2

The Contribution of Pragmatic Skills to Cognition and Its Development

Common Perspectives and Disagreements

Giovanni Pezzulo

Abstract

From both an evolutionary and a developmental perspective, sensorimotor skills precede "higher" cognitive abilities. According to traditional cognitive theories, abilities such as thinking, planning, and mind reading are not action dependent. Action-based theories, however, do not view action and cognition as separate domains; they argue that pragmatic skills (or a "mastery of sensorimotor contingencies") are integral components of higher cognitive abilities, essential for their development. For example, pragmatic skills might afford (active) perception, (active) learning, and the acquisition of conceptual knowledge as well as "intellectual" skills (e.g., thinking or calculating). Despite accumulating empirical support for action-based theories, it is unclear to which extent pragmatic skills contribute to cognition and its development, with contrasting proposals in the field. This chapter reviews three (not mutually exclusive) perspectives: (a) the coordinated self-organization of behavior and cognition; (b) the role of "cognitive mediators" across sensorimotor and higher cognitive domains; and (c) the action-based construction of abstract and amodal cognitive domains. Common perspectives and disagreements between these views are discussed, and open issues, opportunities for theoretical debates, and empirical tests are highlighted, all of which might contribute to a research agenda for the emerging "pragmatic" view of cognition.

Introduction

How agents—living organisms or robots—develop "higher" cognitive abilities based on existing sensorimotor skills is a topic of debate in many disciplines (e.g., cognitive science, psychology, neuroscience, philosophy, and robotics).

From "The Pragmatic Turn: Toward Action-Oriented Views in Cognitive Science,"
Andreas K. Engel, Karl J. Friston, and Danica Kragic, eds. 2016. Strüngmann Forum Reports, vol. 18,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03432-6.

Contrasting proposals have emerged to describe sensorimotor and higher cognitive domains as separate, integrated, or coextensive.

Action-based theories do not regard action and cognition as separate domains. They argue that action is inherent to cognition and its development. Accordingly, the primary role of cognition is to support (and enhance) action-control abilities rather than to produce "encyclopedic" knowledge that is detached from action and perception systems. All cognitive operations usually mentioned in psychology textbooks (e.g., perception, memory, reasoning) are organized around—and ultimately functional to—the demands of action control and goal achievement:

- Perception and learning consist of picking up regularities in the information flow as structured by my actions (O'Regan and Noë 2001; Pfeifer and Scheier 1999).
- Attention is understood as selection-for-action or focusing on the regularities that are useful for the current action demands (Allport 1987; Ballard 1991).
- Conceptualization and memory consist of "encoding patterns of possible physical interaction and reenacting them in the service of the current interaction or to select the next one" (Glenberg 1997) and thus even memory is action dependent because "it is only when patterns of sensorimotor experience have been structured that I can memorize them" (Verschure et al. 2003).
- Cognitive development is "scaffolded by action development" (Byrge et al. 2014; Piaget 1954; von Hofsten 2004).
- Decision making is not independent of action processes; rather, action performance should be considered as a proper part of a decision process, not merely as a means to report the decision (Lepora and Pezzulo 2015).

Despite recent progress, it is unclear if, and how, this new "pragmatic" view extends to the domains of higher cognition.

A common starting point of action-based theories is that higher cognition is not modularized—as in the famous "cognitive sandwich metaphor" which segregates perception, cognition, and action domains—but is instead deeply integrated with, or even dependent on, action-perception systems. Support for this view comes from evolutionary arguments, which hold that the architecture of cognition can be traced back to the sensorimotor architecture of our earlier evolutionary ancestors. The brain's architecture developed to meet the needs of interactive behavior not cognitive abilities, such as playing chess or doing complex exercises in logic and mathematics; the demands of situated action control might also have somehow bootstrapped and shaped higher cognition (Cisek and Kalaska 2010; Pezzulo and Castelfranchi 2009). As a consequence, cognition is better described as a set of adaptive skills that exist *in continuity with* action-control mechanisms; they do not form a separate, modularized domain.

From "The Pragmatic Turn: Toward Action-Oriented Views in Cognitive Science,"
Andreas K. Engel, Karl J. Friston, and Danica Kragic, eds. 2016. Strüngmann Forum Reports, vol. 18,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03432-6.

A similar argument might hold at a developmental timescale. A recurrent theme within action-based theories is that during development, the acquisition of pragmatic skills guides the acquisition of cognitive abilities. In other words, the driving force of development is the acquisition of practical skills (e.g., learning to control one's own actions and to achieve more complex goals; learning to predict longer-range consequences of actions), not the direct acquisition of cognitive skills: the former provides a scaffold for the latter (see Pezzulo et al., this volume). It has been demonstrated that action structures the perceptual domain (Gibson 1979; O'Regan and Noë 2001), the memory domain (Verschure et al. 2003), and the ability to exert proactive behavior, among others (von Hofsten 2004).

The theory of *sensorimotor contingencies* (SMCs) (O'Regan and Noë 2001) offers one possible explanation. Broadly speaking, learning sensorimotor skills means learning to exploit systematic relations between actions and (changes in) the world and leads to a *mastery* of SMCs. In action-based theories, which have both cognitive and enactivist flavors, such systematic relations are key to both action control and cognition (Clark 1998; Maturana and Varela 1980). Under cognitive views, systematic relations must first be internally represented (e.g., as internal generative or forward models) before they can be successively reused during overt sensorimotor interactions or covert tasks, such as imagery. According to an enactivist view, systematic relations are directly enacted to engage in sensorimotor interactions without being internally represented.

Sidestepping this dispute, skilled action control has *epistemic* effects. As clearly recognized by ecological psychologists, sensations help to determine actions but, in the process, they also create new sensations, thus creating a continuous action-perception loop (Gibson 1979). This implies that agents not only change the world by acting (a pragmatic effect), they also unveil and inject information into the world through their actions (an epistemic effect) and, in doing so, they create structure in the sensorimotor flux. The information and structure created by acting can be exploited to steer *active* forms of perception, cognition, and learning. Consider, for example, *optic flow*: only through action are the stimuli required to recognize, for example, the shape, distance, or movements of objects created. A tenet of SMC theory is that action is integral to perception and that sensory experience is an active mode of exploration (O'Regan and Noë 2001), and thus a sole property of active agents. Support for the importance of action for perception comes from studies which show that the development of the latter is impaired in animals that only experience the world passively (Held and Hein 1963). Cognitive robotic experiments have further assessed the importance of sensorimotor engagement for various "active" (i.e., action-mediated) strategies, such as active vision or active learning, where perception and learning depend on the robot's ability to select its next stimuli by acting (Pfeifer and Scheier 1999; Verschure et al. 2003).

Recognizing that actions have epistemic effects is necessary to link the domains of skills and sensorimotor control to knowledge and cognition, which are separated in traditional cognitive theories. In principle, one might argue that action is not only inherent to perception (*no action, no perception*), it is also essential for all domains of cognition, including learning, memory, concept formation, and beyond (*no action, no learning, no concepts, no cognition*). Accordingly, the domain of "action" can also be extended to include mental operations or "intellectual skills," such as *thinking or calculating* (Engel et al. 2013; Rosenbaum et al. 2001). However, mechanistic theories of the relations between sensorimotor and higher cognitive skills are at best largely incomplete, especially in terms of developmental aspects. How might "pragmatic skills" support higher cognitive abilities and/or bootstrap them during development?

Pragmatic Skills and Cognitive Development

The idea that sensorimotor development promotes cognition was pioneered by Piaget (1954). More recently, several working examples have emerged in embodied cognitive science (Barsalou 2008; Byrge et al. 2014; Thelen et al. 2001) and robotics (Nolfi and Floreano 2001; Pfeifer and Scheier 1999; Verschure et al. 2003). However, contrasting proposals exist, which I classify as follows:

1. The *emergentist* perspective emphasizes the self-organization of increasingly more complex (inter)action patterns during development.
2. *Cognitive mediation* stresses that certain abilities, developed initially for the demands of situated action, are mediators of higher cognitive abilities. For example, prediction (or other information-processing mechanisms) can be reused and adapted from the domain of action control to novel, more cognitive domains.
3. The *abstract-and-amodal* perspective emphasizes that action-based processes help develop cognitive abilities, but that once established they become autonomous from perceptual and motor systems, and are thus abstract and amodal domains.

These perspectives are not mutually exclusive and do not clearly cover the full spectrum of possible views. However, they do exemplify the potentialities of action-based approaches and point to current controversies that are ripe for debate.

The Coordinated Self-Organization of Behavior and Cognition

According to an emergentist or interactionist view, cognitive development is an incremental process of self-organization in which the learned products of a given agent-environment interaction (e.g., a grasping skill) can be exploited to acquire increasingly more complex skills (e.g., a reaching-and-grasping skill). For example, robotic experiments show that robots endowed with quite generic *fitness functions* are able to generate an incremental repertoire of behaviors:

From "The Pragmatic Turn: Toward Action-Oriented Views in Cognitive Science,"
Andreas K. Engel, Karl J. Friston, and Danica Kragic, eds. 2016. Strüngmann Forum Reports, vol. 18,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03432-6.

they first acquire low-level behaviors (e.g., locomote, avoid obstacles) and are able successively to acquire higher-level behaviors (e.g., push objects) by integrating and recombining lower-level behaviors (Martius et al. 2013; Nolfi 2009; Verschure and Pfeifer 1993). By acquiring such behaviors, the robots automatically acquire some "cognitive" abilities, at least for some simple forms of cognition. An example of how this is possible is offered by an evolutionary robotics experiment: robots constituted of simple feedforward neural networks learned to "recognize" or "discriminate" objects of different shapes (a wall vs. a cylinder) by producing different behavioral patterns as they interacted with the objects (e.g., linear vs. back-and-forth trajectories) but without using internal states that "represent" the categories (Morlino et al. 2015; Nolfi 2009). In this example, the ability to categorize emerged through situated interaction without corresponding "internal states" or specialized cognitive processes (e.g., the comparison with exemplars of the category or the accumulation of evidence in favor of the perceptual alternatives), as would be more typical in cognitive modeling. This example shows that in naturalistic tasks, "categorization" does not necessarily correspond to an explicit cognitive operation but can be an intrinsic aspect of a successful interaction with objects. This might be different during psychology experiments where subjects are explicitly asked to report a category name. Indeed, further studies suggest that more sophisticated and "explicit" forms of categorization can emerge if the task demands are appropriate. For example, in another evolutionary robotics study, where several robots were able to communicate (with auditory beeps), a more discrete category-recognition ability emerged in which the robot communicated a discrete state (e.g., I am or I am not close to a cylinder) (Nolfi 2009). Other experiments have assessed whether adaptive agents can store the results of previous interactions in memory (e.g., sequences of navigation actions) and use them successively to improve their performance, thus linking pragmatic skills to memory function and planning (Verschure et al. 2003, 2014).

The emergentist idea can, in principle, be extended to most or all cognitive domains. During development, as a child grows and learns new skills, it creates (or unveils) increasingly more structure in the input through its actions (e.g., imagine an infant who learns to crawl and then to walk); in turn, it can learn this structure in the form of novel SMCs to be mastered. By acquiring new SMCs, children influence their developmental trajectory, determining a circular causality between development and cognition. Byrge et al. (2014) summarized this nicely: *changes in the action system due to growth and development change the inputs of the brain and have cascading effects on cognition and behavior in a circular process*. For a child, learning to crawl or to walk opens up entirely new opportunities for visual experiences and manipulations (as the child has many more objects within its reach), as well as for social and linguistic exchanges; these, in turn, support a range of changes in object memory, object discrimination, and view-invariant object recognition (Byrge et al. 2014). Accordingly, changes in the action system support cognitive achievements:

learning more skills implies experiencing more new patterns, and thus new opportunities to learn even more skills or to increase one's memory, decision, and social abilities. In principle, even the most advanced cognitive and social skills should be understood as emerging from brain-body-environment dynamics rather than from some form of "internalized" cognitive process, and for this they are better characterized and studied using the tools of dynamical systems (Kelso 1995; Richardson et al. 2008).

Open Issues and Current Controversies

One important consequence of the emergentist view is that development is not a process that unfolds autonomously or proceeds in predetermined stages. Instead, development depends on continuous interactions between the brain, the body, and behavior. Indeed, it is only through this continuous interaction that the necessary information is acquired to support skill learning and cognitive achievements. Furthermore, a circular causality is introduced by developmental changes in body morphology and size (e.g., growth): brain networks support cognition and action (e.g., connectivity), and the resulting behavior shapes inputs to the brain (Byrge et al. 2014). These ideas have a clear appeal but introduce a number of difficulties for current experimental approaches, given the presence of various factors that influence one another and operate over long timescales (e.g., development), and the fact that agents should be tested in conditions where they can freely explore and select their stimuli. Some of these problems can be mitigated using a "synthetic" approach: robots which develop similar abilities over time, as in developmental robotics (Lungarella et al. 2003), can be analyzed as to how they solve cognitive problems equivalent to those faced by living organisms.

How should the driving forces that guide exploration be understood? Most theories assume an initial sensorimotor exploration phase, which is increasingly considered to be systematic and goal-driven rather than a "random" process (Gottlieb et al. 2013; von Hofsten 2004). After a while, for example, a thirsty child can control its hand to reach for a glass, thus achieving its goal of drinking from the glass. This new skill then provides the basis to learn increasingly more complex skills (e.g., throwing a glass or stacking them on top of each other). This process, however, is severely underconstrained. In most engineered settings, the control problem to be learned is known in advance and therefore fixed. Here, however, it is open-ended: it changes over time due to the combined effect of changes in body morphology (due to growth) and the organism's action and sensation patterns as the organism learns new skills (e.g., crawling vs. walking). From a machine learning perspective, the presence of so many degrees of freedom creates a difficult "autonomous exploration" problem that cannot be easily tackled by existing approaches. Insights from developmental robotics and machine learning can help elucidate the importance of using "developmental tricks" (e.g., freezing some degrees of freedom before

From "The Pragmatic Turn: Toward Action-Oriented Views in Cognitive Science,"
Andreas K. Engel, Karl J. Friston, and Danica Kragic, eds. 2016. Strüngmann Forum Reports, vol. 18,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03432-6.

the system is fully able to exploit them) or of focusing exploration to the most informative regions or where learning progress can be observed (Baldassarre and Mirolli 2013; Oudeyer et al. 2005; Schmidhuber 1991a). Here there are ample opportunities for new experiments.

Within this perspective, there is a huge territory that has not yet been completely charted: the social scaffold of development, or in other words, how other people influence the SMCs that an agent (e.g., a child) experiences and learns. Children develop important skills and learn about important concepts through social interaction, not just through individualistic exploration. Initial social interactions are simple (one child, one caregiver) but they become increasingly more complex (e.g., societal). In terms of information theory, the actions of other agents (e.g., caregivers) structure a child's input and thus influence its learning. In some cases, caregivers use specialized *sensorimotor communication* strategies to structure a coactor's input space (e.g., maximize the information gain of children), as exemplified by child-directed speech (motherese) or gestures (motionese) (Pezzulo and Dindo 2011; Pezzulo et al. 2013b). Furthermore, coactors continuously create "social affordances," which greatly expand action and learning possibilities in the same way that tools do, where words might also be understood as "social tools" (Borghi and Binkofski 2014). Social dynamics are studied in many laboratories, but there is a trade-off between designing ecologically valid scenarios and obtaining controlled data. It is also unclear whether a social, interactive (vs. individualistic) acquisition modality of certain concepts influences their "conceptual content": concepts are based on social SMCs and emerge from patterns of interaction between two agents and two brains (including *linguistic* interactions) rather than from patterns of interaction between one agent and the world.

Finally, this perspective recognizes the role of external stimuli and affordances in shaping perception-action loops but pays less attention to the role of internally generated processes, such as *goals*. Focusing on goal-directed behavior has important consequences: a person can, for example, structure input in different ways depending on the goal in mind; this means that a person can elicit different conceptual content. For instance, I can categorize the chair in front of me as an "obstacle" or "resting place," depending on my goal (walking through the room vs. resting), meaning that affordances can be goal dependent. Furthermore, I can act intentionally to *create* or unveil new affordances that fulfill my goals: a boxer can move closer to or farther from an opponent to create a "left jab" affordance (i.e., the area from which he can hit the opponent from the left side) or a "defense" affordance (Araújo et al. 2006). Thus, even within the boundaries of an emergentist or ecological approach to conceptual processing, an individual might need to move from a stimulus- and action-centered perspective to a *goal-centered* view of cognition.

From "The Pragmatic Turn: Toward Action-Oriented Views in Cognitive Science,"
Andreas K. Engel, Karl J. Friston, and Danica Kragic, eds. 2016. Strüngmann Forum Reports, vol. 18,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03432-6.

The Role of "Cognitive Mediators" (or "Neuronal Mediators")

This perspective emphasizes that sensorimotor processes involved in action control can "mediate" higher cognitive or social abilities. However, contrasting proposals exist as to what constitutes exactly the role of a (cognitive or neuronal) "mediator." In one, the role of a cognitive mediator is played by common neuronal substrates across sensorimotor and cognitive domains. In another, (pre)motor brain areas have been proposed to play several roles well beyond action control to support imagery, planning, and action understanding (Jeannerod 2006; Rizzolatti and Craighero 2004). A related proposal is that the "semantic" systems of the brain, which support most cognitive operations, are deeply based on sensorimotor circuits. Embodied theories of cognition argue that the human's conceptual system is grounded in perceptual experience and the sensorimotor patterns elicited during previous experiences and agent-environment interactions; thus the reenactment of sensorimotor experience supports (or is a mediator of) conceptual processing (Barsalou 1999, 2008). Here, importantly, knowledge is retained in modal and multimodal brain areas, in the same format as used for sensation and action; for example, verbal memory in the articulatory control system and object concepts (and the corresponding lexicon) in the same action-perception system used to interact with the objects (Martin 2007; Pulvermüller 2005). The resulting view holds that action-perception circuits might be the neural basis for higher cognitive operations, including attention, the processing of semantic knowledge, and communication (Pulvermüller et al. 2014).

Within embodied theories of cognition, *body processes* and *resources* mediate cognitive abilities and their acquisition. Useful examples are offered by the SNARC (spatial-numerical association of response codes) effect, which suggests that number coding is highly spatially organized and dependent on the agent's perspective, as well as by experiments that reveal a close link between the way subjects move their eyes and the way they solve complex problems (Grant and Spivey 2003). One explanation as to why bodily processes might be integral to cognitive operations is that they mediate the acquisition of conceptual and problem-solving abilities. Learning a category does not consist in the passive sampling of stimuli (e.g., of exemplars of a category); it has strong situated and embodied components. For example, Barsalou (1999, 2008) proposed that categorical representations might emerge when attention is focused repeatedly on the same kind of thing in the world, where the agent's embodiment, perspective, and current activity all constrain this process. Thus, developmental processes might, in principle, explain why conceptual and even linguistic domains seem to be organized around body- and action-relevant dimensions, as in the case of "actions toward the body" versus "action away from the body" (Glenberg and Kaschak 2002).

A related but different view is the notion of "neural reuse" or "recycling" across simpler to more complex domains. Here the idea involves adapting an

From "The Pragmatic Turn: Toward Action-Oriented Views in Cognitive Science,"
Andreas K. Engel, Karl J. Friston, and Danica Kragic, eds. 2016. Strüngmann Forum Reports, vol. 18,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03432-6.

existing information-processing mechanism to novel, more cognitive contexts. Several examples have been proposed (not all action-centered), including the reuse of (a) hippocampal resources from spatial navigation to episodic memory and mental time travel (Buzsáki and Moser 2013; Pezzulo, van der Meer et al. 2014; Buzsáki et al. 2015), (b) cerebellar internal models from movement control to problem solving (Ito 1993), and (c) parietal systems to coordinate transformation from individual to joint actions (Iriki and Taoka 2012; see also Anderson 2010; Dehaene and Cohen 2007). Furthermore, a related view holds that changes in the brain induced by sensorimotor experience (e.g., tool-use learning) create a *neural niche*—or a newly available resource in the form of extra brain tissue—that can expand an agent's abilities in unprecedented cognitive domains (Iriki and Taoka 2012).

Perhaps the most developed hypothesis of "cognitive mediators" is a *prediction-centric* view of cognition, where predictive mechanisms (e.g., generative or forward models) originally developed for motor control have been "exapted" during evolution to promote and mediate increasingly sophisticated cognitive abilities (e.g., planning, action understanding, and problem solving), all of which now form a *motor cognition* domain (Jeannerod 2006). Hence, the most important route toward "cognitive" operations is the *covert* reuse of predictive mechanisms: "action simulation" or "what-if" loops are steered, both in the future and the past, without associated overt actions (Clark and Grush 1999; Grush 2004).

Predictive mechanisms, used in both overt and covert forms, have been linked to several important cognitive abilities. Engel et al. (2013) have proposed that prediction is central to the acquisition of SMCs. For example, learning a SMC specific for a given object corresponds to determining the conditional probability of making a sensory observation given the past movements and observations. This knowledge, in turn, can be used to categorize and "ground" objects. A sponge, for instance, can be recognized in terms of the (anticipated) softness when it is squeezed (either through imagining or memory) (Pezzulo 2011; Roy 2005). Robots can ground navigation concepts using an internal simulation of possible trajectories. Hoffmann (2007) reports that the distance from obstacles is grounded and estimated by running simulations until an agent encounters the obstacle: *dead ends* are recognized through simulated obstacle avoidance, and *passages* are understood in terms of successfully terminated trajectory simulations. Quinton et al. (2013) demonstrated that pictures of animals can be recognized by learning to predict how sensory features change as a function of eye movements. In an *active inference* generalization of this view, the object or event categorization process becomes even more active because a *hypothesis testing* mechanism allows the system to run "experiments" that disambiguate among competing hypotheses. Here, a possible "experiment" consists of directing a saccade to the most informative and discriminatory parts of the environment, rather than just passively collecting "samples" of perceptual evidence. This would discriminate among multiple competing hypotheses and allow an object or event to be recognized faster (Friston et al. 2010a), thus supporting *counterfactual* forms of reasoning (Seth 2014).

Gorniak and Roy (2007) have proposed that the meaning of words and sentences can also be grounded in sets of anticipations learned via sensorimotor interaction. In the "semiotic schemas" framework (Roy 2005), words for perceptual features are grounded into sensory information; for example, "red" is grounded in some (expected) values of the robot's sensors. Object words are grounded as a result of (actual and potential) actions; for example, the meaning of the word "cup" is grounded in the sensorimotor patterns expected by interacting with the object. Finally, in the social cognition domain, action understanding has been explained in terms of simulated action and the reuse of one's own motor repertoire. Here, an observer agent can simulate performing several actions (e.g., kick and push) and compare their predicted sensory effects with the observed results of a performer agent's action. Using this mechanism, the best-matching hypothesis disambiguates what the performer actor is currently doing (Kilner et al. 2007; Wolpert et al. 2003).

Open Issues and Current Controversies

The most obvious implication of the predictive view is that the maturation of predictive abilities shapes the developmental trajectory of children. This hypothesis has motivated studies with infants and children (von Hofsten 2004) as well as others designed to investigate whether motor expertise and superior prediction abilities (e.g., in athletes) yield cognitive benefits (Aglioti et al. 2008; Pezzulo et al. 2013a). All these studies have elucidated important analogies between the neurocomputational architecture of prediction and advanced social and cognitive skills. They have also exposed new areas of enquiry, in particular, into whether a "simulative" mechanism might be insufficient for social understanding and mind reading and might instead act in concert with nonsimulative mechanisms (Frith and Frith 2008). It is unclear what role (if any) these nonsimulative mechanisms may play in executive function (see Hamilton et al., this volume).

Another problem is whether the "action simulation" approach can explain declarative forms of knowledge. Whereas in traditional cognitive theories, procedural and declarative kinds of knowledge are considered to be completely separated, the idea of simulation offers at least one possible mechanism to reuse (in simulation) sensorimotor skills to access that part of knowledge encoded in procedural format in the internal models and make it declarative. Consider the following example: with what finger do you press the "L" button on your keyboard? The act of imagining your use of a keyboard brings forth a series of predictions and (depending on task demands) elicits various types of information: the finger that is actually used to press the "L" button, the weight and color of the keyboard, and so on. Here, the knowledge elicited through reenactment is not confined to action control but can be used for other

From "The Pragmatic Turn: Toward Action-Oriented Views in Cognitive Science,"
Andreas K. Engel, Karl J. Friston, and Danica Kragic, eds. 2016. Strüngmann Forum Reports, vol. 18,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03432-6.

cognitive operations (e.g., for reasoning or linguistic communication) in open-ended ways (Pezzulo 2011). An intriguing proposal is that the reason why this knowledge is normally not available in declarative format during overt sensorimotor engagement is that action control is too fast; this knowledge might become (consciously) available when the demands are less strict, such as during off-line imagery (Jeannerod 2006). Accordingly, the distinction between procedural and declarative knowledge would not be one of format, but rather depend on usage and temporal constraints. Thus the empirical question becomes: Which part of procedural knowledge is or is not accessible?

A related dispute is whether, and how, predictive and internal simulation abilities can support reasoning in sensorimotor as well as more "abstract" domains. It has been proposed that prediction mechanisms can be temporarily detached from the overt sensorimotor loop to be fully "internalized," supporting thought processes. For example, mechanics can assemble or disassemble an engine in their mind before doing it in practice; a climber can simulate climbing a wall before actually doing it—a form of *embodied problem solving* (Koziol et al. 2014). This ability plausibly involves the coordinated re-enactment of exteroceptive, proprioceptive, and interoceptive information as well as, in some cases, some overt body movements (e.g., eye and arm movements that one would have executed in the real situation, as well as pantomime movements; see Figure 2.1). Despite this suggestive evidence, it is not clear in which domain the idea of *thinking as internalized action* holds (Cotterill 1998; Hesslow 2002; Pezzulo and Castelfranchi 2009). Indeed, it is important to recognize that most prediction-based views focus on *forward models*; these are not "generic" predictors, because they only generate predictions that are conditionally dependent on one's own actions, and are constrained by one's own embodiment, sensorimotor system, and experience (Pezzulo et al. 2013a). Reusing forward models for higher cognition means that all cognitive operations remain in some way tied to the same constraints as sensorimotor operations. Is this, however, true for all domains of cognition, or do some domains require amodal symbolic manipulations that are detached from perception and action systems?

The Construction of Abstract and Amodal Domains of Cognition

Within this perspective, action-control mechanisms are viewed as being capable of directly mediating cognitive processing. They are seen as playing a role during the *acquisition* of higher cognitive skills but they are not as important for their deployment: once they have been acquired, cognitive skills become part of an abstract (or amodal or symbolic) cognitive domain that is separate from perception and action systems. In this abstract-and-amodal view, there is no need to use action-perception loops or "action simulations" for higher cognitive operations. Action-centered mechanisms are considered—at best—to be "facilitators," not "mediators," of higher cognition.

From "The Pragmatic Turn: Toward Action-Oriented Views in Cognitive Science,"
Andreas K. Engel, Karl J. Friston, and Danica Kragic, eds. 2016. Strüngmann Forum Reports, vol. 18,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03432-6.

Figure 2.1 Previewing climbing routes as an example of an *embodied problem solving*. Each figure captures climbers in the process of previewing a (novel) climbing route before a competition. Generally, climbers mimic, imagine, and plan their future climbing movements (both overtly and covertly). In a competition, routes are unknown to the athletes: "route setters" ensure that they include nontrivial sequences of movements. Hence it is important for an athlete to plan in advance how to approach the route. This form of problem solving is embodied in the sense that it requires consideration of the athlete's embodied knowledge (e.g., limb length, finger strength, potential opportunities provided by the various kinds of climbing holds, possible or impossible kinematics). The best climbers are able to anticipate a lot of information (e.g., proprioceptive information, body posture at critical points, how much force to use) and make important decisions (e.g., where to rest) before they begin to climb. Of course, these decisions are subject to revision once they are underway. Pezzulo et al. (2010) reported an advantage of expert climbers in a memory task (i.e., remembering sequences of holds in a route), but only when the climbing route had the right affordances (i.e., when it was "climbable," not when the sequences formed a non-climbable route). Photos reproduced with permission from Luca Parisse (risk4sport.com).

The empirical evidence for or against a separate, amodal domain of higher cognition is at the heart of a debate between proponents and opponents of an embodied view of cognition (Barsalou 2008; Caramazza et al. 2014). From a developmental perspective, a key issue involves which (brain and computational) mechanism might produce "disembodiment" or a complete "detachment" of cognitive processing from sensorimotor experience. In one of the earlier computational implementations of Piaget's ideas, Drescher (1991) proposed a "constructivist" schema mechanism that incrementally constructs new components on top of sensorimotor experience. Here, an artificial agent is initially endowed with simple sensorimotor schemas (essentially,

From "The Pragmatic Turn: Toward Action-Oriented Views in Cognitive Science,"
Andreas K. Engel, Karl J. Friston, and Danica Kragic, eds. 2016. Strüngmann Forum Reports, vol. 18,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03432-6.

context-action-prediction triplets) that permit it to interact with a simple artificial environment. However, the goal of the schema architecture is to build increasingly more complex schemas that include nonperceptual elements and conceptual knowledge. A specific example can help to clarify this point. Initially, the first elements of schema triplets (context) include only sensor measurements but successively the agent interactively enlarges its "ontology" by learning new so-called synthetic items (i.e., common causes from a set of related interactions involving action-outcome effects). This schema learning mechanism essentially discovers nonperceptual states that make schema predictions more accurate, thus increasing the agent's knowledge (e.g., the concept of a "cup" along with a schema that encodes a new regularity: "if a cup is in front of me and I grasp it, it will be in my hand"). These new schemas, in turn, can be used as a new starting point to develop (at least in principle) an open-ended repertoire of skills, including more abstract actions (e.g., counting actions that permit counting the cups, or naming actions that permit linguistical referencing). Although this mechanism might seem similar to the predictive view, the synthetic objects created through this learning process need not have sensorimotor components; in other words, they are amodal, not modal or multimodal, states. Once acquired, the new schemas create new abstract, symbolic domains. In the examples above, this would consist of a numerical domain and a linguistic domain that are segregated from action and perception systems, which is at odds with the aforementioned embodied views of cognition.

Open Issues and Current Controversies

Are domains of abstract cognition (e.g., numerical or linguistic cognition) multimodal (in keeping with embodied theories of cognition) or are they amodal (as described above)? A third possibility is that both conditions exist. It has been argued (Pezzulo and Castelfranchi 2009) that it is possible to compose music in both "Mozartian" (i.e., by hearing or rehearsing auditory information) and "Bachian" (i.e., by using only the symbolic music notation—a new code which has its own rules) ways. This "modality" debate is central in contemporary cognitive science, and numerous controversies exist (Barsalou 1999, 2008; Caramazza et al. 2014; Martin 2007; for discussion of language, action semantics, and the modal versus amodal debate, see Pulvermüller, this volume).

This debate has implications for the design of cognitive models and autonomous agent architectures. To date, we have not yet resolved how to devise an architecture that can incrementally build up new skills and knowledge and extend its abilities from simpler to increasingly more demanding cognitive domains. Should such an architecture have separate modules for sensorimotor skills and higher cognition (Anderson 1983), or should the latter be based on (roughly) the resources that were used in the former (Pezzulo et al. 2011; Pezzulo, Verschure et al. 2014)? The goal of achieving *open-ended*

From "The Pragmatic Turn: Toward Action-Oriented Views in Cognitive Science,"
Andreas K. Engel, Karl J. Friston, and Danica Kragic, eds. 2016. Strüngmann Forum Reports, vol. 18,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03432-6.

development was present in the early "schema mechanism" architecture discussed above. If we adapt (roughly) the same concepts to the modern language of generative models and hierarchical architectures for active inference (Friston et al. 2010a; Friston et al. 2015; Pezzulo et al. 2015), synthetic items and schemas would be latent states and models at high hierarchical levels, which encode the "hidden" (i.e., not directly perceivable) causes of perceptions and actions (see Friston, this volume). These hidden states are not necessarily tied to a specific modality, but can generate predictions in multiple modalities—proprioceptive, exteroceptive, and interoceptive—which in turn directly engage action, perception, and emotion processes (Adams et al. 2013a; Clark 2013b; Pezzulo 2014; Seth 2015; Stoianov et al. 2016). During development, new hidden nodes and models are learned that encode regularities (and permit prediction and control) at longer timescales and in different domains (e.g., social domains), thus expanding the scope and potential for control of the agent, and—at least in principle—supporting cognitively demanding cognitive operations such as counting or reading.

In practice, the potential for this architectural scheme (or others) to mimic cognitive development and extend to the domains of higher cognition remains to be fully explored (for further discussion, see Pezzulo 2011). Furthermore, it is still unclear what resources (e.g., representational) would be required for such an architecture. It has been argued that children face very challenging "structure learning" problems during their development, for which structured and symbolic representations might help (Tenenbaum et al. 2011). However, even in the cognitive modeling literature, it is unclear whether the development of abstract cognitive abilities is based on modal (e.g., interoceptive) processes or amodal and symbolic states, and whether these are innate or can be learned (König and Krüger 2006).

Conclusions

The novel "pragmatic" view of cognition offers a way to change how we conceptualize living organisms, their brains, and their behavior. Since the field is young, controversies exist and important elements of action-based theories await in-depth investigation.

Most action-based theories agree that the acquisition of cognitive abilities is guided, during development, by the acquisition of pragmatic skills or a mastery of SMCs. However, there are contrasting proposals as to how exactly pragmatic skills contribute to cognitive development. These proposals can be grouped into three (not exclusive) perspectives:

1. How pragmatic skills contribute to development is tied to the potential of enabling new SMCs in an incremental process of self-organization of both brain networks and behavior. This view tends to assume that

cognitive abilities (e.g., categorization or problem solving) are not discrete operations but rather emerge during agent-environment interactions, without the necessity of representation or "internalization."

2. The key to cognitive development is the presence of "cognitive mediators" (e.g., a core set of prediction abilities that are reused across action control and higher cognition). A popular example is the idea that cognition can be seen as an *internalized* form of action, where the same mechanisms mediate both. This view tends to "cognitivize" action control in the same way it makes cognition action-based; it emphasizes that intentional action has a complex structure and is mediated by sophisticated mechanisms like internal models.
3. Action-control mechanisms contribute to the development of abstract and amodal domains of cognition. Once established, these domains, however, do not depend on action-perception brain systems for deployment. Thus action-control mechanisms are "facilitators" of cognitive development.

This taxonomy points to numerous open issues: How should we understand and study the circular causality between development and cognition? What (if any) are the most important "cognitive mediators"? Do amodal domains of cognition exist and, if so, how are they developed? In other words, are we "Mozartian" or "Bachian" or both?

From a more epistemological perspective, other questions include: Can prediction-based mechanisms "detach" from the overt sensorimotor loop, and would this count as a "representational" role? A related set of questions involves the neuronal implementation of the proposed architectural schemes: Does the brain use symbols to mediate thought processes? Is there evidence of a causal role of sensorimotor representations in higher cognition? Is there evidence of truly amodal brain representations and, if so, how can we recognize them? Are these innate or do they form (and if so, how) during cognitive development?

We must continue to assess the merits of existing action-based theories by designing novel experiments that test emergentist, modal, and amodal views of higher cognition, and by realizing robots that embody these views. Developing more advanced theoretical proposals must be an important objective within the agenda for a "pragmatic" cognition.

From "The Pragmatic Turn: Toward Action-Oriented Views in Cognitive Science,"
Andreas K. Engel, Karl J. Friston, and Danica Kragic, eds. 2016. Strüngmann Forum Reports, vol. 18,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03432-6.