how must a complete system in the absence of a mediating homunculus be conceptualized? Quite obviously, increased emphasis on embedding in the environment and the embodiment of the active agent does not suffice. This observation in psychology is by no means new or original. The question is rather of *how* an embedding in the environment could theoretically be achieved.

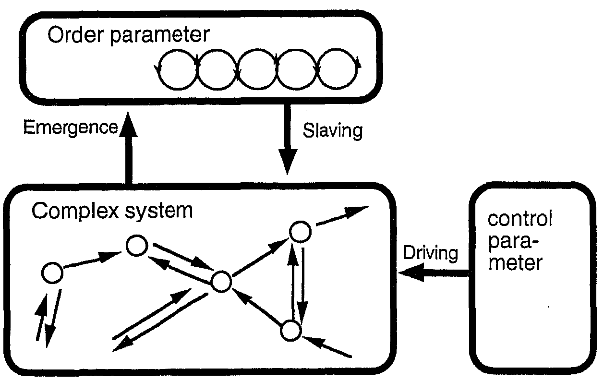
First of all, an attempt is made to control the relationship between the cognitive system and the environment in such a way that the problems discussed in the first section do not arise. How can adaptivity and intentional behavior result *without* either being considered a priori given in the sense of an 'intentional stance' (Dennett 1987)? In the domain of action psychology the definition of action is implicit; yet the very concept of action has proved to be the most difficult aspect in synthetic AI, especially implementing autonomy and adaptivity. Therefore, using intentionalistic terminology a priori would be inadequate for theoretical psychology. What conditions are necessary for systems to exhibit the qualities of *autonomy* and *adaptivity*? In our opinion, the answer to this question should point to a method which would be instrumental in designing artificial cognition in autonomous agents (Tschacher and Scheier 1996). At the same time, a possible fundament could be laid for an explanatory theory that goes beyond the premise of the intentional stance.

With regard to *autonomy*, the ability to generate patterns spontaneously and autonomously is a well-researched characteristic of complex dynamical systems (Haken 1990). Interdisciplinary research programs such as that of synergetics pinpoint the special relationship of systems to their environment as a requisite condition for the creation of emergent order (self-organization). These systems are sensitive to their environment and are constantly being "driven" by energy impulses. The energy throughput of a dynamical system is quantified by means of so-called control parameters. In this, however, the relationship of the system to the control parameters is not one of being "controlled" in the narrower sense; rather these parameters represent unspecific constraints which induce complex systems to produce autonomously, or even creatively, ordered patterns. The schema of a synergetic system is shown in Fig. 1.

Some intensively researched self-organized patterns in complex systems are, for example, the laser in optical systems or the Bénard convection patterns that appear in fluid systems during the transfer of heat. Further examples which are directly related to psychology are the formation of *Gestalts* in perception (Kruse and Stadler 1995; Kriz 1997) and the establishment of movement patterns (Kelso 1995; Leist 1999). The latter was mathematically modeled by Haken et al. (1985) using the simple paradigm of simultaneous index finger movements of a person. It was shown that only a synchronous, parallel or anti-parallel rhythmical movement of both fingers was possible. If the frequency of the finger movements – the control parameter of the system – was changed by setting a metronome to a faster pace, for example, typical phase transitions between the parallel and anti-parallel movement patterns are observed. Both movement patterns are to be viewed as order parameters of the system; they remain stable in the presence of interference and display the typical characteristics of nonlinear systems (such as hysteresis, critical slowing down).

With regard to *adaptivity*, the phenomenon of self-organization described above illustrates that under certain circumstances complex systems become independent of their environment in that they spontaneously create ordered and

Fig. 1. Schema of a synergetic system



stable patterns. The patterns are emergent characteristics of these systems; systems theory in the natural sciences has shown in this connection that "emergence" is not an esoteric term, but rather an ubiquitous characteristic of physically and biologically complex systems. However, defining the relationship of a system to its environment purely from the standpoint of a system creating a self-organized pattern would be an over-simplification. An essential aspect is the *systematically changed* relationship of the system to its environment (i.e., to the control parameters) which is established by the creation of a pattern. The creation of a pattern always occurs in such a way that the distance from thermodynamic equilibrium characteristic to dissipative systems is maximally reduced. If 'distance from thermodynamic equilibrium' is typified by the metaphor 'tension', then it is those patterns which best serve to reduce tensions that are created. The anti-parallel finger movement pattern better fulfills the requirements of higher rhythmical frequency (and is subjectively less strenuous), as galloping, for instance, is the more efficient pattern for quickly moving horses. Self-organization (pattern formation, autonomous order production) therefore follows an *optimization principle* (Tschacher and Dauwalder 1999). Swenson and Turvey (1991) derived the optimization principle of (local) order production from general thermodynamics premises: Self-organized patterns evolve "...because order produces entropy faster than disorder." (p 345). They view this principle as the basis of Darwin's theory of evolution: "The world is in the order production business, including the business of producing living things and their perception and action capacities ... ." This optimization principle is meant when self-organizing systems are characterized as "adaptive".

Within the context of cognitive science, it emerges that the perspective of self-organization may also represent an alternative fundamental theory applicable to cognition processes. In the ongoing discourse in cognitive sciences, this approach is termed the "dynamical systems explanation" (Clark 1997a). The inherent advantage of the dynamical systems explanation is not only that it is fundamentally an interdisciplinary theory of considerable scope, but also most particularly that it permits verifiable predictions of behavior and cognition (e.g., Haken et al. 1985). In conjunction with Haken (2000), therefore, we put forward the thesis that self-organization also plays a decisive role in the information processing of biological systems. "Information processing" is thereby given a novel meaning which markedly differs from the prevailing representational and rule-based models.

Moreover, through this approach links are created to other existing approaches in psychology, such as the ecological theory of perception (Gibson 1979; cf. Reed 1997) and *Gestalt* theory (Lewin 1936; cf. Tschacher 1997). In Gibson's theory of "direct perception", characteristic, invariant properties of the environment act as control parameters or "affordances". The term affordance was derived from the verb 'to afford' and reflects the behavior-"prompting" qualities of environmental stimuli; the evolution of this term can be traced back to Lewin's concept of "Aufforderungscharakter", and the later synonym "valence". According to Gibson, affordances are "directly picked up" from the environment, thus precluding the need of any mediating system. Affordances are analogous to the control parameters of a self-organizing system, which need not be translated at any point into an internal system code because they are "perceived" directly. This approach is highly engaging within the context of the current discourse in cognition theory because it circumvents not only the problems associated with the concept of an internal representation, but also those intrinsic in the competing constructivist concept of a "construction" of the world. The world is its own best model (Brooks 1991).

In his subsumptive approach for the design of autonomous agents, Brooks argues that the foundation of intelligence consists of multiple representation-free sensorimotor layers. Each layer reacts to either one or several specific affordances and performs a simple behavior (this concept corresponds to a large extent to Minsky's definition of "agent" - Minsky 1985). In his discussion of the subsumptive approach, Clark (1997a, p 32) foresees a coherence problem arising once the presence of too many layers complicates the system. How, then, can behavior be coordinated if a great many subsystems must concurrently work together and how is wide-spectrum behavior of the overall system possible in view of the requirements placed upon it by a multifaceted environment and a multitude of affordances? Clark argues that one is sorely tempted to sacrifice the original vision of a coordination among many directly perceiving and acting subsystems for the traditional conception of a centralized symbolic control and planning system. In our opinion, however, this coherence can be afforded by the principles of self-organization applied to cooperating or, alternatively, competing layers or "behavior kernels" (Tschacher 1997). According to the optimization principle, adaptive "intentional" behavior emerges from the synergetic coordination of many such layers and by the selection of those layers best suited to reduce the 'tension' of affordances (Fig. 2).

In the case of complex behaviors (e.g., human behavior), coherence is additionally fostered by a socially and culturally structured environment, and notably also by the evolutionary and far-reaching "discovery" of language. Language enables a form of symbolic representation in that sensorimotor layers need not be performed, but can be "emulated" instead. However, it is not possible for us at this point to examine the conception of language within the context of embodied cognitive science.

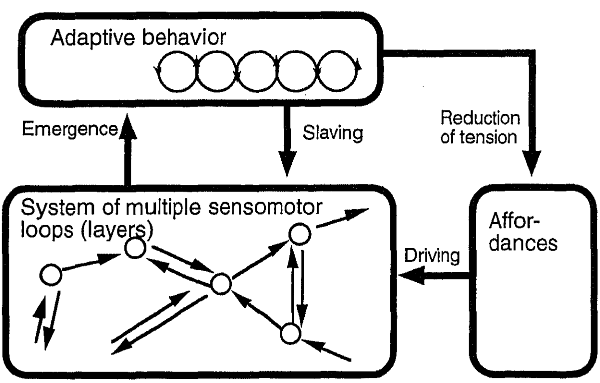


Fig. 2. Schema of a cognitive system from the viewpoint of embodied cognitive science

4
Implications for Psychology

In the 1960s it was a matter of importance to emancipate psychology from the sterility of behaviorism and to develop a theory that would enable the scientific investigation of the mind. The information processing approach presented as the physical symbol systems hypothesis (Newell and Simon 1972) seemed ideal for this purpose, because it permitted the modeling of cognitive processes independent of the machinery in which these processes were implemented. If 'cognition' can 'run' on a computer, then behaviorists could hardly condemn the notion of comparable cognitive processes taking place in the brain as mentalistic and unscientific. This argument overshadowed what had long been of central significance for psychology, namely the *autonomy* in information processing. Although it is implicitly clear in many applications of classical AI that the results generated by data processing must be interpreted by a human user, this implication in the case of a psychological theory inevitably invalidates any explanatory value of the theory. In other words, if the observer cannot fall out of the loop, and if 'symbol grounding' is only possible through his interpretation, then information processing psychology will remain dependent upon a homunculus.

The approach of embodied cognitive science as based on the dynamical systems theory has far-reaching consequences for applications in cognitive science. We will not delve into the ongoing fundamental reorientation that this approach brought about in the field of AI research, but rather investigate possible implications for psychology.

First of all, we discover that only the very first harbingers of this reorientation have succeeded in penetrating cognitive psychology. The unbroken supremacy of the computational metaphor remains to this day widely undisputed (Opwis and Lüer 1996; Pinker 1997). While the "dynamical challenge" is debated in cognitive science (Clark 1997b), in the field of psychology still only a few isolated, mostly clinical applications of "systemic thinking", largely lacking a scientific and empirical design, are present.

A significant part of the discourse on the topic of embodied cognitive science revolves around the conceptualization of cognitive representation and knowledge. Reinmann-Rothmeier and Mandl (1996) provide a summary of various theories of knowledge and action from the viewpoint of psychology. In general, they distinguish between an action-theoretical (e.g., Heckhausen et al. 1987) and a (in the sense of computationalism) system theoretical perspective (Anderson 1983). Both perspectives are based almost exclusively on different concepts of *internal* representation; accordingly, cognition is seen as a manipulation of an "internally" represented world model. However, the prevailing orientation towards a constructivist interpretation does not bring about a change in this. On the contrary, symbol grounding in constructivism becomes even more problematical in that adaptive action in a real world is to be understood within the context of an agent's internally constructed (instead of represented) "reality" alone.

The cognitive psychological approach described here builds on the concept of affordances. As a basic component, this approach always takes situated cognitive entities which are conceptualized as complete sensorimotor layers or "coordinations" (Clancey 1998) as its point of departure. The interaction of an agent with its environment cannot be set aside. Accordingly, knowledge is also not actually stored in the brain, but exists as a coupling of perception and action in the agent's environment (this environment has been additionally structured by the agent itself). Both the structured environment and the body (as the arena of sensorimotor loops) play a decisive role: the environment is to be viewed as an ecological niche originating from past couplings (Reed 1997).

This viewpoint leads to predictions that are empirically verifiable. One of these concerns so-called "external memory" – a topic of discourse in AI. Since autonomous agents do not (need not) construct explicit world models, specific interactions with the environment enable them to use the world itself as an external 'store'. The applicability of this prediction to human cognition is supported by various recent studies.

In a series of investigations, Ballard et al. (1995) examined the deployment of memory resources during the performance of certain tasks. As an example, subjects were required to reproduce a display using differently colored building blocks. In performing this task it was of interest how frequently the subjects referred to the original display by eye movements. It was shown that the subjects referred to the display more frequently than would be expected given an internal store. Instead of accessing an internal model, they referred to the external display directly.

A further phenomenon pointing to the frequency of human reliance on interactional mechanisms is "change blindness" (e.g., Simons and Levin 1997). Change blindness refers to the failure to detect changes in an object or a scene. The aim of this branch of research is to understand better the characteristics of representations. Experiments using highly varied methods and stimulus configurations have yielded converging results, indicating that surprisingly little visual information is represented from one moment to the next. In other words, recall of an object's characteristics and salient features is relatively transient. Simons and Levin's conclusions are consistent with the supposition of an external store for visual information. The authors postulate that people do not construct detailed representations of their environment, but rather that the essence, the "being-at-hand", of a given situation (i.e., its affordances) is actually perceived. Therefore, while unreliable, transient object characteristics are easily filtered out, the primary characteristics of a situation are nonetheless still perceived.

Horowitz and Wolfe (1998) reached a comparable conclusion in their experiments on the role of memory in "visual search" tasks: "Our results show that the visual system does not accumulate information about object identity over time during a search episode. Instead, the visual system seems to exist in a sort of eternal present.... The structure of the world makes it unnecessary to build fully elaborated visual representations in the head." (Horowitz and Wolfe 1998, p 577). As a rule, visual search theories presuppose that, during the search for a particular object, previously perceived, irrelevant objects are also remembered. This strategy intuitively makes sense in that attention is not detracted by unnecessarily searching in places containing irrelevant objects. However, Horowitz and Wolfe's experiments showed that the subjects' performance did not change when the objects in a search area were randomly moved every 100 ms to other areas. It was shown that the performance in both static and dynamic search tasks was the same. The random shifting of objects has no influence on searching time, which in itself indicates the involvement of interactional instead of representational mechanisms.

As a series of studies performed by Esther Thelen and colleagues have demonstrated (Smith et al. 1999; Thelen et al. 2001), these insights and ideas can also be fruitfully applied to developmental psychology. One focal point of these studies is the classic A-not-B task that was originally introduced by Jean Piaget and has since been repeatedly investigated in countless studies. The experimental setup consists of a box with two identical holes ("A" and "B"). While the child is watching, the experimenter puts a toy in hole A, covers the hole, and, after a short time (typically 3 s), pushes the box within the reach of the child, who then usually tries to extract the toy. After repeating this procedure two or three times, the toy is put in hole B (otherwise the procedure remains unchanged). The surprising and extremely robust effect is that children aged approx. 7–10 months persist in looking for the toy in hole A, although they have seen that it was put in hole B. Piaget maintained that this "error" was connected with poor object representation. Without going into the details of other related investigations, it is notable that the majority of conclusions assume a lack of spatial knowledge or a poor capacity to act. Thelen is now investigating an alternative avenue to probe into this phenomenon. To a large extent her approach is based on concepts of embodied cognitive science and dynamical systems theory. Instead of conceptualizing the error as poor information processing, she focuses on the child's behavior. Thelen showed that the error occurred *without* a toy being put in the box, therefore without an object to be represented. The central variable proved to be the number of reaching movements of the child. The more often the child reached for hole A, the greater the probability that it would make the A-not-B error. Moreover, the error was not made if the child's posture was changed (e.g., into an upright, standing position) before it started reaching for hole B. The error is therefore dependent upon the posture of the child. These and other results have been recently integrated in a model, which to a great extent uses the methods of the dynamic approach to embodied cognitive science (Thelen et al. 2001).

On the whole, it has been demonstrated that several of the central implications of embodied cognitive science – for example, the hypothesis of an external memory – find empirical support. However, further studies are necessary to substantiate these concepts. New possibilities are opening up through the use of autonomous agent instruments in psychological modeling, permitting the simultaneous investigation on various levels of behavior and underlying mechanisms. Thus, for example, Almassy et al. (1998) have demonstrated how robots can be used profitably in neuropsychological research. These investigators found that in their model specific characteristics of visual neurons (e.g., certain invariances) only then emerged if they were embedded in an active system.

5 Discussion

This work presented the embodied cognitive science approach and elaborated several implications for the study of intelligence as well as for psychological research and applications. In contrast to the symbol-processing hypothesis, embodied cognitive science assumes that the body and interaction with the environment are decisive for intelligence. Accordingly, in the modeling of cognitive processes, this approach uses autonomous agents (robots) as a methodological aid. Although the intrinsic characteristics of autonomous agents differ from those in connectionist models, connectionist neural networks can nonetheless be profitably integrated in autonomous agents.

Although one may cite literature that substantiates the basic principles of embodied cognitive science, at present the main obstacle for a wide acceptance of this approach in psychology is the lack of well-founded empirical support for these concepts. These innovative theoretical perspectives could be the basis of experiments which put the interactional aspect of cognitive performance in the foreground. In the study of visual perception, Milner and Goodale (1995) argue that the "theoretical commitment to vision qua perception" has contributed decisively to the methodology in perception research. Instead of investigating the relationship between motor output and visual input (in other words, interactional processes), the majority of studies focus on variants of visual discrimination. According to this paradigm, it is assumed that the motor action of a test subject is irrelevant to the performance of a task. Consequently, in animal trials, for example, it is of no relevance whether the test subject pecks, jumps, runs, presses a bar, etc. The only relevant factor is whether the human or animal test subject can differentiate between individual visual stimuli. However, such technically restricted experimental setups do not adequately deal with the situated character of many cognitive processes. The standpoint of embodied cognitive science calls for a focus on the dynamical loops linking sensory, cognitive and motor processes.

In the domain of memory research, Glenberg (1997) poses the question of what memory is for. He suggests that the answer lies in human-environment interaction. Memory develops so that human beings can successfully and adaptively interact with the environment. Here, too, further empirical studies addressing interactional aspects are required.

As educational and clinical-psychological applications have shown, embodied cognitive science stimulates the formulation of relevant questions. The transfer of knowledge into daily action is of at least equal importance as the acquisition of knowledge itself. To a certain extent, the transfer problem ('symbol grounding') is attributable to a blind spot in the computational viewpoint, which ideally should serve to motivate research work towards developing and optimizing teaching/learning strategies on the basis of a situated (embodied) approach (Mandl 1997). Accordingly, in the field of clinical psychology, the task of psychological (cognitive) therapy is not only, or even primarily, the diagnosis and remediation of cognitive dysfunctions as traditionally implemented in 'classic' cognitive-behavior therapy. Rather the actualization of the disorder in the therapeutic setting, i. e., the physical and emotional activation of the problematical sensorimotor patterns, is essential. This indispensable situatedness comes particularly to the fore in exposure-therapeutic interventions. We are of the opinion that it would be useful to emphasize more strongly situative and interactional aspects in other intervention methods. The embodied cognitive science approach offers hypotheses that lend themselves to empirical investigation.

The future will show to what extent the standpoint and rationale of embodied cognitive science will be integrated in the field of psychology. This seems likely since psychology has traditionally followed the developments of both cybernetic and computer technology with a certain time lag. The greatest benefit, however, could emerge once psychology rediscovered its role as an impulse generator in the fundamental research into cognitive modeling and cognitive technology.

References

Almassy N, Sporns O, Edelman G (1998) Behavioral constraints in the development of neuronal properties: a cortical model embedded in a real-world device. Cereb Cortex 8: 346–361

Anderson JR (1983) The architecture of cognition. Harvard University Press, Cambridge

Atmanspacher H, Dalenoort GJ (eds) (1994) Inside versus outside. Springer, Berlin Heidelberg New York

Ballard DH, Hayhoe MM, Pelz JB (1995) Memory representations in natural tasks. J Cogn Neurosci 7: 66–80

Barsalou LD (2001) Perceptual symbol systems. Behavioral and brain sciences (WWW preprint archive)

Bickhard MH, Terveen L (1995) Foundational issues in artificial intelligence and cognitive science. North-Holland, Amsterdam

Braitenberg V (1984) Vehicles – experiments in synthetic psychology. MIT Press, Cambridge

Brooks RA (1991) Intelligence without representation. Artif Intell 47: 139–159

Caspar F, Rothenfluh T, Segal Z (1992) The appeal of connectionism for clinical psychology. Clin Psychol Rev 12: 719–762

Clancey WJ (1991) The frame-of-reference problem in the design of intelligent machines. In: VanLehn K (ed) Architectures for intelligence. Erlbaum, Hillsdale, pp 357–423

Clancey WJ (1998) Situated cognition. Oxford University Press, Oxford

Clark A (1997a) Being there: putting brain, body, and world together again. MIT Press, Cambridge

Clark A (1997b) The dynamical challenge. Cogn Sci 21: 461–481

Dennett DC (1984) Cognitive wheels: the frame problem of AI. In: Hookway C (ed) Minds, machines, evolution. Cambridge University Press, Cambridge, pp 129–151

Dennett DC (1987) The intentional stance. MIT Press, Cambridge

Dreyfus HL (1972) What computers can't do: a critique of artificial reason. Harper and Row, New York

Franklin S (1996) Artificial minds. MIT Press, Cambridge

Gibson JJ (1979) The ecological approach to visual perception. Houghton Mifflin, Boston

Glenberg AM (1997) What memory is for. Behav Brain Sci 20: 1–19

Greeno JG (1989) A perspective on thinking. Am Psychol 44: 134–141

Haken H (2000) Information and self-organization (A macroscopic approach to complex systems), 2nd edn. Springer, Berlin Heidelberg New York

Haken H (1990) Synergetik – eine Einführung. (Nichtgleichgewichts-Phasenübergänge und Selbstorganisation in Physik, Chemie und Biologie). Springer, Berlin Heidelberg New York (2. Auflage)
Haken H, Kelso JAS, Bunz H (1985) A theoretical model of phase transitions in human hand movements. Biol Cybern 51: 347–356
Harnad S (1990) The symbol grounding problem. Physica D 42: 335–346
Heckhausen H, Gollwitzer PM, Weinert FE (eds) (1987) Jenseits des Rubikon: Der Wille in den Humanwissenschaften. Springer, Berlin Heidelberg New York
Hendriks-Jansen H (1996) Catching ourselves in the act: situated activity, emergence, evolution, and human thought. MIT Press, Cambridge
Horowitz TS, Wolfe JM (1998) Visual search has no memory. Nature 394: 575–577
Kelso JAS (1995) Dynamic patterns: the self-organization of brain and behavior. MIT Press, Cambridge
Köhler W (1947) Gestalt psychology: an introduction to new concepts in modern psychology. Liveright, New York
Kolers PA, Smythe WE (1984) Symbol manipulation: alternatives to the computational view of mind. J Verb Learn Verb Behav 23: 289–314
Kriz J (1997) Systemtheorie. Eine Einführung für Psychotherapeuten, Psychologen und Mediziner Wien: Facultas
Kruse P, Stadler M (eds) (1995) Ambiguity in mind and nature. Multistable cognitive phenomena. Springer, Berlin Heidelberg New York
Leist K-H (1999) Synergetische Lehr-Lern-Prozesse des Bewegungssystems. In K Mainzer (ed) Komplexe Systeme und Nichtlineare Dynamik in Natur und Gesellschaft (207–220). Springer, Berlin Heidelberg New York
Lewin K (1936) Principles of topological psychology. McGraw-Hill, New York. Deutsche Ausgabe: Grundzüge der topologischen Psychologie. Bern: Huber (1969)
Mandl H (1997) Wissen und Handeln: Eine theoretische Standortbestimmung. In H Mandl (ed) Schwerpunktthema Wissen und Handeln: Bericht über den 40. Kongreß der Deutschen Gesellschaft für Psychologie in München 1996 (3–13) Göttingen: Hogrefe
Miller GA, Galanter E, Pribram KH (1960) Plans and the structure of behavior. Holt, Rinehart, and Winston, New York
Milner G, Goodale N (1995) The visual brain in action. Oxford University Press, Oxford
Minsky M (1985) Society of mind. Simon and Schuster, New York
Newell A (1980) Physical symbol systems. Cogn Sci 4: 135–183
Newell A (1990) Unified theories of cognition. Harvard University Press, Cambridge
Newell A, Simon HA (1972) Human problem solving. Prentice-Hall, Englewood Cliffs
Opwis K, Lüer G (1996) Modelle der Repräsentation von Wissen In D Albert, KH Stapf (eds) Enzyklopädie der Psychologie, Themenbereich C: Theorie und Forschung, Serie II: Kognition Vol. 4, (337–431) Göttingen: Hogrefe
Pfeifer R, Scheier C (1999) Understanding intelligence. MIT Press, Cambridge
Pinker S (1997) How the mind works. Norton, New York
Pylyshyn ZW (ed) (1987) The robot's dilemma. The frame problem in artificial intelligence. Ablex, Norwood
Reed ES (1997) Encountering the world: toward an ecological psychology. Oxford University Press, Oxford
Reinmann-Rothmeier G, Mandl H (1996) Wissen und Handeln: Eine theoretische Standortbestimmung. Institut für Empirische Pädagogik und Pädagogische Psychologie, Universität München: Forschungsbericht Nr. 70
Scheier C, Pfeifer R (1999) The embodied cognitive science approach. In: Tschacher W, Dauwalder J-P (eds) Dynamics, synergetics, autonomous agents – nonlinear systems approaches to cognitive psychology and cognitive science. World Scientific, Singapore, pp 159–179
Simons DJ, Levin DT (1997) Failure to detect changes to people during a real-world interaction. Psychonom Bull Rev 5: 644–649

Smith LB, Thelen E, Titzer R, McLin D (1999) Knowing in the context of acting: the task dynamics of the A-not-B error. Psychol Rev 106: 235–260

Strube G (1990) Neokonnektionismus: Eine neue Basis für die Theorie und Modellierung menschlicher Kognition? Psychol Rundsch, 41: 129–143

Swenson R, Turvey MT (1991) Thermodynamic reasons for perception-action cycles. Ecol Psychol 3: 317–348

Thelen E, Schöner G, Scheier C, Smith LB (2001) The dynamics of embodiment: a field theory of infant perseverative reaching. Behav Brain Sci (in press)

Toda M (1962) The design of a fungus-eater: a model of human behavior in an unsophisticated environment. Behav Sci 7: 164–183

Tschacher W (1997) Prozessgestalten – Die Anwendung der Selbstorganisationstheorie und der Theorie dynamischer Systeme auf Probleme der Psychologie. Hogrefe, Göttingen

Tschacher W, Scheier C (1996) The perspective of situated and self-organizing cognition in cognitive psychology. Commun Cogn Artif Intell 13: 161–188

Tschacher W, Dauwalder J-P (1999) Situated cognition, ecological perception, and synergetics: a novel perspective for cognitive psychology? In: Tschacher W, Dauwalder J-P (eds) Dynamics, synergetics, autonomous agents. World Scientific, Singapore, pp 83–104

Ullman S (1996) High-level vision. MIT Press, Cambridge

Varela F, Thompson E, Rosch E (1991) The embodied mind. MIT Press, Cambridge

Vera AH, Simon HA (1993) Situated action: a symbolic interpretation. Cogn Sci 17: 7–48

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