Cognition as Coordinated Non-Cognition

Lawrence W. Barsalou
Emory University

Cynthia Breazeal
Massachusetts Institute of Technology

Linda B. Smith
Indiana University

December 15, 2006

Running head: Coordinated Non-Cognition

Address correspondence to:

Lawrence W. Barsalou
Department of Psychology
Emory University
Atlanta, GA 30322

(404) 727-4338
barsalou@emory.edu
http://www.psychology.emory.edu/cognition/barsalou/index.html

Abstract

We propose that cognition is more than a collection of independent processes operating in a modular cognitive system. Instead, we propose that cognition emerges from deep dependencies between all of the basic systems in the brain, including goal management, perception, action, reward, affect, and learning. We also believe that human cognition greatly reflects its social evolution and context, as well as major contributions from a developmental process. After presenting these themes, we illustrate their application to the process of anticipation. Specifically, we propose that anticipations occur extensively across domains (i.e., goal management, perception, action, reward, affect, and learning) in highly coordinated manners. We also propose that anticipation is central to situated action and to social interaction, and that many of its key features reflect the process of development.

Key Words

coordination; development; embodiment; robotics; situated cognition; social interaction

> *"In short, the practically cognized present is no knife edge, but a saddle-back, with a certain breadth of its own on which we sit perched, and from which we look in two directions in time" William James, 1890, Chapter 15.*

There can be little doubt that cognition is informed by the past. There can also be little doubt that much of cognition is about the future. What will happen next? Where must I put my hand so that I pick up the cup? If I say what I'm thinking, how will my listener respond? Human cognition seems magical in its prescience, in its attempts, often successful, to foresee the future. How should we understand this core phenomenon of anticipation (also viewed widely as prediction)? How should we study anticipation so as to understand its underlying processes and principles?

One possibility is to locate anticipation in cognitive computations. This fits the traditional view in which mental life is divided into discrete steps of "sense-think-act." Cognition, by definition, is about the "think" part, the knowledge and processes that mediate perceiving and acting. Knowledge, from this perspective, is amodal and propositional, consisting of relatively fixed representations. Anticipation, taking the form of inference, is one of many processes that operate on this knowledge. Specifically, anticipation operates on stable propositional models of the world, representations of goals states, knowledge of plans, and so forth to produce symbolic representations of what will happen next.

This approach has a long and venerable history in cognitive science. It makes considerable sense for many reasons, and has made major contributions to the field. For example, this view has been adopted widely in the study of text comprehension (e.g., Kintsch & van Dijk, 1978), problem solving (e.g., Newell & Simon, 1972), and reasoning (e.g., Rips, 1994). In general, this approach assumes that people have stable models (or schemas) of the logical, causal, and temporal structure of the world, and that a wide variety of processes operate on this knowledge. For example, these models support understanding and predicting how rain and sun contribute to plant growth (e.g., Gentner & Stevens, 1983), how levers and pulleys work to increase the force on objects (e.g., Chandrasekaren & Josephson, 2000; Forbus, 1993; Hayes, 1985), and how events in a restaurant visit unfold over time (e.g., Schank & Abelson, 1977). In these classic accounts, the knowledge that underlies anticipation is assumed to be fundamentally different in kind, and theoretically separable, from the real time processes of perceiving, acting, and experiencing emotion.

In this article, we present an alternative account that we believe is spreading throughout cognitive science. Although this alternative account takes many forms that are evolving rapidly, several important principles appear to underlie it.

First, knowledge has no existence separate from process, but is instead embedded in, distributed across, and thus inseparable from the real time processes (e.g., Barsalou, 1987,1989, 1993; McClelland et al., 1986; O'Regan & Noe, 2001; Rumelhart et al., 1986; Jones & L. Smith, 1993; Samuelson & L. Smith, 2000; Port & van Gelder, 1995; Spivey, 2006). From this perspective, there is not a fixed and separate representation of anything.

Second, to understand cognition, it is essential to understand fundamental contributions from what have traditionally been viewed as non-cognitive systems, including goal management, perception, action, reward, affect, social interaction, and development. Rather than cognition being a modular system that operates independently of these other systems, we believe that these other systems play central roles in cognition per se, and thus are not really "non-cognitive."

Third, to understand a cognitive process such as anticipation, one must consider the question of how to build it. On way to address this question is to understand how the process develops from birth over time in biological agents. Another approach is to engineer the process in artificial intelligence, for example, in robots.

Fourth, to understand a cognitive process, such as anticipation, it is insufficient to understand the process in isolation. Instead, understanding its coordination with other processes that are typically co-active during real world cognition is also essential. We believe that understanding the coordination of a process with other processes is as important, if not more important, than understanding the internal structure of the process itself.

In the next three sections, we develop the second, third, and fourth themes just described in greater detail. Because the lack of separation between representation and processing has been addressed widely, we do not address it further. Because the latter three themes have received less attention, we focus on them here. In the final section, we sketch an account of anticipation from the perspective of these four principles.

Cognition Depends Intrinsically on Non-Cognitive Processes

We begin with an analogy (or cautionary tale) that may at first seem far from the issue of anticipation. The lesson of this analogy is that one cannot separate cognition from non-cognitive processes, nor from the coordination of all these processes in real time (L. Smith & Thelen, 1993).

The analogy concerns quadrapedal locomotion in cats. Classic theories of motor behavior proposed the construct of a stable central pattern generator (CPG) to explain the patterned regularity of a cat's alternating limb actions during locomotion. Accumulating findings, however, raised problems for this construct. For example, cats walk with alternating limbs on a treadmill even when their spinal cords are separated surgically from their brains (e.g., Delcomyn, 1980). Based on these findings, some researchers made the stretch that the CPG is in the spinal cord. The validity of a CPG has been questioned on many other grounds as well (Thelen & L. Smith, 1994).

Even if a CPG did exist, it could not underlie the stability apparent in the alternating limb action of cats walking across real terrains. Cats walk backward, forward, on grass, on hills, on rubble. They side step objects; they walk when one limb is in a cast. Although a stable alternating limb pattern is apparent in all cases, the adaptability and variability of this pattern is remarkable.

Alternating limb movements in these different contexts requires fundamentally different patterns of muscle firing to maintain the same global stability of alternation. If a CPG exists, then systems outside this structural must constantly make walking happen in globally similar but appropriately different ways across diverse contexts. Thus, the overall system that produces walking is not simply the CPG, but is a distributed system of multiple systems, including the CPG. Another possibility is that there is no CPG, and that the global pattern or alternating limbs emerges from a variety of interacting systems.

We believe that the importance of "outside" systems in walking is analogous with the nature of higher cognition. Theories that separate the "think" part from the perceiving, acting, and affective parts fail to recognize the importance of these other systems in cognitive activity. Indeed, we have increasingly come to believe that these other systems are fundamental parts of the cognitive system—not merely peripheral systems. Furthermore, just as there may not be a CPG module that produces cognition, there may not be a pure cognitive module that produces cognitive processing. Instead, cognitive processing may emerge from the interaction of many different systems, analogous to how cat locomotion emerges from many different systems. As described later, we further believe that anticipation is likely to emerge from many systems in this manner. In the following section, we outline contributions from non-cognitive systems, including the systems that implement perception, action, reward, affect, goal management, motivation, and social interaction.

The brain's modality-specific systems. Increasingly, researchers argue that systems in the brain for perception and action are essential parts of the cognitive system (e.g., Barsalou, 1999b; Damasio, 1989; Glenberg, 1997). Rather than the cognitive system being modular, it utilizes mechanisms in modality-specific systems for its fundamental representation and processing activities (both in the brain and in the body). From this perspective, knowledge is represented as simulations in modality-specific systems, not as amodal symbols elsewhere in a self-contained modular system. Thus, understanding perception and action is essential for understanding cognition.

Furthermore, the coordinated relationships between perception, action, and cognition must be identified to characterize cognition adequately. As a result of viewing cognition this way, cognitive processing becomes integrated with perception and action, something that classic symbolic theories have not achieved. Grounding knowledge is grounded in modality-specific representations provides one solution to this problem. Because the representations that underlie knowledge can be mapped to corresponding representations in perception and action, they link cognition to the sensory-motor interface.

Perusing the evolution of cognition from the simplest organisms to humans, it is clear that cognition in the simplest organisms began with perception coupled directly to action via hard wired circuitry. Over the course of evolution, increasingly sophisticated mechanisms evolved to mediate perception and action. Nevertheless, the abilities to perceive and act remained foundational. Not only are they essential in the simplest organisms, they are central throughout human activity. A system is not fully intelligent if it lacks these systems. Furthermore, the detailed structure of these systems is linked closely with cognitive abilities, given the sophistication and complexity of perception and action systems in humans. Increasingly powerful perceptual and motor abilities may have resulted from increasingly powerful cognition. Alternatively, increasingly powerful perceptual and motor abilities may have prodded the sophistication of cognition forward. Building a cognitive system that lacks this interface is likely to fall significantly short of implementing the magic of human cognition.

The brain's affective and motivational systems. Affect regulation and motivation play substantial roles in human activity. Following Damasio (1994), researchers have become increasingly convinced that cognition divorced from affect is not rationale, and that optimal cognitive performance occurs when emotional information is included in decision making. Over time, affective information accumulates with events and objects, indicating whether they are generally associated with positive or negative affect. On later encountering these events and objects, average affective information becomes available to quickly to support decision making. When this information is lacking, decision making suffers. Thus, affective information plays a central role in cognition. Understanding cognition does not seem possible if these mechanisms are excluded.

A related example is the central role of the brain's reward systems in learning (e.g., Ashby et al., 2005; Granger, 2006). As rewards are experienced for stimuli, cortico-striatal loops capture this information. Later, when these stimuli are perceived again, associated reward circuits become active and to determine actions on the stimuli. In this manner, the reward system fundamentally shapes cognitive processing.

Motivational processes similarly play central roles throughout cognition, from initiating basic response systems for hunger, thirst, etc., to controlling the pursuit of complex social and personal goals. Without these systems, theories of cognition are incomplete. Furthermore, executive systems are necessary for maintaining goals in working memory, and for deciding when to pursue or drop goals.

In general, cognition appears to depend heavily on affective processing, and to be linked closely with affective and motivational systems throughout the brain. Although emotion, reward, and motivation are typically studied as independent processes, they appear to be intimately integrated with each other and with the cognitive system (e.g., Barrett, 2006; Barrett et al., 2007). Rather than being modular systems, it is highly likely that each system is densely interrelated with all the others.

Systems for social interaction and communication. Researchers who study the evolutionary origins of human cognition argue that social pressures shaped the human cognitive system extensively (e.g., Donald, 1993; Rizzolatti et al., 2002; Tomasello et al., 1993). Because increasingly sophisticated social interaction enabled major gains in evolutionary fitness, powerful new mechanisms evolved in the human brain that were absent in earlier species. Included in this list of mechanisms are joint attention, perspective taking, mirroring, imitation, and language. By developing these mechanisms, humans were able to represent and coordinate complex social activity, which in turn yielded major gains in controlling environmental resources and maximizing reproductive outcomes. According to this perspective, if such mechanisms are not included in an intelligent system, it will not come close to achieving the magic of human cognition.

Rather than simply processing inanimate information, such as isolated words and symbols, cognitive scientists should be also be studying the social information that our cognitive systems evolved to process. For example, cognitive science research should address how people infer the mental states of other agents from behavioral signs, which is arguably one of the most frequent and important cognitive activities that humans perform. Analogously, cognitive science research needs to understand how agents learn skills through joint attention, mirror circuits, imitation, and verbal instruction (e.g., Breazeal et al., 2005). Humans probably learn important things more often from social interaction than they do from isolated individual interactions with inanimate stimuli. Cognitive science research also needs to better understand how the cognitive system supports the learning of coordinated activities in division-of-labor settings, as well the learning of competitive strategies, such as deception, in conflict situations (e.g., Hutchins, 1995).

Because cognitive systems in the human brain probably evolved to support unusually extensive and sophisticated social interaction, it is likely that mechanisms for processing social information are densely inter-connected with cognitive mechanisms. To understand the human cognitive system, it will therefore be necessary to study cognition in its social contexts, not just when processing non-social information, such as isolated words.

An example. The emergence of cognition from multiple domains has been explored in research on anthropomorphic robots (e.g., Breazeal et al., 2005). One example demonstrates how a mirror-neuron-like system emerges from multimodal coordination and re-entrance during face-to-face play. The robot brings the following abilities to the task: (1) the ability to visually track the facial features of a person, (2) the ability to "motor babble" by exercising its initial repertoire of facial movements, (3) the ability to sense its own facial configuration, (4) a coarse mapping of "organ relations" that roughly relate regions of the robot's own face to regions of the perceived face of others, and (5) attraction to contingent interactions, where a contingency metric determines whether a visually perceived movement is temporally contingent on the robot's own movement.

The task that couples these systems is an imitation game. As the human developmental literature shows, it is often adults who take the initiative to imitate the facial expressions of infants (e.g., Jones, 2006). Hence, as Breazeal et al.'s robot "motor babbles" by exercising its repertoire of facial expressions, a human participant imitates the robot. Although the robot cannot see its own face, it can sense its facial configuration via proprioception (i.e., position sensors). As the robot moves its face from expression to expression, it observes visually how the human's face responds. Meltzoff and Moore (1997) posit that the desire to match and to be matched by others is innately rewarding to human infants (also see Meltzoff, 1996). A contingency metric allows the robot to determine which regions of its own face the human has chosen to mimic, causing these regions to become salient. Through these processes, the robot attentively selects pairings of matched regions between its own face to and those of the human. Via a standard statistical learning algorithm, these pairings teach the system about how perceived visual movements of the human's face map onto the robot's corresponding motor movements.

The inter-modal representations acquired from the coordination of perception and action allow the robot to directly compare its own motor movements to the observed expressions in the same motor-based coordinate system. Once this representation exists, the robot can mimic facial expressions of the human. The robot can also mimic novel never-produced facial gestures by searching over a weighted blend space of its motor repertoire to find (and generate) an adequate match. Furthermore, this learning increasingly establishes a mirror system that can be used in other tasks (e.g., Rizzolatti et al., 2002). The same representations can be used not only to generate the robot's own actions, but to recognize the same actions in others. Acquiring this ability has profound implications for many other forms of social learning, such as true imitation, social referencing, and other forms observational learning.

Cognition Emerges from Coordinated Sets of Processes

The divide and conquer strategy that isolates processes in empirical research, AI engineering, and formal modeling has important strengths. In empirical research, isolating processes is essential for ruling out confounding variables, demonstrating control, and establishing causal (as opposed to correlational) relationships. In engineering and formal modeling, isolating processes greatly reduces the complexity of systems that must be built and the problems that must be solved. It also makes analytic understanding and formalization easier.

Clearly, the benefits of isolating processes are significant, and this approach has yielded tremendous gains since the cognitive revolution. Nevertheless, we believe increasingly that understanding cognition involves understanding the coordination of cognitive processes at least as much as understanding processes in isolation. We further believe that much of the magic of human cognition results from sophisticated coordination, not just from the processes themselves.

Furthermore, we worry that the characterization of a process acquired from studying it in isolation may differ significantly from the characterization acquired in the context of coordinated activity. In particular, we worry that a stand-alone AI implementation of a process can not be plugged effectively into a larger coordinated system without considerable reprogramming, if not complete redesign. Thus, studying and implementing cognitive processes in a complete system that performs many processes in parallel may teach us more about these processes than studying and implementing them in isolation.

Just because the cognitive science literature is full of research on isolated cognitive tasks, it does not follow that this approach will eventually lead to a complete account of cognition. If we understand how a brain implements many individual processes, it does not follow that we understand how they work together in a coordinated manner. Similarly, just because we can implement many individual processes does not mean that we understand how to implement them operating together.

A related theme is that agents in the real world do not perform tasks in isolation. Instead, they perform sets of coordinated tasks that produce coherent goal-directed behavior. For example, organisms do not perform categorization alone. Instead, they perform categorization together with perception, inference, action, reward, and affect.

Coordination during situated action. Situated action provides one way of exploring coordinated processes (e.g., Barsalou, 2003, in press; Brooks, 1991; Clark, 1997; Glenberg, 1997; Robbins & Aydede, in press). In situated action, agents have goals. As they navigate their environment during goal pursuit, they manage goal priorities, based on motivational states and opportunities in the environment. At each moment, they also perceive the environment, categorize entities and events, and draw inferences that go beyond the information given. They also perform many kinds of memory retrieval about possible actions, rewards, affective states, etc. And ultimately, agents act. At each point, they are not only learning about how to perform individual cognitive processes, they are also learning how to coordinate them.

Thus, there appears to be a core set of coordinated cognitive processes associated with situated action that include goal management, perception, categorization, inference, action, reward assessment, and affect (with learning throughout). These processes do not occur in isolation but together in a coordinated manner. Indeed, how these processes interface with each other is probably as important as how each operates internally. Furthermore, it is likely that the acquired coordination of these processes shapes the internal structure of these processes themselves.

We suspect that the coordinated processes underlying situated action exist in many species, not just in humans (e.g., Barsalou, 2005). It is easy to imagine many other species managing goals, perceiving and categorizing the environment, generating simple inferences about what will happen next, performing actions based on previous rewards, and experiencing affect in response to the outcomes of those actions. Thus, understanding the basic set of coordinated processes that underlie situated action is likely to be informative about the intelligence of many species.

Furthermore, understanding this particular set of coordinated processes may bear on the core kernel of intelligence that evolved into human intelligence. Because this may be the most basic set of coordinated processes in the human brain, it might make scientific sense to understand it first, both empirically and theoretically. Rather than trying to understand the most advanced human abilities first, such as logic and mathematics, it might make more sense to understand how these advanced abilities built upon more basic abilities that existed previously.

Coordination during social interaction. Another important set of coordinated processes are those that support the unusually sophisticated social abilities of humans. As described earlier, humans are unusual in establishing joint attention and in representing other minds. Humans also have unusually good communication systems that allow them to coordinate shared mental states and complex social activities. Humans are also unusually good at learning from each other via observation, imitation, and verbal instruction.

Although many of the basic processes that support situated action probably contribute to social interaction, other processes are probably important as well. Thus, studying the coordination of these additional social processes is probably central to understanding human intelligence. It is also probably important to study social processes together with those for situated action, given that much social behavior revolves around goal-pursuit in the environment (e.g., Barsalou, 1999a; Barsalou et al., 2003; E. Smith & Semin, 2004).

An example. Correlations that emerge from coupled like bodies with like cognitive systems in a shared situational context may lead to “magical human cognitive abilities,” such as “mind reading” inferences about the internal states of others. The key correlations are:

- correlations between the appearance of the self and the appearance of others (e.g., hands to hands, feet to feet),
- correlations between the behavior of the self and the behavior of others (looking to an object),
- correlations between one’s bodily behaviors and internal states (e.g. looking left and remembering what was on the left, maintaining the memory of a goal and looking in the direction of the goal),
- correlations between the external states of others and one’s own internal states (where someone looks, looking there oneself, and thinking about that location).

Through such couplings and coordinations, an artificial device, such as a robot, could develop “human-like intuitions” about the internal states and anticipated behaviors of others.

Mind reading abilities could result from robots developing the ability to simulate the behavior and mental states of others in their own cognitive systems (e.g., Gray et al., 2005). Specifically, a robot could learn to decode emotional messages conveyed through facial expressions by leveraging its own facial imitation abilities to bootstrap emotional empathy. Various experiments with human adults have shown a dual affect-body connection, whereby posing one’s face into a specific emotive expression elicits the feeling associated with that emotion (Strack et al., 1988; Niedenthal et al, 2005). Hence, imitating the facial expressions of others should cause an infant to feel what the other is feeling, thereby allowing the infant to learn the association of observed affective expressions on others with the infant’s own affective states. Other time-locked multi-modal cues may facilitate learning this mapping, such as affective speech that accompanies affective facial expressions between caregivers and infants.

A robot can similarly learn the meanings of affective expressions signaled through another person’s body language and vocal utterances. In humans, affective responses are central to appraising environmental and internal events that are significant to an agent’s needs and goals

(Plutchik, 1991; Izard, 1977). Breazeal (2003) developed a robotic implementation of a simple appraisal process, based on Damasio's theory of somatic markers (1994). This implementation tags perceptual information and internal states with affective information, including valence (positive or negative), arousal (high or low), and novelty (high or low). Various tones of speech, for example, are associated with affective appraisals for positive vs. negative affect (e.g., Breazeal & Aryananda, 2002; Fernald, 1989). During face-to-face interactions between a robot and a human, heterogeneous processes become coupled via these appraisal mechanisms. When the robot imitates affective facial expressions of the human, body-affect pathways evoke the corresponding affective state that ordinarily arises during the robot’s own affective response to a stimulus. These imitations are also reinforced by affective information acquired from the human’s speech signal. All of these multi-modal states are time-locked because of the similarity in bodies and body-affect mappings. As a result, the robot can learn to associate its internal affective state with the corresponding observed expression of a human. Through this “empathic” or experiential approach to social understanding, the robot uses its own cognitive and affective mechanisms as a simulator for inferring the human’s affective state as conveyed through behavior.

The Incremental Process of Development

Developmental researchers have argued for decades that studying adult cognition in isolation will never be successful (e.g., Bjorkund & Pelligrini, 2000). Instead, understanding the mature system requires understanding the history of biological growth, social interaction, and cognitive processing that produced it. L. Smith and Gasser (2005) proposed that the developmental process in humans is successful because the developmental environment contains several important characteristics. First, partially redundant sources of sensory-motor information in the learning environment allow babies to educate themselves, without teachers, simply by interacting with the world. Second, an incremental learning process over the course of development creates capabilities—like understanding the cause and effect relations among actions and rattles—that could not exist genetically at birth. Third, rich statistical structure in the physical world is central to this incremental learning process. Not only does this structure scaffold development of the adult cognitive system, it remains a continual source of constraint and support during adult learning and behavior. Fourth, extended exploration is essential for the development of a mature cognitive system, where exploration discovers structure in the physical world, and also produces important cognitive skills, ranging from simple motor behaviors to creativity. Fifth, the development of a mature cognitive system depends on extended experience with other agents who constitute rich sources of instruction at many levels, including knowledge, skills, and meta-cognition. Sixth, the development of a mature cognitive system depends on extended experience with other agents operating together in a shared environment. As agents share information symbolically and non-verbally, important knowledge, skills, and meta-cognition develop.

System-level properties of human cognition may only emerge from an extended history of coordinating cognitive and non-cognitive processes over the course of development. For example, coordination of the processes that underlie situated action and social interaction are probably some of the major achievements and milestones of the developmental process. Anticipation is another major system-level property likely to develop from a long history of coordinated development. Although the development of individual processes is certainly important, the development of their coordination is no less important. Furthermore, the development of coordination may affect the internal structure of individual processes. If so, then fully understanding an individual process cannot be achieved by studying it in isolation, or even in coordination during adulthood. Instead, understanding the developmental history that coordinated the process with other related processes during its development may be essential.

The training regimens that structure development are likely to play particularly important roles in the development of coordination, where these regimens include experience with the physical world and experience with social agents. The structure of training regimens may often evolve implicitly to support simple coordination initially, followed by more complex regimens later that produce more complex coordination. Interestingly, there may be no other way to establish the kind of flexible, programmable coordination seen in humans. Although genetically-based coordination may be sufficient in simpler species, training regimens may be the only way to achieve human levels of coordination. Indeed, the magic of human cognition may reflect the results of these regimens to a considerable extent.

What kind of architecture? How should we attempt to build artificial agents who have system-level properties that develop from a long history of coordinating processes between classic cognitive and non-cognitive domains? One possibility is to build on the (considerable) successes of the classic sense-think-act approach and implement a hybrid system. Specifically, researchers could start with a mature theory of symbolic cognition, and then add modules that implement goal management, perception, action, affect, reward, social interaction, and development. Building a hybrid system may be feasible, at least to some extent. If so, this would be of considerable theoretical interest and significance, because it would suggest the feasibility of the modular approach. Furthermore, such success might suggest that the cognitive system per se includes something similar to the relatively modular cognitive system in classic symbolic theories. At a minimum, such success would indicate that these theories capture important functionality in the brain.

An alternative strategy, and one that we believe cognitive science should consider, is to develop new architectures motivated explicitly by the attempt to integrate diverse domains of natural intelligence, including goal management, perception, action, cognition, reward, affect, social interaction, and development. From studying these domains together and assessing what it would take to build computational systems that integrate them, alternative architectures may emerge that may be much more powerful than previous architectures.

An example. Recent attempts to engineer cognitively and socially intelligent robots offer good examples of how sophisticated behaviors and learning abilities emerge developmentally from the complex interactions of multiple systems. Social referencing is the ability to use the emotional reaction of another agent to help form one's own affective appraisal of a novel situation, which can then be used to guide subsequent behavior (e.g., Feinman, 1982). In human infants, social referencing arises under conditions of uncertainty and ambiguity, when one’s own intrinsic appraisal processes are not adequate (e.g., Campos & Stenberg, 1981). Social referencing is an important form of emotional communication and is a developmental milestone for human infants in their ability to learn about their environment through social interaction. It is also a behavior built from many component processes, including emotional empathy (bootstrapped from early facial imitation as discussed above), joint attention, and the ability to form affective memories (associating positive or negative valence to stimulus representations in memory).

Thomaz et al. (2005) observed the development of social referencing in a robot, and came to the following conclusions about the developmental process. A robot's social referencing ability emerges in real-time from the dynamic interaction of various capabilities during interactions with a person. When the robot encounters a novel object, the object appraisal mechanism tags the object as novel, which biases the emotion system to evoke a state of mild anxiety. The robot's face expresses a state of heightened arousal as it looks upon the novel object. The robot also looks to the human's face to "soothe" itself. The human often tends to react in naturally instructive ways. The human may notice the robot's initial reaction to the unknown object and decide to familiarize the robot with it. For example, the human may pick up the object and share her reaction to it with the robot.

The robot's attention system determines the robot's focus of attention, monitors the attentional focus of the human, and uses both to track of the referential focus (the object that the interaction is about). For instance, the robot looks to the human's face, thereby allowing the robot to witness her emotional response, and then looks back at the novel toy to share attention with her about the referent. The fact that the human is gazing and reacting toward the novel object draws the robot's attentional focus to it as well. By computing relative-looking-time, the robot establishes the novel object as the referential focus.

As the robot's attentional focus shifts to the human (while maintaining the novel object as the referential focus), the robot extracts the affective signal from her voice by analyzing her vocal prosody for arousal and valence levels. The empathic mechanism described previously enables the robot to also extract the affective meaning from the human's facial expression. The resulting change in the robot's internal affective state triggers a remembering process that establishes a memory of the new object, tagged with the robot's affective state. Thus, the novel object is appraised with socially communicated affective information and committed to long-term memory.

Over time, the robot becomes increasingly sophisticated in its ability to perform social referencing. Although this ability is not present initially, it emerges from the soft assembly of existing processes following interactions with adult agents. If this is how social referencing develops, then assuming that it exists as a pre-existing modularized process is misguided. Instead, this fundamental cognitive ability results from extended practice at coordinating simpler processes.

The "developmental" achievements of this robot should be interesting to researchers of human development for several reasons. First, they provide a measure of our current understanding of developmental process. To the degree that we understand the core developmental principles that produce human intelligence, we should be able to apply these principles when engineering artificial intelligence. Second, these robots serve as physical platforms on which we can model complex coordinated processes across domains (e.g., social referencing). Specifically, robotic platforms allow researchers to present real time tasks in physical and social environments to embodied cognitive systems that are softly assembling complex assemblies of processes repeating across development.

Anticipation as Coordinated Non-Cognition

In this final section, we apply our themes to the cognitive process of anticipation. We first describe how non-cognitive processes enter into anticipation. We then describe how viewing anticipation as coordinated processes differs from thinking about it as a single process.

Choosing situations. We believe that optimal rates of progress in understanding cognition depend on the judicious choice of situations for scientific study. As suggested earlier, we believe that situated action is one particularly important situation. Because this situation occurs ubiquitously across species and is central for survival, it seems particularly important to understand. To the extent that many cognitive processes evolved to support situated action, studying and understanding this situation should be essential to understanding these processes. For these reasons, we begin with situated action. Because social interaction also appears highly central in human intelligence, we consider social situations as well.

As described earlier, we believe that situated action includes the following processes: goal management, perception, categorization, inference, action, reward, affect, and learning. In this situation, an agent is typically moving around with a goal in mind, perceiving the environment, and categorizing what is present. Once categorizations are made about perceived objects and events, the categories accessed generate likely inferences about events, actions, rewards, and affects that could follow. The agent then selects an action and performs it. Events in the world follow that produce rewards, which in turn produce affect and learning. Over time, this cycle iterates constantly, taking many variations that depend on current conditions and the agent's expertise (including developmental level).

In social situations, other agents are also present. They, too, are typically pursuing goals that may be competitive or cooperative. They, too, are having cognitive and affective states. They may further model behaviors that serve as instruction, and may offer instruction explicitly, either verbally or by doing.

Situated anticipation. Where does anticipation arise in these situations? Everywhere. And in intricately coordinated manners. Once a goal is selected, plans and possible courses that plans could take are anticipated. Relevant stimuli in the environment are anticipated as well. When entities and events in the environment are perceived and then categorized, category knowledge generates anticipations in the form of categorical inferences, including potential actions that could help achieve the current goal. As these actions are entertained, their consequences are anticipated, including reward and affect. As feedback is encountered, it specifies whether anticipations were correct or incorrect, producing extensive learning at multiple levels.

The coordination among all these anticipations is extensive. An active goal must coordinate with perception to identify goal-relevant information in the environment. The goal must also coordinate with working memory to maintain goal-relevant information, such as relevant stimuli in the environment to find. Once a relevant entity is identified, extensive coordination between memory, perception, and action must occur to interact effectively with it. As the consequences of actions become available, coordination between anticipated rewards and affects must occur to evaluate what has happened so far, and what to do next, if anything. Further coordination must occur with goal management, starting the cycle all over.

Many additional anticipations arise in social situations, falling under classic topics in social psychology, such as person perception and causal attribution. On perceiving an agent act, for example, a social perceiver anticipates further actions that are likely to follow. Perceiving an embodied state, such as a facial expression or posture, similarly produces anticipations about subsequent actions. Inferring the mental states of agents also leads to extensive anticipations about what the agents will do next. Situations, too, play central roles in producing anticipations, given that particular behaviors occur in particular situations. In general, tremendous amounts of anticipation occur during social interaction.

Anticipation is central to instructional interactions and to collaborative work. In instructional settings, the teacher generally anticipates what should happen next in the domain of study much better than can the student. The teacher’s job is often to teach students sequences of operations that achieve goals, including mental operations and assessments, not just physical actions. As students become increasingly competent, they become increasingly adept at anticipating what to do next and what should happen as a result, no longer requiring the teacher’s assistance. In collaborative work, each co-worker must anticipate what other co-workers will do, and how the collaborative process will evolve. Extensive coordination between agents in all these settings is obviously essential for success.

Studying situated anticipation empirically. How should a rigorous experimental psychologist approach the study of situated anticipation? Obviously, the situations just described are so complex that controlling and analyzing them with classic methods is not feasible. Nevertheless, we believe that these situations should play a central role in motivating experimental research.

Typically, experimental paradigms are chosen with little, if any, interest in their ecological relevance. Instead, the primary reasons for selecting a paradigm are ease of implementation in the laboratory, potential for rigorous control, and tractability in mathematical modeling. For example, research on anticipation in cognitive psychology has been dominated by research on lexical and semantic priming from words. Clearly, much elegant work has resulted from this approach. We believe, however, that this work likely to have little impact until it demonstrates its applicability to real world problems. What potential implications does our understanding of semantic priming have for situated action and social coordination? Because these issues are usually not considered when choosing research paradigms, these paradigms have little potential for informing them.

Thus, we believe that situated action and social interaction have tremendous potential to drive experimental research forward. Consider how framing experimental work in this manner could generate novel and potentially valuable research paradigms. Specifically, consider situated action. Researchers could take each specific form of anticipation in situated action and develop a paradigm for understanding it. Although these researchers would be isolating mechanisms, they would be isolating mechanisms that belong to a larger coordinated system that we know is central to human (and non-human) activity. As researchers increasingly understand specific forms of anticipation in situated action, they could begin to study the coordination of different forms. Although this would increase the complexity of experimental paradigms, studying small subsets of a coordinated system in a controlled manner is certainly feasible.

By framing experimental investigations in this manner, entire systems of coordinated processes from multiple domains could be investigated that ultimately yield the magic of cognition. Anticipation would be studied in goal management, perception, action, reward, and affect. Not only would we understand individual anticipation processes, we would understand a set of individual processes that operate together in an ecologically important situation. And we would eventually come to understand how these individual processes coordinate to yield the magic of cognition.

Along with classic experimental work, qualitative and descriptive methods would be valuable as well. Before beginning analytic laboratory work, it seems essential to describe the component processes of the target situation, and to document their patterns of coordination. Rather than relying on arm chair assessments of what a target situation contains, rigorous assessments of its content should be made, using standard observational and correlational techniques. Such studies could be viewed as analogous to the extensive documentation of phenotypes that preceded more analytic laboratory work in genetics. Before identifying genetic mechanisms, it was necessary first to identify the hereditary distributions of phenotypic patterns to be explained. We believe that a similar state of affairs exists in cognitive science. We first need to identify the components of important situations, such as those for situated action and social interaction, along with their statistical patterns of distribution. Once we have this descriptive information, laboratory paradigms can then be used to isolate important processes, identify their properties, and establish their coordination with other processes.

Implementing situated anticipation. How should situated anticipation be implemented in AI systems? What is the best way to implement all the forms of anticipation that occur across domains during situated action and social interaction, and to create effective coordination between them? The obvious answer is robotics, more specifically, robotics in social environments under developmental training regimens.

Building an autonomous agent that captures the magic of human cognition requires the inclusion of all relevant domains, including goal management, perception, action, categorization, inference, affect, learning, and communication. Furthermore, these autonomous agents need to operate effectively in situated action and social interaction. Given the central role of developmental accumulation, judicious choice of training regimens within these situations is central as well.

Getting all the processes from these domains to work increasingly together across a developmental trajectory requires solving the coordination problem. Solving the coordination problem may also help specify the individual processes correctly in the first place. Clearly, there may be times when implementing a process in a circumscribed toy domain has its benefits. Ultimately, however, the process must work effectively together with other processes across domains in a complete autonomous agent. For these reasons, we believe that the gold standard for implementing anticipation should be implementing it in robots, who experience developmental trajectories in social domains.

An example. Piaget's (1952) book, *The Origins of Intelligence*, presents an insightful example of how anticipation emerges in infants. Piaget placed, for the very first time, a rattle in his 4-month-old infant's hand. As the infant moved the rattle, it came into sight and made noise. The sight and sound aroused and agitated the infant, inducing further bodily motions and causing the rattle to move even more rapidly in and out of sight, and to make even more noise.

Infants at this age have very little organized control over their hands and eyes. They cannot yet reach for a rattle. If given one, they do not necessarily shake it. If the infant accidentally moves the rattle, however, visual, auditory, and somatosensory consequences result, capturing the infant's attention. As these unintentional events repeat, the infant increasingly gains intentional control over shaking the rattle.

Piaget allowed the infant to play with rattle repeatedly for several days and observed the emergence of anticipatory action. At the mere sight of the rattle, the infant would begin to move its hands in the coordinated pattern acquired from past experience. The unplanned and untaught relations between actions and outcomes constituted a self-organizing system that led to anticipatory representations of cause and effect. Piaget referred to such patterns as *secondary circular reactions*, namely, perception-action loops that arise from an embodied multimodal system behaving in the physical world. Piaget believed that secondary circular reactions are foundational to development. If he is correct, then the building robots that implement human cognition will require the presence of secondary circular reactions throughout their training regimens.

Conclusion

We realize that we have made ambitious requests. We have requested that researchers integrate non-cognitive domains with cognition. We have requested that researchers study the coordination of processes, not just individual processes in isolation. We have requested that researchers study the developmental time course of coordinated processes in situated action and social interaction.

These requests arise from our increasing belief that cognition is more than a collection of independent processing operations. Instead, we believe that cognition, and especially the magic of human cognition, emerges from deep dependencies between all of the basic systems in the brain, including goal management, perception, action, reward, affect, and learning. We also believe that human cognition greatly reflects its social evolution and context, as well as major contributions from a developmental process. Because we believe that human cognition reflects all these dependencies, we believe that it is necessary to change how we study it.

The process of anticipation is a paradigm case for our themes. Anticipations occur across all domains in a highly coordinated manner. Furthermore, anticipations are central to situated action and to social interaction, and grow in sophistication as the result of a developmental trajectory. We believe that our understanding of anticipation will proceed most rapidly if examined from this perspective. We also believe that the results of such study will move cognitive science forward significantly.

References

Ashby, F.G., Ell, S.W., Valentin, V.V., & Casale, M.B. (2005). FROST: A distributed neurocomputational model of working memory maintenance. *Journal of Cognitive Neuroscience, 17*, 1728-1743.

Barrett, L. F. (2006). Emotions as natural kinds? *Perspectives on Psychological Science, 1*, 28-58.

Barrett, L.F., Mesquita, B., Ochsner, K.N., & Gross, J.J. (2007). The experience of emotion. *Annual Review of Psychology, 58,* 373-413.

Barsalou, L.W. (1987). The instability of graded structure: Implications for the nature of concepts. In U. Neisser (Ed.), *Concepts and conceptual development: Ecological and intellectual factors in categorization* (pp. 101-140). Cambridge: Cambridge University Press.

Barsalou, L.W. (1989). Intraconcept similarity and its implications for interconcept similarity. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 76-121). Cambridge: Cambridge University Press.

Barsalou, L.W. (1993). Flexibility, structure, and linguistic vagary in concepts: Manifestations of a compositional system of perceptual symbols. In A.C. Collins, S.E. Gathercole, & M.A. Conway (Eds.), *Theories of memory* (pp. 29-101). London: Lawrence Erlbaum Associates.

Barsalou, L.W. (1999a). Language comprehension: Archival memory or preparation for situated action. *Discourse Processes, 28,* 61-80.

Barsalou, L.W. (1999b). Perceptual symbol systems. *Behavioral and Brain Sciences, 22,* 577-660.

Barsalou, L.W. (2003). Situated simulation in the human conceptual system. *Language and Cognitive Processes, 18,* 513-562.

Barsalou, L.W. (2005). Continuity of the conceptual system across species. *Trends in Cognitive Sciences, 9*, 309-311.

Barsalou, L.W. (in press). Situating concepts. In P.Robbins & M. Aydede (Eds.), *Cambridge handbook of situated cognition.* New York: Cambridge University Press.

Barsalou, L.W., Niedenthal, P.M., Barbey, A., & Ruppert, J. (2003). Social embodiment. In B. Ross (Ed.), *The Psychology of Learning and Motivation, Vol. 43* (pp. 43-92). San Diego: Academic Press.

Bjorklund, D.F., & Pellegrini, A.D. (2000). Child development and evolutionary psychology. *Child Development, 71*, 1687-1708.

Breazeal, C. (2003). Emotion and sociable humanoid robots. In E. Hudlika (Ed.), *International Journal of Human Computer Interaction, 59*, 119-155.

Breazeal C., & Aryananda, L. (2002). Recognizing affective intent in robot directed speech. *Autonomous Robots, 12*, 85-104.

Breazeal, C., Buchsbaum, D., Gray, J., Gatenby, D., & Blumberg, B. (2005). Learning from and about others: Towards using imitation to bootstrap the social understanding of others by robots. *Artificial Life*, *11*, 1-32.

Brooks, R.A. (1991). *Intelligence without representation*. *Artificial Intelligence, 47*,139-159.

Chandrasekaran, B., & Josephson, J. R. (2000). Function in device representation. *Engineering With Computers*, *16*, 162–177.

Campos, J., & Stenberg, C. (1981). Perception, appraisal, and emotion: The onset of social referencing. In M. Lamb & L. Sherrod (Eds.), *Infant Social Cognition* (pp. 273-314). Hillsdale, NJ: Erlbaum.

Clark, A. (1997). *Being there: Putting brain, body, and world together again*. Cambridge, MA: MIT Press.

Damasio, A.R. (1989). Time-locked multiregional retroactivation: A systems-level proposal for the neural substrates of recall and recognition. *Cognition, 33,* 25-62.

Damasio, A.R. (1994). *Descartes' error*. New York: Grosset / Putnam.

Delcomyn, F. (1980). Neural basis of rhythmic behavior in animals. *Science, 210*, 492-498.

Donald, M. (1993). Precis of "Origins of the modern mind: Three stages in the evolution of culture and cognition." *Behavioral and Brain Sciences, 16,* 739-791.

Feinman, S. (1982). Social referencing in infancy. *Merrill-Palmer Quarterly*, *28*, 445-470.

Fernald, A., 1989. Intonation and communicative intent in mother's speech to infants: Is the melody the message? *Child Development 60*, 1497-1510.

Forbus, K. (1993). Qualitative process theory: Twelve years after. *Artificial Intelligence*, *59,* 115–123.

Gentner, D., & Stevens, A.L. (Eds.). (1983). *Mental models*. Mahwah, NJ: Lawrence Erlbaum Associates.

Glenberg, A.M. (1997). What memory is for. *Behavioral and Brain Sciences*, *20,* 1-55.

Gray, J., Breazeal, C., Berlin, M., Brooks, A. & Lieberman, J. (2005). Action parsing and goal inference using self as simulator. In M. Bowerman & S. Levinson (Eds.), *Language acquisition and conceptual development. Proceedings of Fourteenth IEEE Workshop on Robot and Human Interactive Communication (Ro-Man05)* (pp. 111-128). Cambridge: Cambridge University Press.

Granger, R. (2006) Engines of the brain: The computational instruction set of human cognition. *AI Magazine, 27,* 15-32.

Hayes, P.J. (1985). The second naive physics manifesto. In J.R. Hobbs & R.C. Moore (Eds.), *Formal theories of the commonsense world*. Norwood, NJ: Ablex.

Hutchins, E. (1995). *Cognition in the wild*. Cambridge, MA: MIT Press.

Izard, C. (1977). *Human emotions*. New York: Plenum Press.

James, W. (1890). *The principles of psychology*. Mineola, NY: Dove.

Jones, S. (2006) Infants learn to imitate by being imitated. *5th Proceedings of the International Conference of Development and Learning.*

Jones, S.S., & Smith, L.B. (1993). The place of perception in children's concepts. *Cognitive Development, 8*, 113-139.

Kintsch, W., & van Dijk, T.A. (1978). Toward a model of text comprehension and production. *Psychological Review, 85,* 363-394.

McClelland, J.L., Rumelhart, D.E., & the PDP Research Group (1986). *Parallel distributed processing: Explorations in the microstructure of cognition: Vol. 2. Psychological and biological models*. Cambridge, MA: MIT Press.

Meltzoff, A., & Moore, M.K. (1997). Explaining facial imitation: A theoretical model. *Early Development and Parenting*, 6, 179-192.

Meltzoff, A.N. (1996) The human infant as imitative generalist: A 20 year progress report on infant imitation with implications for comparative psychology. In C.M. Heyes & B.G. Galef (Eds.), *Social learning in animals: The roots of culture*. New York: Academic Press.

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

Niedenthal, P.M., Barsalou, L.W., Winkielman, P., Krauth-Gruber, S., & Ric, F. (2005). Embodiment in attitudes, social perception, and emotion. *Personality and Social Psychology Review, 9,* 184-211.

O'Regan, J.K. & Noe, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences, 24*, 939-1011.

Piaget, J. (1952). *The origins of intelligence in children*. New York: International Universities Press.

Plutchik, R. (1991). *The emotions.* Lanham, MD: University Press of America.

Port, R., & van Gelder, T. (Eds.). (1995). *Mind as motion: explorations in the dynamics of cognition*. Cambridge, MA: MIT Press.

Rips, L. J. (1994). *The psychology of proof: Deduction in human thinking.* Cambridge, MA: MIT Press.

Rizzolatti, G., Fadiga, L., Fogassi, L., & Gallese, V. (2002). From mirror neurons to imitation: Facts and speculations. In A.N. Meltzoff & W. Prinz (Eds), *The imitative mind: Development, evolution, and brain bases. Cambridge studies in cognitive perceptual development* (pp. 247-266). New York: Cambridge University Press.

Robbins, P., & M. Aydede (Eds.) (in press). *Cambridge handbook of situated cognition*. New York: Cambridge University Press.

Rumelhart, D.E., McClelland, J.L., & the PDP Research Group (1986). *Parallel distributed processing: Explorations in the microstructure of cognition: Vol 1. Foundations*. Cambridge, MA: MIT Press.

Samuelson. L.K., & Smith, L.B. (2000). Children's attention to rigid and deformable shape in naming and non-naming tasks. *Child development, 71*, 1555-1570.

Schank, R.C., & Abelson, R.P. (1977). *Scripts, plans, goals, and understanding: An inquiry into human knowledge structures.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Smith, L.B. & Gasser, M. (2005). The development of embodied cognition: Six lessons from babies. *Artificial Life*, 11, 13-30.

Smith, E.R., & Semin, G. R. (2004). Socially situated cognition: Cognition in its social context. *Advances in Experimental Social Psychology, 36*, 53-117.

Smith, L.B., & Thelen, E. (1993). (Eds.). *A dynamical systems approach to development: Applications.* Cambridge, MA: MIT Press.

Spivey, M. (2006). *The continuity of mind.* New York: Oxford University Press.

Strack, F., Martin, L. L., & Stepper, S. (1988). Inhibiting and facilitating conditions of the human smile: A nonobtrusive test of the facial feedback hypothesis. *Journal of Personality and Social Psychology, 54,* 768-777.

Thelen, E., & Smith, L.B. (1994). *A dynamic systems approach to the development of cognition and actions*. Cambridge, MA: MIT Press.

Thomaz, A.L., Berlin, M. & Breazeal, C. (2005). An embodied computational model of social referencing. Proceedings of Fourteenth IEEE Workshop on Robot and Human Interactive Communication (Ro-Man05).

Tomasello, M., Kruger, A., & Ratner, H. (1993). Cultural learning. *Behavioral and Brain Sciences, 16*, 495-552.

Author Notes

DARPA contract BICA FA8650-05-C-7255 awarded to three authors supported their collaboration on this project. In addition, Lawrence Barsalou was supported by DARPA contract FA8650-05-C-7256 and National Science Foundation Grant BCS-0212134. Address correspondence to Lawrence W. Barsalou, Department of Psychology, Emory University, Atlanta, GA 30322, barsalou@emory.edu, http://www.psychology.emory.edu/cognition/barsalou/index.html).