THE IMPLICATIONS OF EMBODIMENT

Cognition and Communication

edited by Wolfgang Tschacher and Claudia Bergomi
1. Towards a general formulation of self-organization

In this chapter, we wish to study embodied cognition on the basis of self-organized pattern formation. We will interpret pattern formation in a rather wide sense: one may think of spatial, temporal, spatio-temporal, or behavioral patterns. We will introduce the idea of pattern formation by means of a few examples taken from various fields of science.

In physics, a fluid in a container heated from below may spontaneously form hexagonal cells (establishing the so-called Bénard convection, Fig. 1 left). In each cell, a fluid compartment rises due to the heating, cools at the upper surface of the container, and eventually sinks down at the borders of the rising compartment. This revolving conical pattern is dynamical in nature. An example of pattern formation in optics is the laser, which has become a paradigmatic system for synergetics; we will elaborate it below. Well-known further examples of self-organizing dynamics come from chemistry: specific substances, when poured together, may spontaneously form rotating spirals or concentric waves that run outwards (the
Belousov-Zhabotinsky reaction, Fig. 1 right). In biology, we observe the growth of organisms with their highly sophisticated structures, which may also be conceived as patterns. Finally, one may think of behavioral patterns, the simplest ones perhaps occurring in movement coordination such as the gaits of quadrupeds or hand movements (Haken et al., 1985). An important aspect of all such pattern formation is the fact that these patterns arise spontaneously, i.e. there is no ordering hand that creates them like a sculptor would do.

Some 40 years ago, one of us initiated an interdisciplinary field of research that aimed at studying the self-organized formation of patterns from a unifying point of view. We called it synergetics, signifying a science of cooperation (Haken, 1977). Synergetics combined methods of dynamical systems theory (which it developed even further) with methods of statistical physics (thereby taking into account the important role of fluctuations, i.e. of chance events). Synergetics is focused especially on those situations where new patterns arise. Initially, in order to express the synergetic approach as clearly as possible, we chose examples from physics, especially the light source laser, as paradigmatic systems. As an interdisciplinary field, however, applications of the synergetic principles have ranged from physical to social to psychological and neuronal dynamics. This was the agenda of a series of international interdisciplinary symposia starting in 1972\textsuperscript{1}. The field has generated numerous publications and a book series with proceedings, monographs and edited volumes (the Series on Synergetics, published by Springer Verlag, Berlin).

A major goal within synergetics was to base brain functioning on abstract assumptions, which can be formulated independently of a specific physical substrate. One of us (Haken, 1996) has suggested to treat the brain as an open physical system, which on the one hand obeys the fundamental laws of physics, but is also subject to laws at the superordinate level of synergetics. We may thus say that the brain is an open and complex physical system, to which the mathematical tools and concepts of synergetics may be applied. The corresponding systems concept can be summarized as follows: A system is conceived of as an ensemble of individual components interacting with each other; at the same time, the system is embedded in an environment, with which it may exchange energy, matter and/or information. Open systems are those, which are maintained in their ordered dynamical state by an influx of energy, matter or information with a corresponding outflux of degraded energy or matter (on this degradation process focuses the free-energy principle: See Friston, this volume). In mathematical terms, the level or amount of such influx serves as control parameter(s) of the system. Additionally, the individual components of the system are subject to continuous stochastic

\textsuperscript{1} for a list see www.upd.unibe.ch
fluctuations, which may be of internal or external origin. Such stochasticity has turned out to be quite essential for the initiation of ordering processes.

When a control parameter exceeds a critical value, fluctuations may trigger the establishment of a variety of spatial, temporal or behavioral patterns. In many cases, the single potential patterns are mutually exclusive, i.e. in general they appear to ‘compete’ with one another. In these cases, one specific pattern wins the competition and consequently acts as order parameter. In some cases, several order parameters may result from the competition, which then govern a coexistence of different patterns. While the order parameters are brought about by the cooperation of the individual components, they in turn prescribe the behavior of the latter parts, or in other words, they enslave the components (‘slaving principle’). The slaving principle is an example of circular causality.

Of particular interest to us are so-called transients close to the critical values of the system. Transient dynamics occurs during the buildup of a pattern; as soon as the control parameter is switched off the pattern decreases. Thus we may distinguish three phases:

Phase 1: Buildup of a highly ordered coherent state by means of the recruitment of components;
Phase 2: Fully developed pattern of the coherent state;
Phase 3: Decay of that pattern.

A paradigmatic model of the spontaneous formation of ordered states is a gas laser (Fig. 2). The gas, composed of atoms, is enclosed in a glass cylinder. At its end faces, mirrors are mounted that serve to reflect light running in axial direction, so that it can stay comparatively long inside this device before it eventually exits through one of the half-silvered mirrors. The laser allows us to exemplify the various aspects of the buildup of self-organization mentioned above.

The laser system is composed of atoms as well as of light waves emitted by the atoms. The environment contributes to self-organization in two ways. First, there is a static environmental contribution: The glass tube provides confinement of the atoms; the mirrors at the end faces serve for a pre-selection of permitted light waves. Second, the environment contributes dynamically: An electric current, generated by a battery as an energy source, is sent through the device. This free energy lifts atoms from their resting states to excited (energy-richer) states. The control parameter here is the strength of the electric current — if strong enough, it can generate a sufficiently large number of excited atoms. As soon as this number is larger than that of the atoms in the resting state, laser physics speaks of an ‘inversion’.
Fluctuations are provided by the spontaneous emission of light wave tracks that are rather short, e.g. of one meter length. At this point in Phase 1, competition sets in: When a positive inversion is reached, avalanches of light waves may be formed. That means, a light wave impinging on an excited atom forces the latter to enhance this wave. This process, when continued, leads to the avalanche. There are different kinds of avalanches of different wave lengths that compete with each other. The most efficient avalanche ‘wins’ and becomes the order parameter of the laser system. This is a wave that is practically infinitely long. Laser action does evolve only if the inversion is kept sufficiently high by means of the electric current. The order parameter then emerges as a macroscopic wave of high amplitude (Phase 2). According to the slaving principle, the order parameter forces the individual atoms to behave in such a way that they maintain its existence, by the mechanism of circular causality (Tschacher & Haken, 2007). Through the ongoing laser process, however, the inversion is continuously depleted: Stronger order parameters entail faster depletion. This can be shown both mathematically as well as empirically when the electric current through the glass tube is suddenly switched off (Phase 3). This fast depletion of inversion is achieved by the coordinated action of the atoms of the gas. In other words, depletion is achieved by the action of the order parameter until, eventually, the resting state of all atoms is again realized.

We have now arrived at a picture, that of a transient system, needed to illustrate our approach to embodied cognition. To summarize this process in the laser: By means of a strong enough electric current suddenly a critical, sufficiently high inversion is reached. Then in Phase 1, by means of fluctuations, a competition of initiated patterns sets in. In Phase 2 a coherent wave, the order parameter, emerges. Phase 3 is reached when the coherent wave decays as a consequence of the inversion being lowered by the action of the wave.
2. Embodied cognition, not only computation

In this section we will describe the shift of paradigm that has occurred in cognitive science during recent decades. This shift has resulted in a general focus on the embodiment of most, if not all cognitive acts. In its general form, the concept ‘embodied cognition’ conveys the idea that cognition must always be viewed in context. A shared conviction has emerged in the cognitive science community as well as in cognitive psychology that one should address more deeply the intrinsic relationship between cognition and its environment. Due to the embodiment perspective, the ecology of cognition deserves more profound and careful investigation.

The increasing appeal of embodiment derives from a confluence of different fields of cognition research. To begin with, there is a continuous line of philosophical thinking, that of phenomenology, which has always emphasized the role of the body for the mind (Heidegger, 1927). Another important origin goes back to the quite distant field of informatics and computer science, which has encountered a large-scale engineering fiasco in the last century: Despite the work of decades, the creation of artificial intelligence (AI) has been largely unsuccessful. The consequence drawn by a majority of AI researchers and cognitive scientists has been to move into the field of embodied agents and robotics (Brooks, 1991; Hoffmann & Pfeifer, this volume). It is widely recognized today that intelligent cognition on the basis of symbol manipulation alone is unattainable; one must therefore regard classical AI as a failed paradigm and contemplate the reasons of its failure (Dreyfus, 1992).

This changing of perspectives in informatics has entailed marked reverberations in cognitive psychology. Why especially in psychology? — in the 1960s, after decades of behaviorism, psychologists were actively searching for a justification to scientifically investigate (again) cognition and thinking. Cognition has attributes that appeared elusive to scientific study (it is a largely subjective phenomenon; unlike physical processes, it is intentional in the sense of Brentano, 1874). In a strictly behaviorist scientific framework, the promise that intelligent and conscious computers would be a reality well before the end of the 20th century (Minsky, cited in Dreyfus, 1992) was perceived as a great relief by psychologists. Their conclusion was that, if even machines can be developed to become thinking machines, it should be well justified to again explore thinking in humans. Important implications followed from this, especially that human cognition should be conceptualized along the lines of computation, of computer-like information processing and manipulation of symbols. In the computational framework, these symbols are conceived as (per se) meaningless tokens, as mere stand-ins for items of the outside world, which they represent internally. The methodology to describe (and maybe
eventually synthesize) cognition was derived from propositional logic, predicate logic and set theory.

When the classical concept of AI finally failed, however, and the new research program of embodied agents was gradually instantiated (Brooks, 1991), the message for cognitive psychology was modified once more: Psychology likewise developed in the direction of embodied cognition. This road was additionally paved by experimental findings in social psychology, which had shown over the years to what extent social cognition was influenced by motor behavior and posture (e.g., Strack et al., 1988). The conclusion of abandoning pure computation in favor of embodied cognition was supported by a further development: The rise of neurobiology. To a biologist, cognition is a natural product of the body, especially of the activity of neurons. Embodiment is not synonymous with neuroscience; yet the ‘decade of the brain’ (proclaimed by a US president in 1990 and now entering its 21st year) likely provided an additional boost for the research program of embodied cognition.

3. Embodiment: The ecology of cognition

Our argument in this section will be that, since the metaphor of cognition as computation has been largely abandoned, the concept of cognition (as embodied cognition) needs a new formalism. We think that the framework and principles of synergetics can provide such a new formalism.

To align our proposal with the development of psychology, let us consider Kurt Lewin’s topological psychology, an early realization of a dynamical systems theory in psychology (Tschacher & Dauwalder, 1999a; Tschacher, 1997). Lewin (1936) defined the ensemble of all psychologically active variables as ‘life space’ (Lebensraum). Life space consists of a person \( P \) together with this person’s psychological environment \( U \). Any behavior \( V \) (i.e. any reorganization of life space) was regarded as depending on the state of life space itself at a given moment in time, and was described by Lewin simply as a function of \( P \) and \( U \):

\[
V = f(P, U)
\]

(1)

Life space is itself embedded in a ‘foreign hull’ of non-psychological variables (e.g., the architectural environment). This hull comprises the static environmental contribution mentioned in section 1. The explicit consideration of the environment as a set of psychological variables \( U \) on the one hand and as a hull of behavior on the other was among the concepts that stimulated the psychological field of ecological psychology (Barker, 1968).

Life space is in itself not a dynamical construct; therefore, a dynamic environmental contribution is needed to understand how change comes about. In psychology this influence is called ‘motivational’. Therefore,
Lewin introduced the concept of valence, which imposes a psychological vector field onto life space. Valence consequently became the focal concept in Gibson’s (1979) theory of ecological perception (then termed ‘affordance’): “...affordances and only the relative availability (or nonavailability) of affordances create selection pressure on the behavior of individual organisms; hence, behavior is regulated with respect to the affordances of the environment of a given animal.” (Reed, 1996, p. 18).

Affordances may be conceived of as environmental resources encountered by an agent. Thus, affordances exist as properties of environmental niches, independent of an agent. Reed (1996) rejected the mutualist position that affordances exist only through the interaction of a specific environment with an agent. For example, a barstool affords sitting for most adult humans. It does not afford sitting for an elephant or a human infant simply because they could not realize the chair’s affordance, even though it exists. In other words, in the view of ecological psychology, agent and environment are connected by affordances, which are environmental properties. Affordances are simply there to be ‘picked up’ or utilized. The notion of a pick-up of information constitutes the Gibsonian view of direct perception (Gibson, 1979; Greeno & Moore, 1993). In its emphasis on ecological perception, the notion of a pick-up of information conforms with a Darwinistic, functional view: The functionality of an object, its resourcefulness and Zuhandenheit (Heidegger, 1927), are perceived primarily, the ‘ontological’ object-per-se is represented in second line (if at all).

This ecological perspective can be easily developed towards a unified concept of ‘situated cognition’ (see Tschacher & Dauwalder, 1999a). But how about embodied cognition? Situatedness and embodiment are closely linked in that they both point to the mechanisms by which variables afford behavior of an agent. Situative and bodily variables both comprise the environment of the agent.

4. A model of embodied cognition

There is agreement in the dynamics community that cognitive pattern is not pre-programmed but is ‘soft-assembled’ by pattern formation and self-organization (Thelen & Smith, 1994; Tschacher & Dauwalder, 2003). This assumption was introduced into neurocognitive science by synergetics and complexity theory (e.g., Haken, 1996; Kelso, 1995) in continuation of previous cognitive theory in Gestalt psychology, especially by Köhler (1920) and Lewin (1936). The alignment of synergetics and Gestalt psychology was motivated by evident and deep similarities between the properties of dynamical attractors and those of Gestalt perception (e.g. Haken & Stadler, 1990).

The computational paradigm in classical cognitive science has shown an inclination to address ‘higher’ cognitive functions. Given the develop-
ments and problems that we have outlined in the previous section, it was wise to put cognitive theory back on its feet again. Accordingly, dynamics would start in a bottom-up fashion at the sensorimotor level, considering perception-action loops (Clark, 1997) and the ecological embeddings in which these loops occur. Thus cognition can be grounded by an embodied and situated approach. Several of the hotly debated problems consequently appear in a different light: The dynamical view would not investigate ‘symbol grounding’, but rather the emergence of symbols as pattern formation. The dynamical view would not regard mental representation as primary, but rather how perception-action loops are evoked in their valent environment.

One of us (Tschacher, 1997) has conceptualized agent-environment coupling starting from Lewin’s life-space treatment. In this model, self-organizational processes in agent-environment interaction can be illustrated. According to Lewin, temporal progression of behavior takes place ‘contemporally’, without causal dependence on past or future states of the life space (see equation (1)). Instead, behavioral change was represented in his model by vectors and force fields (the dynamical components of life space), which attach to the objects in life space supplying temporal causation. For our present purposes, we prefer to avoid Lewin’s notion of contemporality, and ‘dynamise’ life space (1) by:

\[
\frac{dP}{dt} = f_1(P, U, \nabla) + \varepsilon_t, \tag{2}
\]

\[
\frac{dU}{dt} = f_2(P, U, \nabla) + \varepsilon_t, \tag{3}
\]

This dynamical and recursive reformulation of (1) is in the spirit of dynamical systems theory. The motivational terms of Lewin’s vector psychology are substituted by the differentials (i.e. changes) of agent/person $P$ and environment $U$. $P$ denotes the state of a person in $m$-dimensional person space, comprising cognitive and emotional variables needed to describe the person. $U$ denotes the state of the environment in $n$-dimensional environment space. $U$ contains all further variables that can affect the cognitive and emotional variables of $P$, such as bodily states of $P$ and objects perceived by $P$. $\nabla$ is an operator for partial derivatives (to deal with possible inhomogeneities of life space), $\varepsilon_t$ stands for stochastic noise. Expressed in Piagetian terminology, (2) addresses the processes of accommodation (i.e. how cognitive schemata adapt to environmental facts), and (3) those of assimilation.
For reasons of simplicity, we assume that the environment $U$ is constant over the time scales that are of interest here, i.e. we focus on a person’s accommodation. Equation (3) then collapses into a set of parameters $\mu$, which act as environmental control parameters of the person’s change. Note that we thus imply that bodily variables of $P$ are a part of the environment of the person. Consequently, the system (2), (3) can be written as:

$$\frac{dP}{dt} = f_\mu(P, \nabla) + \varepsilon$$

This formulation has significance for the empirical analysis of multivariate time series, e.g. in psychotherapy research. It means that the change of a person is a function of previous states of the person. With some assumptions (ergodicity, linearity) this can be translated to a simpler approach, which is accessible for vector autoregression (VAR) modelling (Tschacher & Ramseyer, 2009).

(4) can be examined by linear stability analysis (see Haken, 1988, p. 46ff; Haken & Wunderlin, 1991, p. 219ff) to check the stability of a point attractor, which is given by

$$\frac{dP_0}{dt} = f_\mu(P_0, \nabla) + \varepsilon = 0$$

$P_0$ is the state of the person at some stable point induced by environmental (including embodied) constraints $\mu_0$. This stable point attractor in state space is continuously challenged by the fluctuations $\varepsilon_t$, which add small perturbations $\partial P$ to the person’s state $P_0$. Thus, the solution for $\frac{d(P_0 + \partial P)}{dt}$ is required which can be written as

$$\frac{d(P_0)}{dt} + \frac{d(\partial P)}{dt} = f_\mu(P_0 + \partial P, \nabla)$$

The right hand side of (6) can be expanded in a Taylor series in $\partial P$: $f_\mu(P_0, \nabla) + l_\mu(P_0, \nabla)\partial P + l_\mu(P_0, \nabla)(\partial P)^2 + ...$. The first term of the series is identical to the first term on the left side of (6) and can therefore be eliminated. Squares and higher powers may be neglected if $\partial P$ is very small. Thus, a solvable linearized equation remains where eigenvalues $\lambda_i$ can be computed. The eigenvalues characterize the stability of a state $P_0$ of a person.

$$\frac{d(\partial P)}{dt} = l_\mu(P_0, \nabla)\partial P$$

This formalization shows that attractors in the life space of a person exist
H. Haken & W. Tschacher

and demonstrates, in principle, how they can be found. Such attractors at $P_0$ may be realized as cognitive schemata or emotional states.

We would generally assume $P$ to consist of a large number of variables. Thus, $P_0$ denotes a complex, i.e. a very high-dimensional state. Correspondingly, the number of characterizing $\lambda_i$ would be high. At this point the core of synergetic theory comes to bear, i.e. the mathematical formulation of the ‘slaving principle’. Haken (e.g., 1988, p. 48f) demonstrated that the number of degrees of freedom may be drastically reduced. He showed that one or a few variables $P_u$ — those variables with index $u$ that become unstable at critical values of control parameters $\mu$ — succeed in entraining and synchronizing all remaining variables. This was described by the examples of section 1. These variables are called the order parameters of the system $P$. They reduce the dimensionality of $P_0$ enormously by application of the slaving principle to all stable components of $P_0$.

We have used the formalization in this section to demonstrate that, building on Lewin’s psychological assumptions, the principles of synergetics can be applied to life space, a complex psychological system. Cognitive or emotional attractor states of a person are conceived of as being affected by environmental constraints, including constraints due to bodily variables. Such attractors are therefore examples of embodied cognition. In terms of ecological psychology, the body creates affordances that then shape cognition. The mathematics used here is of course tautological, reflecting predominantly our theoretical assumptions. Tautologies do not prove or disprove hypotheses, but they demonstrate that the frameworks that we have combined — ecological psychology, dynamical systems theory and synergetics — provide a consistent picture for embodied cognition. Ensuingly, it is the task of empirical studies to support the various predictions that result from our theoretical platform.

5. Discussion and implications for empirical work

Efficiency and intentionality

The synergetic model addresses the three phases of a transient system as described in section 1. In Phase 1, the buildup of pattern is initiated. In a competition between components of the system, one of many potential internal states $P_u$ of the person eventually prevails. Hence, which features are specific for the successful component $P_u$?

Although fluctuations $\partial P$ play a decisive role in the emergence of a stable state of embodied cognition, the resulting $P_u$, however, is definitely not arbitrary. This can be easily shown in any of the paradigmatic self-organizing systems mentioned in the introduction: When the buildup of
pattern in Phase 1 is repeated in different runs of an experiment, the same pattern (i.e. the same order parameter) is generated if the conditions (control parameters and the static environment) are equal (we disregard bistability here, as in clockwise vs. counterclockwise rotation of Bénard cells, because it has no influence on the efficacy of a pattern).

We have proposed previously that those patterns are established which are ‘efficient’ or ‘optimal’ (Tschacher, 1997; Tschacher & Haken, 2007; Haken & Tschacher, 2010). This proposal must obviously specify the reference, quasi the ‘purpose’, of efficiency or optimality: Self-organizing systems show pattern formation about what? We have suggested that this reference relates specifically to the control parameters of the system: These are always reduced by the coordinated action. In the case of the laser system, the inversion is continuously depleted by the coordinated laser light (the order parameter): The stronger the order parameter, the faster the depletion. Laser action dies out as soon as the inversion drops below a critical level (Phase 3). Below that level, only incoherent waves and eventually no waves are emitted any more. If, as in many applications of laser physics, continuous laser action is desired, then the depletion of inversion must be counteracted by a continuous electric current.

It is true that the buildup of laser action in Phase 1 is afforded (to use Gibson’s terminology) by the electric current that caused the inversion. The depletion of inversion by laser action in Phase 3 is thus the flip-side of Phase 1 buildup, but depletion is of specific interest with respect to efficiency: It gives us a clue that laser action is about the reduction of its control parameter. The same dynamics occurs in any self-organizing system: The pattern reduces the free energy influx, which is described by the control parameters, in an efficient and targeted manner.

Why is this meaningful? The aboutness of self-organized action apparently provides us with a physical analog for aboutness in the philosophy of mind: Intentionality (Tschacher, 2009). Intentionality has been proposed as a characterizing property of mental acts, as the distinguishing feature of a process being mental (Brentano, 1874). It is intriguing to consider that material and mental dynamics may not be categorically different with respect to intentional behavior. Hence self-organizing systems may be those physical systems that can mimic intentional behavior of the mind to a certain degree. This is likely the way how the brain can host the mind (this admittedly expressed in dualistic language). Quite evidently, the mind-body problem is concerned here. According to the interpretation suggested by synergetics (Haken, 1996) we may also be dealing with mind and body as identical, like two sides of the same medal, the medal being the order parameter.
Efficient embodied agents

One goal is to understand intelligent behavior, a further goal is to synthesize it. The synthetic goal has been the main entry on the agenda of the Artificial Intelligence (AI) research program, which has come to a standstill in the so-called ‘AI winter’. It would be too complicated here to diagnose the reasons that have led AI into this state of affairs. Nevertheless, as indicated above, such explanations can be highly illuminating for understanding intelligent behavior, i.e. both for the philosophy of mind and for psychology.

It is our impression, from a synergetic point of view, that AI has neglected two aspects of intelligence: First, embodiment variables have not been considered as control parameters of cognition; second, the property of intentionality was insufficiently covered in the propositional framework. Let us briefly discuss these two shortcomings. The problem of lacking embodiment has been remedied largely in recent years, at least theoretically (Pfeifer & Scheier, 1999; Tschacher & Dauwalder, 1999b). In the meanwhile, embodied autonomous agents rank high in contemporary robotics research.

We assume that especially the second point needs closer attention. In order to synthesize intelligent agents one must focus on the prerequisites of intentional or intentional-like systems. How can these prerequisites be created? It has proven dysfunctional to prewire intelligent cognition in computational architectures. A complex driven system is needed, in which patterns are allowed to emerge spontaneously; these same patterns must then be used by the system to perceive and categorize events. In other words, perception and action should be engineered in a complementary fashion, as sensorimotor couplings (Jordan, 2003), treating perception and action as a unity described, in terms of synergetics, by a single order parameter. Complexity is a trivial yet necessary further prerequisite for self-organized patterns to emerge: All Darwinistic, competitive processes demand a large number of components upon which selection pressure can be applied. Complexity is a condition for intelligent processes to come to the fore — there can be no intelligence in non-complex circumstances, unless an intelligent agent is already present. Given that intentionality is necessary for autonomous intelligent agents, is embodiment likewise a necessary condition? The only known types of intelligence, i.e. animals, are definitely embodied, suggesting that embodied cognition is required. This is only an induction, not ruling out the feasibility of disembodied intelligence: But neither AI nor metaphysics have as yet provided support for this hypothesis.
The Transfer of Principles of Non-Equilibrium Physics to Embodied Cognition

References


