20

Implications of Action-Oriented Paradigm Shifts in Cognitive Science

Peter F. Dominey, Tony J. Prescott, Jeannette Bohg,
Andreas K. Engel, Shaun Gallagher,
Tobias Heed, Matej Hoffmann, Günther Knoblich,
Wolfgang Prinz, and Andrew Schwartz

Abstract

An action-oriented perspective changes the role of an individual from a passive observer to an actively engaged agent interacting in a closed loop with the world as well as with others. Cognition exists to serve action within a landscape that contains both. This chapter surveys this landscape and addresses the status of the pragmatic turn. Its potential influence on science and the study of cognition are considered (including perception, social cognition, social interaction, sensorimotor entrainment, and language acquisition) and its impact on how neuroscience is studied is also investigated (with the notion that brains do not passively build models, but instead support the guidance of action).

A review of its implications in robotics and engineering includes a discussion of the application of enactive control principles to couple action and perception in robotics as well as the conceptualization of system design in a more holistic, less modular manner. Practical applications that can impact the human condition are reviewed (e.g., educational applications, treatment possibilities for developmental and psychopathological disorders, the development of neural prostheses). All of this foreshadows the potential societal implications of the pragmatic turn. The chapter concludes that an action-oriented approach emphasizes a continuum of interaction between technical aspects of cognitive systems and robotics, biology, psychology, the social sciences, and the humanities, where the individual is part of a grounded cultural system.

Embodied Cognition and Enactivism

The concept of embodied cognition generally considers extra-neural bodily structures and processes important for cognition. There are a number of different theories of embodied cognition. Wilson (2002) distinguishes between theories that emphasize:

- Situatedness
- Online or real-time processes
- Off-loading cognitive processing onto the environment (also referred to as embedded or scaffolded cognition)
- The idea that the environment itself is part of the cognitive system (sometimes called extended mind or distributed cognition)
- The idea that cognition is for action (sometimes called action-oriented or enactive cognition)
- The idea that offline cognition is body-based

We focus on enactive or action-oriented cognition, but we do include concepts that involve embedded/scaffolded and extended/distributed cognition. Extended mind approaches include the idea of action-oriented representations, whereas more radical versions of enactivism eschew representationalism (Hutto and Myin 2013).

Action-oriented theories that emphasize sensorimotor contingencies (O'Regan and Noë 2001; Noë 2004) are usually put under the heading of enactivism. More radical forms of enactivism include the idea of sensorimotor contingencies, but also emphasize other aspects of embodiment such as affectivity (interoceptive, autonomic, emotional aspects), reward, interest (motivation), and embodied social interaction (Varela et al. 1992). Both enactive and extended approaches endorse the idea that cognition is not just "in the head." Most embodied approaches accept the Gibsonian idea of affordances. Despite certain common elements, however, it is safe to say that not all embodied cognitive theories agree on all issues.

Action-oriented theories of embodied cognition (including enactive and extended) are prefigured in the work of American pragmatists such as Peirce (1887) and Dewey (1896, 1916, 1938). Dewey, for example, spoke of the organism environment as a single unit of explanation, suggesting that the environment is not only physical but also social, and emphasizing movement and our use of tools and instruments as part of cognition (see Menary 2007, 2010). Enactive theories find important sources in the European philosophical tradition of phenomenology (Husserl, Heidegger, and Merleau-Ponty) where perception is characterized as pragmatic, guided by what the agent can do (i.e., by embodied skills and motor possibilities). Phenomenological contributions to cognitive science champion embodied-enactive aspects. Such approaches include neurophenomenology (Varela 1996), where experimental subjects are trained in phenomenological methods that focus on first-person experience,

and "front-loaded" phenomenology, where phenomenological concepts or distinctions (e.g., sense of agency vs. sense of ownership for actions) are incorporated into experimental design (Gallagher and Zahavi 2012).

The enactive approach can be characterized by the following background assumptions (Varela et al. 1992; McGann 2007; Di Paolo et al. 2010; Engel 2010; Engel et al. 2013):

1. Cognition is considered as the exercise of skillful know-how in situated and embodied action.
2. Cognition structures the world of the agent, which is not pregiven or predefined.
3. Cognition is not viewed as happening only in the brain; it emerges from processes in the agent-environment loop. The intertwinement between agent and world is constitutive for cognition.
4. System states are thought to acquire meaning by their functional role in the context of action, rather than through a representational mapping from a stimulus domain.
5. The approach aims at grounding more complex cognitive functions in sensorimotor coordination.
6. In contrast to classical cognitive science, which has an individualistic and disembodied view of cognition, the enactive approach strongly refers to the embodied, extended, and socially situated nature of cognitive systems.
7. The approach implies strong links to dynamical systems theory and emphasizes the key relevance of dynamic coupling and dynamic coordination.

As illustrated in Figure 20.1, the pragmatic turn can be characterized in terms of various brands of action-oriented cognition. Parallel to the advent of enactivism, such a turn has independently occurred in classical representational approaches as well. Importing action into the study of cognition cannot be considered a unique signature of enactivist agendas. Three major moves that one may discern in representational approaches to cognition and action are:

1. Theoretical: Cognition-for-action takes a new look at cognitive functions (e.g., perception and attention) and emphasizes their role in action. The aim is to find out how these functions are constrained and shaped by the requirements of action control.
2. Experimental: Action is viewed as an object of study itself (with regard to production and perception as well as their mutual relationships). The aim is to investigate the representational underpinnings of action representation, with particular emphasis on scenarios of social interaction and communication (where action production and perception are combined).
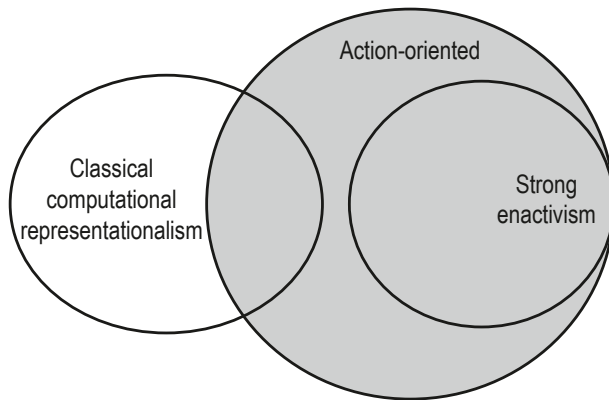
Figure 20.1   The primacy of action is constitutive of enactive and action-oriented approaches. Representational approaches address aspects of cognition for action. This common ground constitutes the pragmatic turn.

  3.   Theoretical and experimental: Action is seen as a constitutive ingredient of cognition (particularly high-level cognition). The aim is to trace the latent and implicit involvement of action in cognition (coming fairly close to what nonrepresentational enactivism claims as well).

Action-related approaches to cognition have important characteristics. First, they allow for actions to generate and modulate perception, which allows for much more dynamic and active perception. Second, they allow for an ontogenetic perspective that can account for dynamic changes of the cognitive system over time and development. This approach is thus much more flexible than the more or less invariant information-processing machinery assumed by the cognitive approach. Third, the assumption that cognition is for action allows for a reinterpretation of various cognitive phenomena, especially limitations of processing. For instance, considering that attention may ultimately serve action allows one to understand that the purpose of selection is not to shield a capacity-limited system against stimulus overflow but to allow actions to be optimally fed with actual relevant information. If there is a limitation, it is on the action side—not on the input side.

Does This Pragmatic Turn Constitute a New Paradigm?

Kuhn (1962/1970) did not invent the term paradigm but he is largely responsible for how it is used today. A *paradigm* describes a collection of ideas—theories, concepts, laws, and experimental methods—that are shared by a community of scientists as a basis for describing the current status of knowledge about their domain and plans to advance knowledge within that domain. Although Kuhn initially talked about "pre-paradigmatic science," he later changed his

view and proposed that at all stages in the development of a field there are paradigms around which scientific communities cohere. For Kuhn, the history of science is described by the changing status of different paradigms, including periods of more incremental research ("normal science"), where scientists work to develop the ideas within a paradigm, as well as more radical events (paradigm shifts), where a previously dominant paradigm crumbles and is replaced by another. Paradigm shifts happen when results that do not fit within the existing approach (so-called anomalies) accumulate and the new paradigm better addresses the anomalies and is consistent with much of the previous data accounted for under the earlier view. Experiments which identify key anomalies can be viewed as critical ones, although attribution of such significance is always easier in hindsight. It is the nature of scientific paradigms that ideas in one do not necessarily translate into another. For example, Einstein's theory of relativity reformulates fundamental concepts in physics in ways that cannot be fully expressed in terms of classical Newtonian physics. As a result of the changed view that results from a paradigm shift, scientific understanding of the world is qualitatively different in some important ways from how it was previously.

Looking at the action-oriented view in cognitive science, it is difficult to argue that it constitutes one paradigm. Instead, we might say that there are a number of related paradigms which share some common emphases on the role of action. Do we have a paradigm shift? Twenty years ago, not many people thought motor activity could influence perception. Now there is a consensus that the motor system contributes to understanding. Perhaps the important point is that to make progress from this proposed paradigm shift, we must focus not only on high-level definitions but also on how these concepts will lead to new *experimental paradigms*,[1] both in the natural sciences and medicine as well as in the engineering sciences. To understand cognition, it has to be viewed in the context of meaningful behavior. Independently of whether the resulting dominant paradigm might be strong enactivism or a hybrid paradigm, the crucial point is to take action into account as a major constituent of cognition.

How Does the Pragmatic Turn Change How We Do Science?

A New Perspective on Perception

The implications of the pragmatic turn can be illustrated by referring to one of the most discussed versions of an action-oriented approach: the sensorimotor contingency (SMC) theory developed by O'Regan and Noë (2001). In this framework, the agent's acquired knowledge of SMCs (i.e., the rules governing sensory changes produced by motor actions) are critical for both development

---

[1] The term "experimental paradigm" refers to specific procedures and protocols and should not to be confused with the term "paradigm" defined above.

and maintenance of cognitive capacities. Vision, as an example, involves much more than the processing of retinal information. According to this view, seeing corresponds to a way of acting, to exploratory activity mediated by deployment of specific skills and knowledge of SMCs. Neural activity patterns that emerge in visual cortex do not in themselves constitute seeing. Rather, the brain supports vision by enabling exercise and mastery of SMCs. This concept is in strong opposition to classical representation-centered approaches of vision, such as those of Marr (1982) or Biederman (1987), which assume that perception consists essentially in building context-neutral descriptions of objects that are stored in the brain, irrespective of any actions that might be performed by the agent. In contrast, according to the sensorimotor account, the concept of an object would consist in a family of related SMCs, and thus clustering across these contingencies provides the basis of learning object structures. One of the interesting implications of this account is that the distinction between declarative and procedural knowledge, which figures prominently in cognitivist approaches, dissolves in favor of the notion that different types of knowledge might be seen as repertoires of contingencies with different degrees of complexity. The SMC approach also stipulates interesting predictions for changes in patients with brain disorders. A key prediction is that patients with motor deficits (e.g., Parkinson disease or amyotrophic lateral sclerosis) should exhibit perceptual and cognitive deficits that result from an impoverished repertoire of SMCs that can still be utilized by the patient.

Concept of Attention as Action-Constrained Sensory Refocusing

Research on neural mechanisms of attention serve as an example of how an action-oriented account may inform the analysis of presumed "high-level" cognitive capacities. The "premotor theory of attention" (Rizzolatti et al. 1987) introduced the idea that the selection of sensory information should be modulated and focused by constraints arising from current action planning and execution. In agreement with this prediction, several studies have shown that movement preparation can lead to attentional shifts (Collins et al. 2008, 2010) and to changes in the acquisition of object-related information (Craighero et al. 1999; Eimer and van Velzen 2006; Fagioli et al. 2007). Functional imaging studies and neurophysiological recordings have provided evidence that the modulatory bias imposed by attention may indeed arise from premotor regions, in particular the frontal eye fields (Donner et al. 2000; Moore and Fallah 2001). MEG experiments on visual attention have shown that premotor regions like the frontal eye field are involved in top-down modulation of sensory processing through selective enhancement of dynamic coupling, expressed by phase coherence of fast neuronal oscillations between premotor, parietal, and sensory regions (Siegel et al. 2008). Similar evidence has also been obtained in a recent EEG study using an ambiguous audiovisual stimulus. Analysis of neural

coherence revealed that large-scale interactions in a network of premotor, parietal, and temporal regions modulate the perception of the ambiguous stimulus (Hipp et al. 2011). Taken together, these findings provide clear support for an action-oriented account of attention. Actually, these studies suggest a shift in the concept of attention, which may indeed most appropriately be described as a bias in sensory processing that is procured by the current action context.

A New Perspective on Social Cognition

Different versions of the action-oriented view have had a major impact on social cognition research. Previously, social cognition research focused on explicit symbolic communication, on mind reading as theoretical inference, and on social categorization. One focus was on person perception, for example, in the mechanisms of face identification (Kanwisher et al. 1997), assigning traits to a person or categorization of others as in-group/out-group (Macrae and Bodenhausen 2000). Action-oriented views have brought social interaction into the foreground (Prinz 2012; Rizzolatti and Craighero 2004): simulationist models emphasize common coding and mirroring, two related representational and brain mechanisms for matching one's own actions and others' actions. This has sparked new experiments on imitation as well as the planning and execution of joint actions, such as carrying a table or playing a piano duet (Knoblich et al. 2011). Dynamic systems approaches highlight the importance of temporal entrainment in interpersonal action coordination (Riley et al. 2011). Enactivist approaches have brought to the foreground peoples' experiences during joint action (De Jaegher et al. 2010). Together, the different action-centered mechanisms have defined a new perspective on social cognition, and they have started to influence views of how individual cognition works (e.g., Pickering and Garrod 2013b).

Studying Social Interaction through Sensorimotor Entrainment

In classical representation-oriented approaches of social cognition, agents are thought to interact with conspecifics based on their capacity to develop a "theory of mind"; that is, to derive complex models of the intentions, beliefs, and personalities of other agents (Carruthers and Smith 1996). In this framework, which has also been termed the "spectator theory" of social cognition (Schilbach et al. 2013), the primary mode of interaction with the social environment is that of a detached observer who theorizes and produces inferences about other participants. The pragmatic turn inspires an alternative view which assumes that even complex modes of social interaction may be grounded in basic sensorimotor patterns enabling the dynamic coupling of agents (Di Paolo and De Jaegher 2012). Such sensorimotor patterns, or contingencies, are known to be highly relevant in cognition (O'Regan and Noë 2001; Engel

et al. 2013; Maye and Engel, this volume). A key hypothesis that follows from this approach is that learning and mastery of action-effect contingencies may also be critical for predicting consequences of the action of others and, thus, to enable effective coupling of agents in social contexts. According to this notion, social interaction depends on dynamic coupling of agents; the interaction dynamics provides a clue to social understanding and shares aspects with the interactionist concept of social cognition (Di Paolo and De Jaegher 2012; Gallagher 2004). This concept also agrees well with the joint action model by Knoblich and colleagues, who predict that shared intentionality can arise from joint action (Sebanz et al. 2006).

Action-Oriented Word Learning in Language Acquisition

The advantage of considering an enactive approach in cognitive development has been clearly demonstrated in the domain of child language acquisition.[2] Word learning has often been characterized in terms of the potential ambiguity when a word is uttered and the referent could be any one of a number of possible objects in the visual scene. In this characterization, the infant passively attends to the words and the cluttered scene. How would this scene change in an enactive setting? There is evidence that gestures (especially pointing gestures) by caregivers correlate with gestures by infants and, critically, that gestures produced by infants at 14 months predict vocabulary at later age (Özçalışkan et al. 2009; Rowe and Goldin-Meadow 2009). This indicates the importance of the action-oriented communicative context for language acquisition. In this context, Yu and Smith (2013) examined child-parent dyads during word learning, while the children wore head-mounted cameras that indicated where they were looking. The data revealed that children were not passive, but grasped and held the objects in question, creating conditions where referential ambiguity was eliminated, and word learning efficacy was maximized. This action-oriented shift in the study of language acquisition will likely find useful application as well in the context of the social dynamics that will similarly reduce referential ambiguity (Dominey and Dodane 2004).

Action-Centered Approach to Neuroscience

Peter König said: "If you only pose simple questions, you only get simple answers." In this spirit, we need to reformat questions about the nervous system as a complex system in an action-centered context that also involves extra-neural components. A reductionist approach misses elements of the system that were designed or which evolved to work in the real world. The problem is that mainstream neuroscience is stuck in the old model where subjects are often

---

2 Embodied language processing is another hallmark area where the action centeredness of cognition (in this case language) is primary.

passive or anesthetized. In classic vision and sensory experiments, at best, subjects report what they sensed passively; at worst, they are stimulated while anesthetized and unable to generate behavior that determines what they sense. Results from these open-loop experiments have been interpreted with static maps consisting of neurons that perform single, discrete functions. This points to another shortcoming of conventional neuroscience. It is as if each cell has one function and "causes" a change in a single output region. This is far from accurate. We now know that each cell encodes many parameters simultaneously and sends output to many places. These neurons "operate" continuously, not just when they receive input from a single source. The "causal" output (defined as the ability to change another neuron's probability of discharge) is typically very weak. It takes many neurons acting together to generate an action potential in another neuron. The "causal" chain is very noisy, can only be defined statistically, and is likely context dependent. It is difficult to decode information from a single neuron because its firing rate-parameter correlation is very weak; the change in discharge rate associated with any given parameter is small. However, across a population, if these small changes are consistent, the representation of the parameter value will emerge clearly. Describing causality in a complex system is difficult and can best be described as statistical structure. Common input to the population generates correlation between members of the population, and this structure can be recognized with proper analyses.

Changes in Conceptual and Methodological Orientation of Neuroscience Suggested by the Pragmatic Turn

A pragmatic turn in cognitive science will profoundly change our view of the brain and its function. Brains will no longer be considered only as devices for building models of the world (the "representationalist" view), but will instead be hypothesized to support the guidance of action. A key assumption is that action shapes brain structure and is constitutive for brain function. Importantly, this would hold not only during development, but also for the functioning of the adult brain.

These considerations suggest a refocusing of the conceptual premises of neuroscientific research. The following premises contribute to defining a framework for action-oriented, pragmatic neuroscience:

1. The primary concern of the experimenter is not the relation of neural activity patterns to stimuli, but to the action at hand and the situation in which the subject under study is currently engaged.
2. The functional roles of neural states might be viewed as supporting the capacity of structuring situations through action, rather than as "encoding" information about pregiven objects or events in the world.
3. Investigation of neural function encompasses the view that cognition is a highly active, selective, and constructive process.

4. Sensory processing should be considered in a holistic perspective, as being subject to strong top-down influences that continuously create predictions about forthcoming sensory events and eventually reflect constraints from current action.
5. The function of neurons and neural modules might not be considered in isolation, but with proper reference to other subsystems and the actions of the whole cognitive system.
6. Investigating the intrinsic dynamics of the brain becomes increasingly important, because interactions within and across neural assemblies are constitutive for the operations of the cognitive system.

Reorienting the focus of scientific investigation of cognitive processing will likely require the repertoire of neuroscientific methodology to be geared toward action-oriented experimental strategies. Some requirements of such methodological reorientation include:

1. Experiments must avoid studying passive subjects and, instead, allow for active exploration (e.g., free viewing, manual exploration).
2. Improved technologies are needed to track the actions of one or several interacting subjects.
3. Current neuroscientific methods (e.g., EEG, MEG, and fMRI) are limited with respect to subjects' ability to execute movements, either due to the design of the measurement apparatus or to artifacts being introduced into the signal by movement. Consequently, novel and improved technologies for acquiring neural activity and biosignals during movement are needed, as well as for analysis strategies that separate movement artifact from true signal.
4. Exploration of the relationship of neural activity with action parameters calls for the recording of nonneural, action-related signals. These may include, but are not restricted to, EMG, displacements, forces, heart and breathing rate, and possibly changes of the surroundings.
5. New data analysis techniques will be required to (a) examine correlations between behavior and high-dimensional neurophysiological signals and (b) develop methodologies that consider massively distributed coding both within and across brain structures.
6. Technologies allowing controlled manipulation of SMCs will likely become increasingly important. Virtual reality setups may become crucial experimental tools.

In summary, the pragmatic turn has had an important influence in changing the way that science is done. But what can we draw from these examples? From a strong enactivist view, it appears that as soon as time is critical for immediate interactions in a joint task, representational theories may have nothing to say. This suggests that within the framework of action there is a particularly highlighted status for interpersonal action. Interpersonal action is different, in

the relational aspect: the object of my interaction is also interacting, generating reciprocity in prediction when interacting with others.

Implications for Robotics and Engineering

In defining a shift in perspective on the status of agents acting in the world, the pragmatic turn has immense potential to change the fields of robotics and engineering. One could say that robots have always been embodied and pragmatic in the sense that they are physical devices designed to perform useful actions in the world. However, with respect to high-level tasks, roboticists adopted the representationalist stance which came from good old-fashioned artificial intelligence (GOFAI; Haugeland 1985). This yielded the classical sense-think-act control architectures which emphasized on the "think" part that involved building and updating world models and planning the next action using AI techniques operating on these models. Sensing and acting were initially regarded as straightforward interfaces with the real world and were thus considered less interesting or challenging research problems. A good example is the Stanford Cart (Moravec 1983). As a result, real-time responsiveness was lost.

Embodiment, Compliance, and Soft Robotics

The pragmatic turn goes hand in hand with an embodied turn. The properties of physical bodies were completely neglected in the strand of robotics that came out of the GOFAI tradition. Even in more recent and much more impressive examples that successfully interact with the world in real time, like the DARPA Grand Challenge winner autonomous car Stanley (Thrun et al. 2006), a clear separation between body and brain is apparent. Indeed, the design philosophy in the DARPA challenge was: "*treat autonomous navigation as a software problem*" (Thrun et al. 2006, their emphasis). This philosophy is in stark contrast with the embodied perspective that posits tight coupling between brain/controller, body, and environment (e.g., Pfeifer and Scheier 1999).

A classic and extreme example of the fact that behavior may be generated by a completely brain-less mechanical system are the passive dynamic walkers (McGeer 1990). Although such contraptions are highly dependent on their ecological niche (e.g., slope of particular inclination), to some extent "pure physics walking" can already display simple "adaptive" properties, like robustness to perturbations. This is achieved through mechanical feedback and has been called self-stabilization (e.g., Blickhan et al. 2007). Another ingenious demonstration from robotics where morphology decisively contributes to the generation of behavior is the "universal gripper" (Brown et al. 2010): A bag of ground coffee is pressed onto an object and let to conform to its shape. Afterward, a vacuum pump evacuates air from the gripper, making the granular material jam

and stabilizing the grasp. Another concerns the amazing climbing capabilities of geckos, which rely on van der Waals forces between their feet and the surface on which they are climbing; these are strong enough thanks to a hierarchical structure of compliance from centimeter to 500 nanometer scales (Autumn et al. 2002). These findings inspired the design of Stickybot, a robot that can climb smooth vertical surfaces (Kim et al. 2007).

Unlike "high-level" robotics capitalizing on the GOFAI (thus overlooking the importance of the body and closed-loop real-time interaction with the environment), control engineers would consider exactly these aspects as their "bread and butter." They would understand the robot as a dynamic system governed by differential equations. It responds to control input in a way that is defined by the body's structure and its interaction with the environment. Furthermore, adding a feedback loop allows shaping this response to achieve a desired behavior or goal state. The challenge includes identifying and modeling the dynamic system (i.e., finding the linear or nonlinear differential equations, deciding around which sensors to place the feedback loops, and designing the controller itself). For linear systems, a large body of powerful and well-understood mathematical tools exists to analyze the system's response and design appropriate controllers. However, the majority of real systems are governed by nonlinear dynamics. While a number of techniques exist to cope with these kinds of systems, they are usually far more complex to compute, apply only to a subset of systems, or approximate the nonlinear with linear dynamics.

As a result, the plant in a control system is typically treated as fixed, and the overall tendency has been to suppress its complex nonlinear dynamics in favor of stiff and linear behavior. Often this approach works surprisingly well. However, "linear systems are not rich enough to describe many commonly observed phenomena" (Sastry 1999:2). Complex and soft bodies offer new possibilities which can be exploited; however, they also pose difficulties for classic control approaches (see Hoffmann and Müller 2014).

Complementary to *passive compliance*,[3] there is an interest in *active compliance* where a controller mimics elasticity using otherwise stiff actuators. Although these kinds of systems are not as energy efficient as passive systems, they offer the possibility to study variable stiffness as well as to study where it is beneficial to introduce compliance in a system. The importance of compliance has been demonstrated for robotic locomotion (e.g., Kalakrishnan et al. 2010; Semini et al. 2013) and manipulation (e.g., Righetti et al. 2014; Deimel et al. 2013) where it increases robustness against perturbations that may be due to inaccuracies and noise in perception and actuation. Instead of precisely planning a movement in joint space, a feedback loop is closed around the

---

[3] Compliance is the opposite of stiffness. A stiff system can move along a desired trajectory or toward a goal position and will remain in this state no matter what external forces are applied to it. A compliant system allows deviations from these desired positions that may be introduced by external forces.

interaction forces between robot and environment such that it can gracefully give in when experiencing unexpected contact with the environment. These challenges are taken up by the growing field of "soft robotics" (e.g., Albu-Schäffer et al. 2008; Pfeifer et al. 2012; Trimmer 2013).

Action in Robotic Perception

One of the interesting claims of the enactive approach is that there may be no principle difference between "high-level" cognitive processes and "low-level" sensorimotor functions. Instead, the former is seen as grounded in and emerging from the latter. Thus perception cannot be treated as a passive and disembodied process: it is critically shaped by the morphological properties of the whole agent (including, of course, its sensory apparatus) as well as by the actions executed by the agent. This is in contrast with the prevailing approaches to robotic perception, vision in particular (e.g., Horn 1986).

Object concepts and object recognition may be considered to illustrate this. From an enactive perspective, object concepts do not consist of feature-based descriptions stored in memory agnostic to any action contexts. Objects would be defined by sets of possible actions that can be performed on them, and knowledge of an object concept would consist in the mastery of the relevant object-related SMCs (Engel et al. 2013; Maye and Engel, this volume). This view is supported by studies on object recognition in humans and robotic systems. Evidence in humans clearly demonstrates a dependence of exploratory eye movements on the specific task given prior to viewing an image (Yarbus 1967). The influence of semantic information, also in terms of object identity, is visible in the results. Furthermore, there is evidence that visual object recognition depends on exploratory eye movements during free viewing of images. Using ambiguous images, a recent study showed that eye movements performed prior to conscious object recognition predict the object identity recognized later (Kietzmann et al. 2011). For artificial vision systems, these ideas were first explored in the area of active vision and perception (Ballard 1991; Bajcsy 1988). They focused on implementing systems both in hardware and software that have the ability to choose actively what to sense. More recently, the area of interactive perception goes beyond the mere selection of percepts toward actively changing the state of the environment to increase information gain. In studies on robot vision, it has been shown that a visual scene can be disambiguated by actively manipulating objects in the scene through a robotic arm (Fitzpatrick and Metta 2002; Björkman et al. 2013; Högman et al. 2013). Similar implications apply to other sensory modalities. For example, a quadruped robot running on different ground is critically able to improve terrain discrimination if a history of the actions taken (gaits used) is considered together with sensory stimulations induced in tactile, proprioceptive, and inertial sensors (Hoffmann et al. 2012).

A Holistic Distributed Approach to Control: An Action-Oriented Systems Engineering Perspective

The representationalist view of cognition has some commonalities with traditional systems engineering[4] in the sense that it heavily relies on top-down design and modularization. Processing is often central and sequential, and different modules are recruited to perform their part in a pipeline-like manner (like sense-think-act). This approach is dominant in the engineering disciplines, because it allows for efficient separation of subtasks and distribution of work among different units.

However, as has already been argued extensively, humans and animals do not seem to rely on the same type of architecture. Instead, an effective and smooth interaction with their environment emerges from the interplay of a plethora of physical and informational processes that operate in parallel and have different couplings between them. This was articulated by Rodney Brooks when he openly attacked the GOFAI position in seminal papers: "Intelligence without representation" (Brooks 1991b) and "Intelligence without reason" (Brooks 1991a). Through building robots that interact with the real world, such as insect robots (Brooks 1989), Brooks realized that "when we examine very simple level intelligence we find that explicit representations and models of the world simply get in the way. It turns out to be better to use the world as its own model" (Brooks 1991b:396). Inspired by biological evolution, Brooks created a decentralized control architecture consisting of different layers: every layer a more or less simple coupling of sensors to motors. The levels operated in parallel, but were built in a hierarchy, hence the term subsumption architecture (Brooks 1986).

This approach proved to be effective in "low-level" tasks, such as walking or obstacle avoidance. However, to our knowledge, it has not demonstrated, to date, how it could scale to more "cognitive" tasks, beyond the immediate "here-and-now" timescale of the agent. Growing evidence from biology suggests that this approach has been adopted by humans and animals and thus can be scaled up. Therefore, if we take the enactive approach seriously, we need to revise the way robots are designed. Here, we elaborate on a few points toward this goal:

1. *From centralized modules to parallel, loosely coupled processes*: "Intelligence is emergent from an agent-environment interaction based on a large number of parallel, loosely coupled processes that run asynchronously and are connected to the agent's sensorimotor apparatus."

---

[4]   Systems engineering was initially developed in the context of complex spacecraft system design (including the spacecraft and the associated ground data systems). Our perspective here is to in no way call into question the validity of this methodology, but rather to recognize that it was developed for a class of systems which do not at all have the same interaction requirements that one finds in real-time social interaction.

(Pfeifer and Scheier 1999:303). This contrasts with the classical approach consisting of centralized, functional modules operating sequentially. Furthermore, unlike the classical scheme, all processes operate essentially in a closed loop through the environment and on different timescales. In addition, as we are learning from the brain, process specialization is much weaker than representationalist views of the mind envisioned: different circuits are dynamically recruited for different tasks when need arises. This insight should flow into action-oriented and embodied design methodologies. Rather than starting out with a modularization of the system into functional blocks, it may be more beneficial to begin thinking about the interfaces. Modules (in a weaker sense) can subsequently be plugged in. This may have a profound influence on how the modules are designed by the end of the process. An even stronger version of this methodology would first be to design entire feedback loops. This would shift the focus away from modules toward combinations of different feedback loops.

2. *Timing is important*: Real-time interaction with the world and with other agents is critical. For example, in human-robot interaction, we are still very far from the natural "fluid" interaction that humans have among themselves. A similar observation can be made for physical contact interactions between the robot and its environment. When it comes to manipulation or locomotion in unstructured environments, the abilities of robots lag far behind those of humans in terms of dexterity, robustness, and speed. The low-level bodily and sensorimotor levels with fast feedback loops as well as their combination, integration, and coupling should play an important role in solving this lack of fluidity. We realize that for temporal coordination (either during verbal interaction with a person or physical interaction with the environment), underlying processes cannot be sequential and feedforward but have to be parallel, asynchronous, and predictive.

3. *Self-organized task solving*: Imagine the task of cleaning a table. If a household robot were to solve it in the classical approach, it would perform image segmentation first and then try to identify all the objects in the scene. Based on a model acquired this way, it could set itself a goal state (e.g., clear the table) and then apply search techniques from AI to obtain a complete plan on how to proceed. Intuitively, it is evident that this is not how a human would solve such a task: a human would most likely remove objects within reach without fully planning out the remaining steps. In the end, paraphrasing Brooks, the table is there and one does not need a model to clean it. Thus, seemingly complex problems can often be greatly simplified through sensorimotor coordination and by off-loading complex planning to the brain-body-environment interaction.

In summary, we are convinced that the embodied, action-oriented, or enactive approach has important consequences for the way robots and their control systems are designed. However, for some of the implications—in particular, to move away from specialized, clearly separated modules—the path to their adoption by industry may not be easy.

Practical Applications of the Action-Oriented Approach

We have come to be convinced of the potential impact of this change in perspective. Clearly, such a change in perspective must generate practical applications. These practical applications can be seen in domains that include education, clinical therapy, and the development of enactive prosthetic devices.

Education

Historically, the pragmatic approach has had a significant impact on education. While we will not treat this subject in detail, we note that John Dewey, in particular, had a profound commitment to and influence on public education. Dewey (1897, 1900, 1902, 1916, 1938) advocated the importance of education proceeding in a context of interaction, where curiosity and discovery motivated the student's active search for knowledge. Similarly, action-oriented education would emphasize the importance of participatory aspects of learning situations. Instead of "sitting in place" in a classroom, listening or working with books or even desktop technologies, action-oriented education would emphasize the use of one's whole body for learning. In contrast to conceptions of learning and cognition modeled on information processing and amodal problem solving, for example, the use of computer-generated interactive simulations (mixed reality immersive technologies)—where students can enter into a whole-body engagement with a subject matter—has been shown to increase learning speed and accuracy compared to desktop computer use (Lindgren and Moshell 2011; Lindgren and Johnson-Glenberg 2013).

Clinical Applications in Aphasia

Research in action-centered cognitive neuroscience has led to the development of a new translational method for the treatment of language deficits, or aphasias. This method, called intensive language action therapy, has yielded significant improvements in language and communication abilities, even at chronic stages of poststroke aphasia, in contrast to conventional approaches, which have not demonstrated comparable effects (Berthier and Pulvermüller 2011). Further translational progress may be achieved by developing similar methods for other neurological deficits affecting language and action.

Enactive Approaches to Developmental Disorders

Sensorimotor problems can be found in a number of developmental disorders. In autism spectrum disorder (ASD), for example, infants who are later diagnosed with autism display sensorimotor problems (e.g., postural stability, gait, timing and coordination of motor sequences, anticipatory adjustments and face expression) before they reach the developmental age associated with theory of mind, when typically developing children engage in joint attention and joint action with others and are learning to communicate (Trevarthen and Delafield-Butt 2013; Gallese et al. 2013; Cattaneo et al. 2007; Cook et al. 2013; Gallagher 2004; Hilton et al. 2012; Fabbri-Destro et al. 2009; Whyatt and Craig 2013; David et al. 2014). In individuals with ASD, Torres and colleagues show the occurrence of disrupted patterns in re-entrant (afferent, proprioceptive) sensory feedback that usually contributes to the autonomous regulation and coordination of motor output, and supports volitional control and fluid, flexible transitions between intentional and spontaneous behaviors (Torres 2013; Torres et al. 2013). In ASD, as well as in other developmental disorders (e.g., Down Syndrome), disruptions in motor processes may partly explain why individuals show difficulties in distinguishing goal-directed from goal-less movement (Torres 2013; Brincker and Torres 2013), anticipating the consequences of their own impending movements and applying fine-tuned discriminations to the actions and emotional facial expressions of others during real-time social interactions. These studies hold important implications for future research and therapeutic interventions, although further research is required to understand differences among sensorimotor problems in the different disorders and what precisely they contribute to each one (Gallagher and Varga 2015).

Action-Oriented Approaches to Schizophrenia

The early work by Elaine Walker and colleagues (e.g., Walker and Lewine 1990) has motivated investigations of motor activity in people with schizophrenia. Ford and coworkers showed that 2-year-old children, who became schizophrenic in adulthood, differed from their healthy siblings by exhibiting motor awkwardness and social withdrawal. Ford argues that this could be due to an inaccurate, slow, or faulty forward model of motor control starting in infancy. Using EEG-based methods that allow excellent temporal assessment of neural processes, they found abnormal premotor activity in patients with schizophrenia in the ~100 ms preceding talking (Ford et al. 2007), with abnormalities being greater in patients with more severe auditory hallucinations and worse amotivation and avolition, respectively.

Several studies have investigated action-perception loops in schizophrenia with particular reference to the understanding of delusions of control, and the belief that one's actions are controlled by another agent (C. D. Frith 2012; Blakemore, Smith et al. 2000; Frith et al. 2000a). The idea stems from

Helmholtz who pointed out that when we move our eyes, the world remains stationary, but if we poke our eye with our finger, the world appears to move. The normal attenuation of movement depends on relating the intention to move (or speak) with the anticipated sensory changes (the forward model). This is why we are unable to tickle ourselves. Schizophrenic patients with delusion of control can tickle themselves, presumably because something has gone wrong with this perception-action loop. The same framework has been applied to other disorders in the experience and production of action.

## Exploiting the Link between Movement Disorders and Cognitive Disorders

Based on the foregoing examples and related data, a view has emerged that movement disorders and cognitive disorders are linked. Thus, a framework that considers the crossover of action and cognition would be useful. For example, in schizophrenia and depression, disturbed notions of agency are evident in individuals with these syndromes, but movement components are considered less frequently. In Parkinson disease, one finds the opposite: a bias toward movement.

The application of principles that are consistent with the enactive agenda can be found in the domain of neurorehabilitation. One example is the rehabilitation gaming system (RGS), which exploits virtual reality-based interaction to incorporate embodiment, first-person perspective, and goal-oriented action in rehabilitation protocols. RGS has been used by over 400 stroke patients in controlled clinical impact studies, and significant impact on functional recovery has been demonstrated (Cameirão et al. 2011, 2012). In a virtual reality therapy setup, Cameirão et al. have demonstrated the increased effectiveness of embodied versus disembodied therapy, noting the particular effectiveness of embodiment with a first-person perspective, and that goal-directed action is more effective than repetitive action. Likewise, effective results are also seen when patients are placed in a social context (e.g., working with others), thus improving self-image.

Neural Prosthetics

Neural prosthetics offers a clear example of action-based learning and control (Schwartz 2004). In the scheme illustrated in Figure 20.2, we consider brain operation to be the creation of behavior and movement to be behavioral output. The cost function is generation of the desired movement.

One aspect of this approach is the need to map each neuron's firing rate to movement. With intact subjects, this is relatively straightforward; movement parameters (e.g., arm direction) can be regressed against firing rate to capture a tuning function for each unit. This is more difficult with a paralyzed subject. Schwartz and colleagues have developed a "calibration" procedure
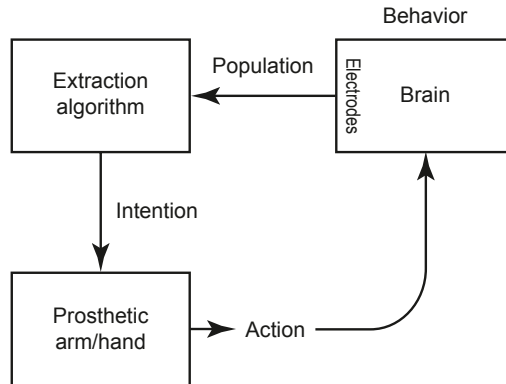
Figure 20.2   Neural prosthetic scheme: Electrodes in the brain record single-unit activity. Many units are recorded simultaneously and this population is processed with an extraction algorithm. This algorithm is a model of the transformation between neural activity and movement intention. The model generates an estimate of the intended movement signal, and this serves as the control signal for an external device (here, a prosthetic replica of the arm and hand). The prosthetic device moves according to the estimated intention signal, and its action is visualized by the subject who registers any errors in the desired movement. The error drives a learning process in the brain which modifies the neural output to reduce the error signal. Essentially, the brain is learning the extraction algorithm. This closed loop induces powerful learning, leading to high prosthetic performance.

based on observation-driven cortical activity (Velliste et al. 2008; Tkach et al. 2008; Collinger et al. 2012). By demonstrating different movements to the subject who is actively monitoring the motion of the prosthetic device, neural activity in the motor cortex is elicited that can be used to estimate an initial tuning function. After an initial estimate, the process is iterated; the subject's activity is gradually used to control the device by mixing their volitional signal with a set of parameters that attenuate errors which would move the device away from the displayed target. The amount of assistance (error attenuation) is gradually reduced while the proportion of volitional drive is increased until the control signal is composed only of the subject's extracted neural output.

The method of mixing volitional and autonomous control is called "shared mode" control (Clanton et al. 2013). For instance, suppose a paralyzed subject is operating the arm/hand to reach, grasp, and manipulate an object. Advanced robotic algorithms have been developed to generate autonomous manipulation (Katz et al. 2013), and these autonomous signals can also be mixed with those derived from brain activity. This could be used to test the idea of agency (e.g., by querying the subjects as to whether they were controlling the device, or if it was being controlled automatically as different ratios of automatic/volitional signals were mixed together within a neural prosthetic experiment). An open question is how the conscious experience of an artificial limb as one's own

adds to the (motor) abilities of the agent over and above tool use, in which the tool is usually experienced as clearly separate from the body (e.g., a hammer, a fork). Ehrsson et al. (2004) suggest that parietal regions mediate multisensory matching of the rubber hand, but that premotor cortex is related to the conscious experience of embodiment. In contrast, current work (Seth, pers. comm.) is trying to identify how action may determine whether an object is perceived as embodied.

These experiments also show that single neurons encode multiple parameters in their firing rate patterns. For instance, many neurons in the motor cortices of human and nonhuman primates have been found to contain neurons with tuning functions describing 10 degrees of freedom for hand and arm movement (Wodlinger et al. 2015). These tuning functions are not clustered in parameter space but seem to be distributed uniformly. Furthermore, neurons with different tuning functions are located together anatomically in the same region of cortex. Finally, neurons with the same type of tuning function (cosine shaped) for movement can be found throughout the neural axis (Van Hemmen and Schwartz 2008). This suggests that the principle of directional tuning is widely distributed and not localized to specific structures.

Societal Implications: How Can We Connect to Everyday Human Life?

A central goal of enactivist theory as proposed by Varela et al. (1992) was to address what it means to be human in everyday life, the ultimate aim being to address existential and ethical questions about the human condition: Who am I? What am I? How can I live a good life? Following Merleau-Ponty, the human body can be viewed as both a physical structure and as a "lived-in" phenomenological one. Thus, to move from a science of embodied cognition to an enactive understanding of human nature, it is necessary to bridge this gap. Varela et al. thought that progress in this direction could be made by connecting ideas which emanate from cognitive science—embodiment, distributed cognition (e.g., connectionism and Minsky's "society of mind"), and self-organization and emergence—with insights external to science (e.g., from traditions such as Buddhist views on the groundless nature of the self). How does this enterprise to apply action-oriented cognitive science to the challenge of the human condition stand today?

Within contemporary action-oriented views, we discern a number of positions. For some, significant progress toward addressing existential issues can be identified through recent advances in our understanding of consciousness and the self (see Seth et al., this volume). By translating these theoretical insights into more everyday conceptual language, action-oriented cognitive science could play a more direct role in, for instance, the ongoing debate about the relationship between science and religion. For others, the action-oriented

approach, despite its focus on subjectivity, is seen as providing only third-person insights into consciousness and personhood, which are then of no immediate help to understanding the first-person human predicament. From this perspective, action-oriented cognitive science can only impact indirectly on existential considerations through the filter of nonscientific frameworks. Although the range of views on these issues is broad, there is general agreement that through its emphasis on embodiedness, extended cognition, and intersubjectivity, the action-oriented approach opens up new pathways for dialogue between the biological and psychological sciences, on one hand, and the social sciences and humanities, on the other.

Availability of Phenomenological Expertise

That phenomenology can contribute to cognitive science is a relatively new notion (Gallagher and Varela 2003). If the task of cognitive science is to explain how we experience, engage, and interact with the world, it is claimed that phenomenology provides a controlled descriptive method that is able to characterize the explanandum, without necessarily specifying the cognitive mechanism that explains it. This is not so much a division of labor—since the two tasks (description and explanation) depend on each other—but rather more a case of mutual enlightenment. This raises the following questions: Even if all people have access to experience, are they all equally good at describing it? Do we need experts in phenomenology to provide such descriptions? Do such descriptions differ between cultures?

While it is likely that there are clearly some aspects of variability related to culture and experience, we may ask whether there are some universal invariants. If "experts" perceive things that novices cannot perceive, as numerous studies in athletics show (see, e.g., Mann et al. 2007), this suggests that the depth and precision of access to experience can be increased through practice, and that a new vocabulary can be developed to allow precise communication. Thus, for example, Buddhists who practice mindfulness mediation have a precise language for discussing detailed notions of phenomenological experience of different aspects of self (Trungpa 2002). Phenomenology, as well as Buddhist mindfulness, may offer important tools to allow us to develop insight into our everyday experiences.

The claim here is not that phenomenology is primary. Rather, according to what Varela (1996) calls the methodological principle of "mutual constraints," phenomenology and cognitive science should maintain some consistency, so that if there is some disagreement between them about some particular phenomenon or process, further research will be necessary to resolve the difference.

Phenomenology has informed enactivist approaches to cognition. Phenomenological philosophers have argued that world or environment is not independent of mind or organism. We need to think of the organism-environment (or the mind-world) as an epistemic system. The dynamic systems

perspective of enactive thinking provides descriptions of agent-environment systems in which basic processes, including problems and pathologies, are described in relational terms: where problems involve some tension or conflict in the relation between the agent and environment. Brain-centered approaches tend to identify these problems entirely as occurring within the organism. An enactive approach would spread this out to examine the relations between brain, body, and environment and consider treatments in terms of those relations. Remedies could then be achieved by manipulating or intervening with the brain, the body, the environment, or some mix of factors. We note that many domains of practical application already use such a systems view without it being inspired by enactive thinking per se. Enactive researchers would like to push such approaches into much wider use than is currently the case.

Enactive Autonomy and Embodied Psychosocial Existence

As an example of how the enactive view could be applied to everyday life, we conclude this section by summarizing Kyselo's theory of the human self as embodied psychosocial existence (Kyselo 2014), as it incorporates some key principles of the broader action-oriented approach. According to Kyselo, the enactive self can be operationally defined as a socially enacted autonomous system, whose systemic network identity emerges as a result of an ongoing engagement in processes of *distinction* (which promotes the existence of the individual in his/her own right) and *participation* (which promotes connectedness with others). Some implications for understanding and improving everyday human life are as follows.

This view emphasizes that human identity depends on others. The enactive view presupposes a deep dynamic interrelation between agent and world—a relation of mutual co-constitution. For the self, this means that the social world is not merely a context or developmentally relevant, but that without continuously engaging with the social world of other people, without their contribution, no individual self can be generated or maintained. Human nature is not egocentric but genuinely relational. We are not embodied islands but existentially dependent on each other as long as we live. Further, because we care for the maintenance of our identities, we continuously and adaptively evaluate ourselves and our interactions with others accordingly.

Self-maintenance involves tension and vulnerability. Cognitive identity is generated and maintained under precarious circumstances. For the human self, this means that both kinds of network processes—those that enable distinction as well as those that enable participation—are required together to bring about the individual as a network of autonomous self-other organization. Without distinction, the individual risks dissolving in social interaction dynamics. Without participation—acts of openness toward others—the individual eschews structural renewal, risking isolation. Both goals are in opposition, bringing about a

deep tension for the self that has to be negotiated. A useful insight for everyday life is that this tension is a necessity that cannot be avoided. Vulnerability and openness to others are appreciated as enabling the second dimension of self: the sense of being connected.

Human identity and understanding others requires continuous negotiation. Tension and conflict with others are to some extent unavoidable, since each individual is engaged in an interaction that strives to maintain this twofold sense of self. When two agents interact, their goal is not merely to reach consensus and harmony, but to reach it under particular conditions; namely, by acknowledging that their respective individual goals are met. Combining these ideas with dynamic systems theory, Kyselo and Tschacher (2014) propose an approach for understanding relationship dynamics and the negotiation of closeness and distance within couples, and what this means for personal well-being and the likelihood of relationship maintenance.

Finally, the twofold structure of the self, as distinct and participating, entails a further useful insight: individuals are limited in their personal control over their self-constitution. Other people have a say in the construction of our identities and, given that they have perspectives and interests of their own, others do not always comply with what we would like or need. The enactive view contrasts with the increasing emphasis on individualism in Western society—that we should seek to be the omnipotent creators of our own lives—and suggests that a better understanding of the role of others in constructing the self could help people to be more at ease with themselves.

Concluding Remarks

An action-oriented perspective changes the role of an individual from a passive observer to an actively engaged agent interacting in a closed loop with the world as well as with others. Cognition exists to serve action within a landscape that contains both. Here we have offered an overview of this landscape, where action is not just the output of the system but also where the system is there for action.

This perspective emphasizes a continuum of interaction between biology, psychology, the social sciences, and the humanities and has already had an impact on science by changing the way we consider perception, social cognition, and interaction, and the bases of their neurophysiological implementation. Other impacts can also be expected. For example, approaches in engineering need to change to what we refer to as action-centered systems engineering. Social implications are equally visible. From an action-oriented perspective, a human being is not an isolated individual responsible alone for his/her destiny, but rather a member of a grounded cultural system. The true test of these proposed implications cannot be fully evaluated today but will require the test of time to determine if we have seen clearly into the future.

Acknowledgments

We would like to thank our distinguished colleagues from the Forum for providing additional input to the chapter, including Judith M. Ford, Jürgen Jost, Miriam Kyselo, Friedemann Pulvermüller, Gottfried Vosgerau, Peter König, Paul Verschure, Marek McGann, Chris Frith, and Gabriella Vigliocco.