

## Introduction

This book is about contemporary philosophical and neuroscientific perspectives on action, perception, and cognition as they are lived in embodied and socially embedded experience. This emphasis on embodiment and embeddedness is a change from traditional theories, which focused on isolated, representational, and conceptual cognition. In the new perspectives contained in our book, such “pure” cognition is thought to be undergirded and interpenetrated by embodied and embedded processes.

This new perspective on what John Dewey called “lived experience” (1916) has emerged from a number of sources detailing how we navigate complex social environments (e.g., Barton, 2004; 2006; Johnson, 2007). We also see research into the diverse cognitive systems that organize action (Gallistel, 1980; 1992; von Holst, 1973) and underlie perception (Lakoff and Johnson, 1999; Schulkin, 2000).

What’s important and new is that these systems are also now seen to be rich in bodily sensibility (Gallagher, 2005; Johnson, 2007; Schulkin, 2004; Varela et al., 1991); the perceptual/conceptual split is held to be quite porous (Dewey, 1925/1989) and cognitive systems are not divorced from action or perception, but are endemic to them (Barton, 2004; Berthoz, 2000; Dewey, 1896; Gigerenzer, 2000; Peirce, 1878; Schulkin, 2000; 2007). This enables us to investigate the biological derivation of the cognition/action/perception nexus.

Lakoff and Johnson (1999) have nicely depicted the relation of perception and action:

Our Introduction will guide us through six aspects of this new research: 1) evolution and the brain; 2) tool use; 3) social cognition; 4) the extended mind; 5) the enactive approach; and 6) the bounded yet self-corrective nature of human inquiry.



## 2 *Introduction*

*Table I.1*

Thinking is perceiving	Imagining as moving
Knowing is seeing	Attempting to gain knowledge is searching
Representing as doing	Becoming aware is noticing
Communicating as showing	Impediments to knowledge are impediments to vision
Searching as knowing	Knowing from a “perspective” is seeing from a point of view

### Evolution and the Brain

New theories of brain evolution stress communication and sociality as essential to our capacity to represent objects as intersubjectively accessible. How did we grow as a species to be able to recognize objects as common, as that which can also be seen in much the same way by others? Such constitution of intersubjectively accessible objects is bound up with our flexible and sophisticated capacities for social cognition – understanding others and their desires, intentions, emotions, and moods – which are crucial to the way human beings live.

Evolution shapes all life processes; the organization of the brain is no exception. Brains vary between species and between taxa, so to grasp the evolution of the human brain requires attention to the nervous system, its core outline, variation, and common themes (e.g., Swanson, 2000).

The nineteenth-century British neurologist J. Hughlings Jackson (1884/1958) formulated a conception of the nervous system that places the development of the brain within evolution; for our species, the neocortex is a crowning achievement, increasing the possible range of action. This traditional view made the cortex the only cognitive part of the brain (e.g., Jackson, 1884/1958; Jackson and Decety, 2004; James, 1890/1917; Swanson, 2000; 2003). However, cognitive systems run across the brain; from cortex to brainstem, the central nervous system is knotted to cognitively rich information processing resources (Barton, 2004; Schukin, 2004; 2007).

The brain contains many rich information processing systems that reflect an expanding cognitive capability. For instance, a broad repertoire of social perception and action capacities is codified across diverse regions of the brain (Decety and Jackson, 2006; Jeannerod, 1997; Rizzolatti and Lupino, 2001). An elaborate set of neural structures widely distributed throughout the brain is linked to keeping track of others, watching what they do, and getting a foothold in the world

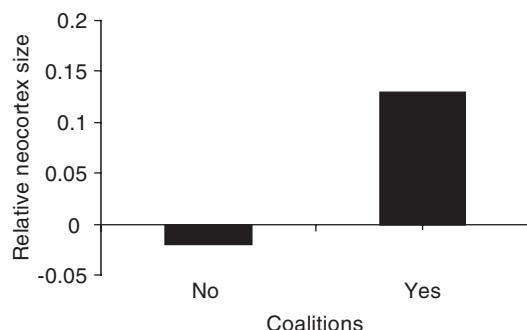
AQ2



of approachable and avoidable events (Jacob and Jeannerod, 2003; Schukin, 2000). We are visual animals, and the expansion of the visual primary cortex reflects our social embeddedness and the expansion of our options (Barton, 2004; 2006; Dunbar and Shultz, 2007).

Social objects appear in diverse contexts (Levinson, 2006) and diverse preadaptive cephalic systems are expanded in use [Mayr, 1942/1982; Rozin, 1976; 1998]. The automatic perception of events and the orchestration of action vital for diverse social behaviors have long been noted (Dewey, 1896; 1925/1989). A large body of evidence links the degree of social interaction with neocortical expansions. Diverse models of group size have been linked to neocortical enlargement (Dunbar, 1998; Dunbar and Shultz, 2007) and cognitive competence (Barrett et al., 2003; Byrne and Corp, 2004) across distributed cognitive systems (Barton, 2004; 2006; Schukin, 2007). The greater the degree of social contact and social organization experienced by diverse primates, and particularly our species, the greater the trend for cortical expansion (Barton, 2006; Dunbar, 1998; Dunbar and Shultz, 2007) (see Figure I.1).

We can conclude that we evolved under pressure for physiological cognitive systems that are oriented to social systems. Their evolution and expression underlie the diverse forms of complicated social assessments; group size, for instance, is correlated with cortical expansion (Dunbar and Shultz, 2007). Consider the complex social relationships of primates, their hierarchy, distribution of food resources, shelter, protection, dominance, and comfort through co-alliances. Such systems are quite varied, and all involve cephalic innervations and expression.



*Figure I.1* Social contact in primates is consistently linked to neocortical expression and size (adapted from Dunbar and Shultz, 2007).

## 4 *Introduction*

Social behavior is just one element in our evolutionary ascent, however; structural changes in our visual system and bipedalism are also most definitely core features in our evolutionary ascent. We also see an evolved motor system in which subcortical brain regions (e.g., basal ganglia, Lieberman, 2000; 2002; Rizzolatti and Luppino, 2001; Ullman, 2004) are pregnant with cognitive capacities. Researchers now recognize that diverse regions of the cortex are tied to motor function (Berthoz, 2000; Jackson, 1884/1958; Jeannerod, 1997) and the recognition of motor regions being rich with cortical functions is growing (Barton, 2004; Jackson and Decety, 2004; Rizzolatti and Lupino, 2001).  
AQ4

Many species use diverse tools in adapting to their environment, and so we see the evolution of cortical and subcortical systems in the brain that participate in tool use, tool-making and tool recognition (Gibson and Ingold, 1993; Johnson-Frey, 2004; Martin, 2007). These diverse and expanding cognitive systems are prevalent in motor regions of the brain (Barton, 2004; Jackson and Decety, 2004; Schukin, 2007) and are fundamental to our evolutionary ascent. Tool use is an expression of an expanding cortical motor system in which cognitive systems are endemic to motor systems (e.g., Martin, 2007; Ullman, 2004).  
AQ5

## Cephalic Capabilities for Action and Tools

We are a highly linguistic tool-using species with advanced social abilities and an elaborate diversity of cognitive skills (Gibson and Ingold, 1993). Many forms of instrumental expression and tool use have arisen in our history, as have diverse forms of cognitive adaptations, some broad and some specific; cognitive expansion and fluidity, as we move from narrow adaptive abilities to expanding use across diverse problematic context, is a feature of the human mind (e.g., speed vs. diversity, flexibility and expansion of responses) (Donald, 1991; Mithen, 1996; Rozin, 1976; 1998).

Preadaptive cognitive functions are expanded in use, in diverse and broader contexts; the greater the connectivity, the less modular the access and the greater the expansion (Rozin, 1976; 1998). Key features of this cognitive fluidity are the integration of several orientations to coping with the world: social intelligence, technical abilities, diverse expression of natural knowledge, and, of course, language use (Geschwind, 1974; Mithen, 1996; Pinker, 1994; Rozin, 1976; 1998).

We should never forget that we are linguistic animals. Language, with its rich syntactical elegance and powers of expansion, is a unique feature of our species. Once syntactical language use emerged, our cognitive

# PROOF

Introduction 5

Table I.2 (Adapted from Mithen, 1996)

- Natural history
- Social intelligence
- Natural history intelligence
- Technical intelligence
- Language
- General intelligence

abilities increased in great measure (Levinson, 2006; Pinker, 1994), particularly our social discourse (Mellars, 2006). The focus of our evolution is on our social contact and discourse, our communicative competence and praxis (Donald, 1991; Tomasello et al., 2004). This social contact and preadaptation, in turn, are vital for the formation of basic regulatory events that traverse a wide range of rewards within the behavioral biology and the central nervous system.

Cognitive/behavioral capacities are, after all, ways to engage the world (Dewey, 1925/1989; Gigerenzer, 2000); we compute probability and assess friendly or nonfriendly events (Gallistel, 1992). The representations are often social in nature, about others (Moreno, 1995; Sabini and Schulkin, 1994).

As a species, we are rooted with others and what they do. But representations of others, of those that we care about, do not divide us from others; rather, they guide the organization of cognition, which is necessary for the organization of action (Gallistel, 1980). This is important when we consider others, their beliefs and desires, the way they are oriented, to what they are oriented, learning from them, the tools they use, and so on.

One fundamental feature of our species is that our social discourse, social transactions, language, eating and tool use are embedded in the social milieu and not divorced from them; they evolved together in the more elaborate context of remaining in touch and of keeping track of one another.

Language and checking in on others, amid an endless array of social transactions (Dunbar, 1998; Mithen, 1996; Levinson, 2006), are recurrent and evolving features of our experiences. Of course, the expansion of our motor capacity and our ability to see greatly coevolved with these social capacities, which are perhaps best seen as a linguistic expansion of grooming behaviors and larger social behaviors, tied to corticalization of function.

The motor regions are gravid with cortical functions (Barton, 2004; Jackson and Decety, 2004). As we noted above, many species use diverse

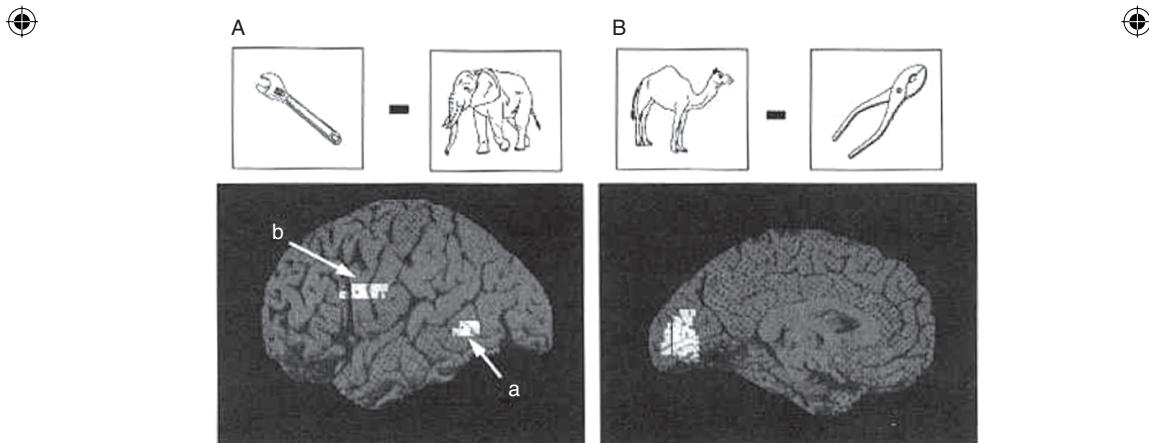
6 *Introduction*

tools in adapting to their environment, capacities enabled by the evolution of cortical and subcortical systems in the brain that participate in tool use, tool-making, and tool recognition (Gibson and Ingold, 1993; Martin, 2007).

Tool use is an expression of an expanding cortical motor system in which cognitive systems are endemic to motor systems (e.g., Lieberman, 2000; Martin, 2007; Ullman, 2004) and in which premotor regions of the neocortex play diverse preparatory functions (Gallese et al., 1996; Passingham, 2008). Moreover, regions of the brain are prepared to recognize differences between different kinds of objects, one being mechanical tools (Martin, 2007; Martin et al., 2000). It is also important to see that frontal motor regions are linked to the motor features of tool use (Johnson-Frey, 2004; Martin, 2007; see Figure I.2).

It is the expansion of cephalic function that underlies the tool use that serves physiological/ behavioral regulation. An expanded motor system with diverse cognitive capacities no doubt is pivotal in our evolutionary ascent.

Tool use requires an evolving cognitive/motor system that underlies perception and action such that tool use enhances many aspects of brain



**AQ6** *Figure I.2* (A) View of the left side of the brain showing areas in the left posterior lobe (a) and premotor cortex (b) that were more active when subjects silently named pictures of tools than when they silently named pictures of animals. (B) View of the inner (medial sulcus of the left side of the brain showing the region of the occipital lobe that was more active when subjects silently named pictures of animals than when they silently named pictures of tools (Martin et al., 1996).

# PROOF

## Introduction 7

function (Heelan, 1983/1988; Heelan and Schulkin, 1998; Jackson and Decety, 2004; Johnson, 2007). What emerged in us via the tool/action/perception/communication nexus, ultimately, were improved methods of passing information to others, including other generations.

Tool use and brain expansion are coevolutionary phenomena. Cephalic expansion set the stage for technological creations, expanding our sensory systems and mnemonic capacities (Clark, 1997; 2008; Donald, 1991; Heelan and Schulkin, 1998). Seeing by magnifying became an evolving theme as our capacities were extended and we turned from managing nature towards understanding nature; tool-making was critical for this development (Foley, 1998) (see Figure I.3).

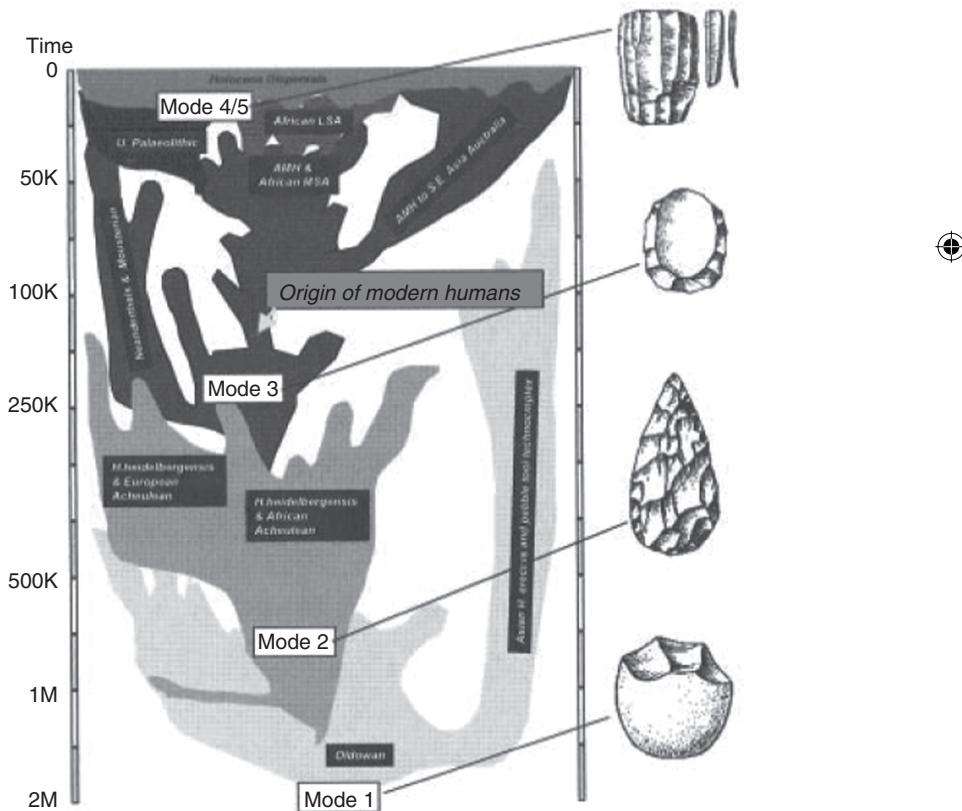


Figure I.3

## 8 *Introduction*

### Social Cognition: Keeping Track of Others

Our brains are prepared to recognize animated objects, motion and action; the detection of motion and our sense of being a causal agent are embedded in the concept of agency (e.g., Premack, 1990; Whitehead, 1927/1953). Moreover, the detection of movement is knotted to visual cognition and is an important adaptation in understanding others (e.g., Frith and Wolpert, 2003; Gallese, 2007; Jacob and Jeannerod, 2003; Martin, 2007). The brain comes prepared to discern, or at least try to discern, such social relationships.

Cognitive systems are embedded in sensorimotor explorations (e.g., Clark, 1997; 2008). Moreover, one could suggest quite reasonably that the imitative processes seen in neonates are an elementary form of learning (Meltzoff and Moore, 1977). Infants do not just track events in their environment; they imitate facial expressions. Infants change from simply following others to engaging others, learning from others, and eventually challenging others; they eventually generate hypotheses in an expanding, self-corrective process (Gallagher and Meltzoff, 1996; Meltzoff, 2007).

This ability of “perspective taking” is an evolved central state (see AQ7 Figure 7). It is an active state in the consideration of the experiences of others, and is tied to communicative competence, essential human bonds, and well-being (e.g., Baron-Cohen, 1995/2000; Decety and Jackson, 2006). Recognizing the intentions of others is one critical feature in prosocial behaviors. This cognitive capacity begins early in ontogeny and is tied in to visual sensibility (Premack, 1990; Tomasello and Carpenter, 2007). Recognizing the intentions of others is knotted to a broad array of cephalic tissue that underlies perspective taking (Lamm et al., 2007) and human social judgment (Adolphs, 1999; Greene and Haidt, 2002; Moll and Schulkin, 2009).

While perception and action are represented in similar regions of the brain, they are not identical (Berthoz, 2000; Jeannerod, 1999). But embedded in perception and action is the commonsense justifiable assumption of something shared, of forms of bodily experience, of cognitive systems embedded in the sensorimotor experiences (e.g., Lakoff and Johnson, 1999; Merleau-Ponty, 1962/1970; Schulkin, 2004; Varela et al., 1991; Wilson and Knoblich, 2005) – what Dewey called “lived experience.”

The imagery of action – watching others, and getting anchored to a world in which perception and action are linked – is fundamental (Jackson and Decety, 2004). Looking at action words (Hauk et al., 2004;

# PROOF

Introduction 9

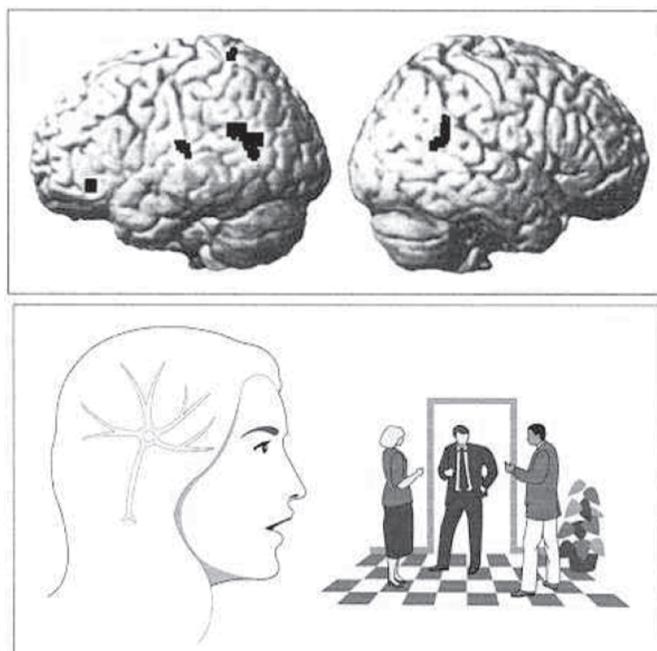


Figure I.4

Pulvermuller et al., 2005), and the performance of an action, activate many of the same regions of the brain (e.g., Martin et al., 2000; Perani et al., 1995). The coherence in the organization of the brain is the tight link between cognitive systems and action and function (Handy et al., 2003).

Representations of action words and category-specific prepared knowledge about objects linked to action show that cognitive systems traverse most, if not all, regions of the brain (Jackson and Decety, 2004; Schulkin, 2004). Importantly, these areas of the brain underlie the adaptation of shared experience, a sense of community (Dewey, 1925/1989; Flanagan, 2007) and meaning (Jaspers, 1913/1997). Embodied cognitive systems are expressed in our human experience (e.g., Clark, 1997; Gallagher, 2005; Lakoff and Johnson, 1999; Varela et al., 1991; Wheeler, 2005), and they underlie a multitude of forms of practice across human symbolic expression (Bourdieu, 1980/1990).

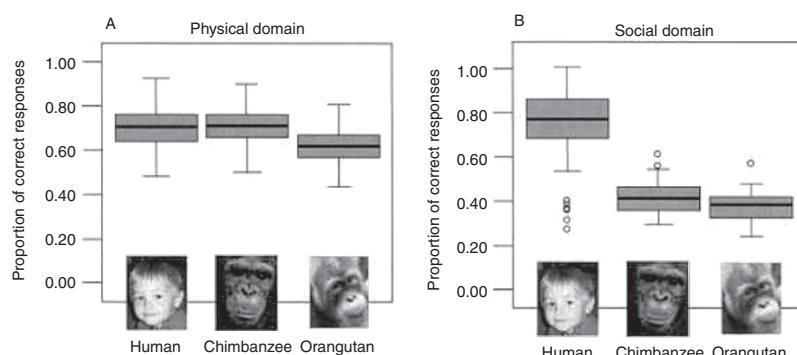
10 *Introduction*

The focus of our evolution is on our communicative abilities (Donald, 1991; Tomasello et al., 2004). The representation of objects paralleled a desire to depict them in diverse ways, particularly those important to our everyday sustenance; depicting and predicting edible objects is an important adaptation. A cognitive capacity to represent objects of diverse kinds is central to the evolution of the primate brain, particularly in our species.

Cognitive categories figure in our recognition of social and live objects (Tomasello et al., 2004). There is much unresolved debate with regard to the range, innateness, and developmental expression of these capabilities. What is not debated is the fact that they are anchored to our social milieu (getting oriented to others) and to the ecological and social surroundings.

Thus, children are quite similar to the common chimpanzee or orangutan in the first few years of development when it comes to nonsocial problem-solving (Herrmann et al., 2007). For instance, when given problems concerning objects in space, the responses of very young humans, chimpanzees, and orangutans look similar to each other (Figures I.5A and I.5B), though the proportion of social responses was different. What becomes quite evident early on in ontogeny is the link to the vastness of the social world in which the human neonate is trying to get a foothold for action and problem-solving capabilities (Herrmann et al., 2007).

As we have noted above, cognitive/behavioral capacities are ways to engage the world (Casebeer, 2003; Dewey, 1925/1989; Sterelney, 2007).



*Figure I.5* Young children's and two related primates' solution to both simple physical and social problems. Physical domain (A) and social domain (B). The box plots show the full distribution of the proportion of correct responses for physical and social domains for each species (from Herrmann et al., 2007).



to compute probability, and to assess friendly or nonfriendly events (Gallistel, 1992). These representations are often social in nature; they are typically about how to associate with or avoid others (Levinson, 2006; Moreno, 1995).

“Representation” often has the connotation of something divorced from the object, something that detaches. Of course, some forms of representations are like that, but many are not (Clark, 2008; Dewey, 1896; Heelan and Schulkin, 1998; Johnson, 2007), especially when it comes to our social existence. We are interested in others and what they do; our representations of them are not simply divorced pictures from a distance, they are active ways in which we engage the world, so much so that human flexibility, survival and reproductive success are all tied to the social world of others (Mead, 1934)...

In other words, representations of others do not divide us from others. Instead, they guide the organization of cognition, and cognitive or informational systems are inherent to the organization of action (Gallistel, 1980). This is important when we consider others, their beliefs and desires, the way they are oriented, to what they are oriented, the tools they use, and so on.

### The Extended Mind: Expanding Cephalic Capabilities

Human cognitive capacities include the ability to offload, and then to access, information stored in the environment. Clark and Chalmers called this the “extended mind” (Clark, 1997; 2003; Clark and Chalmers, 1998). Merlin Donald’s term “exogram” is an apt term for such environmental memory.

Humans are endlessly “bootstrapping” onto external objects to enrich our memory capacities. Some features of internal memory include what Karl Lashley (1951) called the “engram,” and external memory, what Merlin Donald (1991) called the “exogram” (Table I.3).

We unload memory storage onto our surroundings, and into the spaces we inhabit (Clark, 1997; 2008). As we decrease our cognitive load, we make what is around us familiar. The exogram is a metaphor for this. It is an important adaptation that, after learning something, we can externalize it and make it part of the everyday (e.g., Noë, 2004; Rowlands, 2010; Schulkin, 2009; Wheeler, 2005).

One adaptive feature in our evolution is the extent to which, after we learn something, we create cues in the environment to activate systems for performance (e.g., recognition of an event, person, or how to do something). Memory is, thus, not strictly in the head. Our cognitive



# PROOF

## 12 *Introduction*

Table I.3 Properties of internal and external memory compared. (Donald, 1991)

Internal Memory Record (engram)	External Memory Record (exogram)
Fixed physiological media	Virtually unlimited physical media
Constrained format, depending on type of record, and cannot be reformatted	Unconstrained format, and may be reformatted
Impermanent and easily distorted	May be made much more permanent
Large but limited capacity	Overall capacity unlimited
Limited size of single entries (e.g., names, words, images, narratives)	Single entries may be very large (e.g., novels, encyclopedic reports, legal systems)
Retrieval paths constrained; main cues for recall are proximity, similarity, meaning	Retrieval paths unconstrained; any feature or attribute of the items can be used for recall
Limited perceptual access in audition, virtually none in vision	Unlimited perceptual access, especially in vision
Organization is determined by the modality and manner of initial experience	Spatial structure, temporal juxtaposition may be used as an organizational device
The “working” area of memory is restricted to a few innate systems, like speaking or subvocalizing to oneself, or visual imagination	The “working” area of memory is an external display which can be organized in a rich 3-D spatial environment
Literal retrieval from internal memory achieved with weak activation of perceptual brain areas; precise and literal recall is very rare, often misleading	Retrieval from external memory produces full activation of perceptual brain areas; external activation of memory can actually appear to be clearer and more intense than reality

world is rich with exograms that serve to guide behavior and that are continuous with endogenous cognitive mechanisms; the important point is to store less and recognize and utilize more. Less can be more; adaptation is the key factor in the organization of action.

### **Lived Experience and the enactive approach to cognition**

The “enactive” approach to cognition, and its relation to the pragmatist tradition, is another trend with which our book deals. Enactive approaches enable us to see how action and perception are linked in sensorimotor systems, which are in turn linked with cultural symbol systems. Both action–perception and cultural symbol systems come together to shape our spatiotemporal experience, our relation to the short, mid, and long-term past and future.

# PROOF

## Introduction 13

From an enactive perspective, we are neither Cartesian machines thinking in an abstract divorced vacuum nor random inductive machines. We bring with us diverse forms of cognitive devices that underlie Dewey's "lived experiences" (1916) or what others have called "embodied cognition" or "enaction" (e.g., Gallagher, 2005; Lakoff and Johnson, 1999; Schulkin, 2004; Stewart et al., 2011; Varela et al., 1991).

Our sense of self is rooted in our life histories, which are built up from patterns of action and movement. The relation of lives, purposes, action, and motion is depicted below in Figure I.6 adapted from Johnson (1993).

In the enactive model, thinking is understood in the context of actions with others, and is quite close to a pragmatist position in which cognitive systems are embedded in the organization of action (see also Dewey, 1896; James, 1890/1917; Schulkin, 2009). The emphasis is on embodied, expanded cognitive systems (Barsalou et al., 1995; Clark, 1997; 2008; Noë, 2004; Wilson, 2005); the sensorimotor systems are themselves knotted to functions across all regions of the brain (see also, e.g., Barton, 2004; Berthoz, 2000; Dewey, 1896; Jeannerod, 1997; Schulkin, 2007). In other words, cognitive systems are not just a cortical affair, but are endemic to widespread brain function, though that certainly does not mean that cognitive systems are exactly the same as motor systems; they are not (Jacob and Jeannerod, 2005; Schulkin, 2007). "Embodied" cognition is more a way of understanding ourselves in the context of action and perception, "lived experience" as Dewey

AQ9

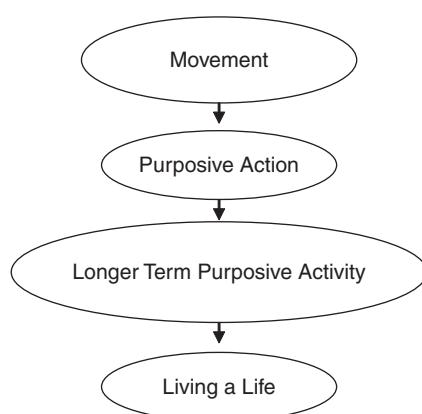


Figure I.6

# PROOF

## 14 *Introduction*

put it, than a settled suggestion across all the many cognitive systems that are operative (cf. Gibbs, 2006; Goldman and de Vignemont, 2009; Grafton, 2009; Mahon and Caramazza, 2008; Wilson and Knoblich, 2005).

The construction of objects shapes the way we orchestrate action; the shared experience with one another is a fundamental adaptation [Gallagher, 2005; Tomassello and Carpenter, 2007] and the cephalic innervations of bodily sensibility are a pervasive common element in our evolution (e.g., Berthoz, 2000; Damasio, 1994; James, 1890/1917; Merleau-Ponty, 1962/1970; Schulkin, 2004).

The use of our cognitive arsenal with the onset of memory systems external to us resulted in greater motor and cognitive flexibility. The expansion of symbols culminated in theoretical speculations (Donald, 1991; 2004). The onset of culture expanded memory beyond our internal processes. But, from the onset, memory has been anchored to our sense of objects, the adaptive systems that tie us to the discernment of events, our sense of action and embodied existence (Gibbs, 2006; Glenberg, 1997; [Glenberg and Kaschar, 2002]). Context pervades memory fields (Clark, 1997; Gibson, 1966; Glenberg, 1997; Glenberg and Kaschar, 2002).

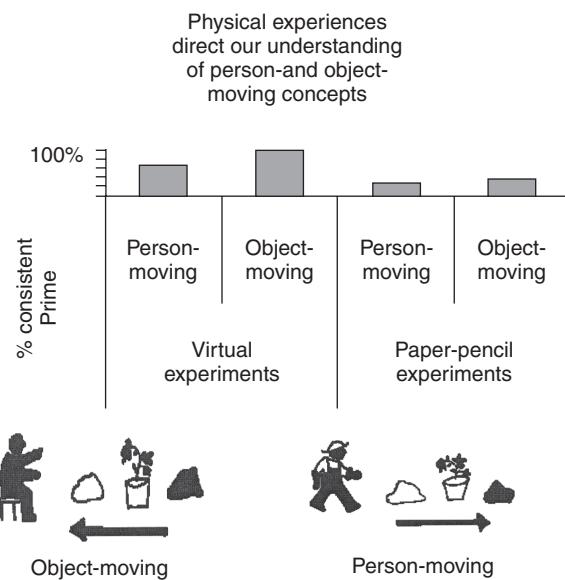
Experience and language influence our sense of moving through space and of projecting ahead. That is, one's sense of time is influenced not only by space, but also by movement (Alloway et al., 2006; Levinson, 1996). Perhaps not surprisingly, our sense of motion, or physical sensibility, impacts our sense of time and space (see Figure I.7). So, there is a bodily component of our sense of time and memory of events, which is rich in sensorimotor experiences (Boroditsky and Ramscar, 2002; Glenberg, 1997), as well as agency and action (Boroditsky, 2000; Matlock et al., 2004) – and the experiences are smoothly automatic (Bargh and Chartrand, 1999; Knoblich and Sebanz, 2006).

Our evolved brain and cognitive predilections are oriented towards context and flexibility. Our embodied perception of objects is conceptually rich; this is vital for the behavioral adaptation of action, perception, and the brain (cf. Barsalou, 1999; Clark, 2008; Jacob and Jeannerod, 2003; Johnson, 2007). The emphasis of much research is on the adaptive nature of these systems (Dewey, 1896; Gigerenzer, 2000; Schulkin, 2009).

The lenses with which we interpret the world and make inferences are oriented to events (Hansen, 1958/1972; Heelan and Schulkin, 1998; Peirce, 1899/1992) and the orientation to objects is part of the adaptive specializations for action.

# PROOF

Introduction 15



AQ14 *Figure I.7* Subjects in a virtual-reality computer environment were likely to act consistently with a spatial prime (Alloway et al., 2006).

We come into this world prepared to learn, inquire, and theorize. The range of hypotheses one can create is constrained by abduction, that is, by seeing an object as something already presupposed by some context of understanding (e.g., Hansen, 1958/ 1972; Lakoff and Johnson, 1999). Abduction highlights central tendencies about object relations, for instance, and inductive inferences are placed in a warranted context (Barsalou, 1999; Levinson, 1996).

Our inferences are constrained by an orientation to events and by the kinds of objects that are detected. For example, categorizing mammals as viviparous and finding what seem like counterexamples (such as the platypus, an egg-laying mammal) first requires a broad way to link diverse kinds of events, which may (perceptual) or may not (conceptual) have clear common properties (Barsalou, 1999).

The taxonomic and thematic conditions to categorization may be simple or complex, but there is always a background cognitive framework (e.g., Levinson, 2006; Murphy, 2002). Moreover, the inductive devices are broadly conceived in a mind/brain ready to compute statistical probability, draw diverse inferences, and construct models (Johnson-Laird, 2001) essential for information processing and coherent action. These

## 16 *Introduction*

cognitive events are apparent in taxonomic organization in diverse human societies about basic objects (e.g., plants, animals, etc.; Carey, 2009).

### An Evolving and Shared Sense of Human Inquiry

Epistemic diversity, like biological diversity, is both the reality of the scientific enterprise and an adaptation of it. A combination of instruments, experiments, and theory cuts across the range of inquiry (Dear, 2006; [Galison, 1988; 1999] Shapin, 1996). Today's scope of scientific knowledge is far too vast for one person to embrace completely, hence the emergence of specialties and subspecialties in science and medicine. In an era of exponential leaps in scientific knowledge, the multiple levels of analysis required in the scientific enterprise necessitate that we work together, because the knowledge of each level is not – indeed, cannot possibly be – shared by all members of the group.

So, in addition to inherent competition and collaborative bonds, the recognition of others and our link to them are essential to advances in knowledge. As C. S. Peirce (1898/1992) would say, a “laboratory frame of mind” is at the heart of this endeavor. Peirce, an essential pragmatist thinker, established a philosophy of self-correction and set up the first experimental laboratory in psychophysics in America. He understood that all experience is embedded in practices that are rich with frames of reference (e.g., Hanson, 1958/1972; Heelan and Schulkin, 1998); as we perceive the world around us, we presuppose background sets of inferences as we are embedded in diverse social practices.

Any form of inquiry needs to be self-corrective. Assumptions must be questioned; biases permeate any endeavor and recognition of them is essential to self-correction. Moreover, the recognition of biases in human decision-making functions by what social scientists call “sensitizing concepts,” which provide a context for decisions. Some of these biases are simple, rough heuristics that work well, have been selected by evolutionary factors, and are part of our cognitive apparatus. Multiple cognitive systems operate across the many areas of inquiry that we pursue (see Rozin, 1976; Simon, 1962; 1982).

Research in the cognitive sciences reinforced a shift to the logic of heuristics in reasoning, or what Herb Simon (1962; 1982; see also Gigerenzer and Selten, 2001) called “satisficing” – what has also been called “bounded rationality” (see Figure I.8). All decision-making reflects the frameworks out of which one interprets one's world, the lifeworld in which one participates. The ideal of perfect reason inherited from the

AQ15



# PROOF

Introduction 17

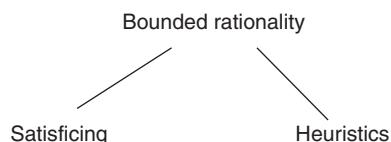


Figure I.8

Enlightenment (Israel, 2001) is replaced with demythologized reason, problem-solving rooted in human experience; looking at everyday life, we see existential and pragmatic – less than perfect – decision-making (Gigerenzer and Brighton, 2009; Gigerenzer and Selten, 2001).

The demythologizing of decision-making, whether within philosophy or within cognitive sciences, requires that one recognize both that science works and that there are advances to be proud of, and that real-life human reasoning, scientific or otherwise, is replete with flaws. The realization that optimization is not a pedestal or Archimedean point on which one can stand for all time is not demoralizing – quite the contrary (Gigerenzer and Selten, 2001; Kahneman et al., 1982; Simon, 1962; 1982). Advances are made by self-correction; explanations evolve, get better, or are discarded (Kitcher, 1993).

Diverse embodied cognitive systems are inherent in the organization of action. Our cognitive evolution puts an extra premium on memory, language, and diverse tools for problem-solving. Our cultural evolution draws on all resources, as our memory and communicative systems are both internal and external to us (Clark, 1997; Donald, 1991). John Dewey (1916), well appreciated at the beginning of the last century, emphasizes that the experience of the knower is essentially embodied in ecologically rich and historical diverse social practices (Galison, 1988).

Cognitive systems are endemic to all human action, and codified frames endlessly and effortlessly make up the tapestry that we live in. We take the social milieu and the rich sources of information for granted, but they are bodily understood; cephalic sensibility, haunting spaces, and place, as Merleau-Ponty (1962/1970) suggested, are pervasive facts about us. Minds are embodied and embedded in practices of coherence and adaptation within the “affordances” of direct action (Gibson, 1966) with objects and with others; we are endlessly and easily scaffolding upon practices that are already in place in the culture that surrounds us (Clark, 1997; 1999; 2003; 2008; Donald, 1991; 2004; Noë, 2004; Wheeler, 2005).

# PROOF

## 18 *Introduction*

These are fundamental adaptations that underlie the organization of action. Philosophy, cognitive science, and the broader sense of cephalic function are continuous; each involves satisficing and contextual problem-solving, using heuristics that underlie our rapid ability to grasp the gist of actions we need to do. Amidst the diverse readily available resources and the diverse forms of problem-solving are also abductive moments of putting events together into forms of understanding through the genesis of ideas, which are a fundamental part of the human experience and human adaptation and understanding.

### Forecast of the book

The logic of the chapters is as follows. The first three chapters are on evolution, cognition, and adaptive systems. The next three deal with the philosophical understanding and perplexities of human cognitive capability and bodily expression. The final three focus on cognitive, neural, and pragmatic/hermeneutical issues that underlie human social contact and meaning.

Chapter 1: Merlin Donald, Department of Psychology, Queens University: A Hypothetical Cognitive Adaptation for Distributed Cognitive Networks. Donald will describe the evolution of memory systems in the context of action, perception, and memory.

Chapter 2: Susanne Shultz and Robin Dunbar, Department of Evolutionary Anthropology, Oxford University: Social Cognition and Cortical Function: an Evolutionary Perspective. Shultz and Dunbar will describe the evolution of cortical and subcortical structures that underlie social adaptation.

Chapter 3: Henry Brighton and Gerd Gigerenzer, Max Planck Institute for Human Development: *Homo Heuristicus* and the Bias–Variance Dilemma. Brighton and Gigerenzer demythologize human reasoning and place problem-solving in an adaptive context.

Chapter 4: Mark Johnson, Department of Philosophy, University of Oregon: Action, Embodied Meaning, and Thought. Johnson depicts the diverse cognitive events that underlie human purpose and action.

Chapter 5: Shaun Gallagher and Katsunori Miyahara, Department of Philosophy, University of South Florida: Neo-Pragmatism and Enactive Intentionality. Gallagher describes the events that underlie intersubjective discourse.

Chapter 6: Michael Wheeler, Department of Philosophy, University of Stirling, UK: Minds, Things, and Materiality. Wheeler engages some of the philosophical issues that surround the extended mind hypothesis.

AQ16



# PROOF

Introduction 19

- AQ17 Chapter 7: Arthur Glenberg, Department of Psychology, University of Wisconsin: Contributions of Mirror Mechanisms to the Embodiment of Cognition. Glenberg presents evidence of embodied cognitive systems tied to mirror neurons.
- AQ18 Chapter 8: Sébastien Hétu and Philip Jackson, Departments of Psychology, University of Laval and University of Chicago: The Neural Systems Involved in Motor Cognition and Social Contact. Hétu and Jackson will describe the neural systems that underlie social contact and action.
- AQ19 Chapter 9: Jay Schulkin and Patrick Heelan, Departments of Neuroscience and Philosophy, Georgetown University: Action, Perception and Semantic Cephalic States. Shulkin and Heelan will describe the pragmatic/interpretative events that underlie action.

## References

- Adolphs, R. (1999). Social cognition and the human brain. *Trends in Cognitive Sciences*, 3, 469–79.
- AQ20 Alloway, T. P., Corley, M., and Ramscar, M. (2006). Seeing ahead: experience and language in spatial perspective. *Memory and Cognition*, 34, 38–386.
- Bargh, J. A., and Chartrand, T. L. (1999). The unbearable automaticity of being. *American Psychologist*, 54, 462–78.
- Baron-Cohen, S. (1995/2000). *Mindblindness*. Cambridge, MA: MIT Press.
- Barrett, L., Henzi, P., and Dunbar, R. (2003). Primate cognition: From ‘what now?’ to ‘what if?’ *Trends Cogn Sci*, 7, 494–7.
- Barsalou, L.W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22, 577–660.
- Barsalou, L. W., Barbey, A. K., Simmons, W. K., and Satos, A. (2005). Embodiment in religious knowledge. *Journal of Cognition and Culture*, 5, 14–57.
- Barton, R. A. (2004). Binocularity and brain evolution in primates. *PNAS*, 101, 10,113–15.
- Barton, R. A. (2006). Primate brain evolution: integrating comparative neurophysiological and ethological data. *Evolutionary Anthropology*, 15, 224–36.
- Berthoz, A. (2000). *The brain’s sense of movement*. Cambridge, MA: Harvard University Press.
- Boroditsky, L. (2000). Metaphoric structuring: understanding time through spatial metaphors. *Cognition*, 75, 1–28.
- Boroditsky, L., and Ramscar, M. (2002). The roles of body and mind in abstract thought. *Psychological Science*, 13, 185–9.
- Bourdieu, P. (1980/1990). *The logic of practice*. Stanford: Stanford University Press.
- Byrne, R. W. and Corp, N. (2004). Neocortex size predicts deception rate in primates. *Proceedings of the Royal Society*, 271, 1,693–9.
- Carey, S. (2009). *The origin of concepts*. Oxford: Oxford University Press.
- Casebeer, W. D. (2003). *Natural ethical facts*. Cambridge, MA: MIT Press.
- Clark, A. (1997). *Being there*. Cambridge: MIT Press.

20 *Introduction*

- Clark, A. (1999). An embodied cognitive science? *Trends in Cognitive Sciences*, 3, 345–51.
- Clark, A. (2003). *Natural-born cyborgs*. Oxford: Oxford University Press.
- Clark, A. (2008). *Supersizing the mind*. Oxford: Oxford University Press.
- Clark, A., and Chalmers, D. J. (1998) The Extended Mind. *Analysis*, 58, 7–19.
- Damasio, A. R. (1994). *Descartes' error*. New York: Putnam Press.
- Dear, P. (2006). *The intelligibility of nature*. Chicago: University of Chicago Press.
- Decety, J., and Jackson, P. W. (2006). A social neuroscience perspective on empathy. *Current Directions in Psychological Science*, 15, 54–8.
- Dewey, J. (1896). The reflex arc concept in psychology. *Psychological Review*, 3, 357–70.
- Dewey, J. (1916) *Democracy and Education*. New York: Macmillan.
- Dewey, J. (1925/1989). *Experience and nature*. La Salle, Illinois: Open Court.
- Donald, M. (1991). *Origins of the modern mind*. Cambridge: Harvard University Press.
- Donald, M. (2004). The virtues of rigorous interdisciplinarity. In J. M. Luraciello, J. A. Hudson, R. Fivush, and P. J. Baver (Eds), *The development of the mediated mind*. Mahwah, NJ: Lawrence Erlbaum Associates, pp. 245–56.
- Dunbar, R. I. M. (1998). The social brain hypothesis. *Evolutionary Anthropology*, 6, 178–90.
- Dunbar, R. I. M., and Shultz, S. (2007). Understanding primate brain evolution. *Philosophical Transactions of the Royal Society*, 363, 649–58.
- Flanagan, O. (2007). *The really hard problems*. Cambridge: MIT Press.
- Foley, R. (1998). The context of human genetic evolution. *Genome Res*, 8, 339–47.
- Frith, C., and Wolpert, D. (2003). *The neuroscience of social interaction*. Oxford: Oxford University Press.
- Galison, P. (1988). History, philosophy and the central metaphor. *Science in Context*, 47, 197–212.
- Gallagher, S. (2005). *How the body shapes the mind*. Oxford: Oxford University Press.
- Gallagher, S., and Meltzoff, A. N. (1996). The earliest sense of self and others: Merleau-Ponty and recent developmental studies. *Philosophical Psychology*, 9(2), 211–33.
- Gallese, V. (2007). Before and below “theory of mind”: Embodied simulation and the neural correlates of social cognition. *Phil. Trans. R. Soc. B*, 362, 659–69.
- Gallese, V., Fadiga, L., Fogassi, L., and Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, 119, 593–609.
- Gallistel, C. R. (1980). *The organization of action: A new synthesis*. Hillsdale, NJ: Lawrence Erlbaum.
- Gallistel, C. R. (1992). *The organization of learning*. Cambridge: MIT Press.
- Geschwind, N. (1974). *Selected papers on language and the brain*. Boston: Reidel Publishing Company.
- Gibbs, R. W. (2006). *Embodiment and cognitive science*. Cambridge: Cambridge University Press.
- Gibson, J. J. (1966). *The senses considered as perceptual systems*. New York: Houghton-Mifflin.
- Gibson, K. R., and Ingold, T. (Eds) (1993). *Tools, language and cognition in human evolution*. Cambridge: Cambridge University Press.

AQ21

# PROOF

Introduction 21

- Gigerenzer, G. (2000). *Adaptive thinking*. Oxford: Oxford University Press.
- Gigerenzer, G., and Brighton, H. (2009). Homo heuristicus: Why biased minds make better inferences. *Topics in Cognitive Science*, 1, 107–43.
- Gigerenzer, G., and Selten, R. (2001). *Bounded rationality: The adaptive tool box*. Cambridge: MIT Press.
- Glenberg, A. M. (1997). What memory is for. *Behavioral and Brain Sciences*, 20, 1–55.
- Glenberg, A. M., and Kaschar, M. P. (2002) Grounding language in action. *Psychonomic Bulletin and Review*, 9, 558–65.
- Goldman, A., and de Vignemont, F. (2009). Is social cognition embodied? *Trends in Cognitive Science*, 13, 154–9.
- Grafton, C. T. (2009). Embodied Cognition and the Stimulation of action to understand others. *Ann. NY Acad. of Sciences*, 1156, 99–117.
- Greene, J. D., and Haidt, J. (2002). How (and where) does moral judgment work? *Trends in Cognitive Science*, 6, 517–23.
- Handy, T. C., Grafton, C. T., Shroff, N. M., Katay, S., and Gazzaniga, M. S. (2003). Graspable objects grab attention when the potential for action is recognized. *Nature Neuroscience*, 6, 421–7.
- Hansen, N. R. (1958/1972). *Patterns of discovery*. Cambridge: Cambridge University Press.
- Hauk, O., Johnsrude, I., and Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron*, 41, 301–7.
- Heelan, P. (1983/1988). *Space-Perception and the Philosophy of Science*. Berkeley: University of California Press.
- Heelan, P. A., and Schukin, J. (1998). Hermeneutical philosophy and pragmatism: A philosophy of the science. *Synthese*, 115, 269–302.
- Herrmann, E., Call, J., Hernandez-Lloreda, M. V., Hare, B., and Tomasello, M. (2007). Humans have evolved specialized skills of social cognition: The cultural intelligence hypothesis. *Science*, 317, 1,360–6.
- Israel, J. I. (2001). *Radical enlightenment*. Oxford: Oxford University Press.
- Jackson, J. H. (1884/1958). Evolution and dissolution of the nervous system. In J. Taylor (Ed.), *Collected Works of John Hughlings Jackson* (1958), Volume 11. London: Staples Press.
- Jackson, P. L., and Decety, J. (2004). Motor cognition current opinion. *Neurobiology*, 14, 259–63.
- Jacob, P., and Jeannerod, M. (2003). *Ways of seeing*. Oxford: Oxford University Press.
- Jacob, P., and Jeannerod, M. (2005). The motor theory of social cognition: a critique. *Trends in Cognitive Science*, 9, 21–5.
- James, W. (1890/1917). *Principles of psychology*. New York: Henry Holt.
- Jaspers, K. (1913/1997). *General Psychopathology*, Vols I and II. (J. Hoenig and M. W. Hamilton, Trans.) Baltimore: The Johns Hopkins University Press.
- Jeannerod, M. (1997). *The cognitive neuroscience of action*. Oxford, England: Blackwell Publishers.
- Jeannerod, M. (1999). To act or not to act: Perspectives on the representation of action. *Quarterly Journal of Experimental Psychology*, 52, 1–29.
- Johnson, M. (1993). *Moral imagination*. Chicago: University of Chicago Press.
- Johnson, M. (2007). *The meaning of the body*. Chicago: University of Chicago Press.

# PROOF

## 22 Introduction

- Johnson-Frey, S. H. (2004). The neural bases of complex tool use in humans. *Trends in Cognitive Science*, 8, 71–8.
- Johnson-Laird, P. N. (2001). Mental models and deduction. *Trends in Cognitive Science*, 5, 434–42.
- Kahneman, D., Slovic, P., and Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge, England: Cambridge University Press.
- Kitcher, P. (1993). *The advancement of science*. Oxford: Oxford University Press.
- Knoblich, G., and Sebanz, N. (2006). The social nature of perception and action. *Curr. Dir. In Psychol. Sci.*, 15, 99–104.
- Lakoff, G., and Johnson, M. (1999). *Philosophy in the Flesh: The embodied mind and its challenge to Western thought*. New York: Basic Books.
- Lamm, C., Batson, C. D., and Decety, J. (2007). The neural substrate of human empathy: Effects of perspective-taking and cognitive appraisal. *Journal of Cognitive Neuroscience*, 19, 42–58.
- Lashley, K. S. (1951). *The problem of serial order in behavior*. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior*. New York: Wiley.
- Levinson, S. (1996). Language and space. *Annual Review of Anthropology*, 25, 353–82.
- Levinson, S. (2006). Cognition at the heart of human interaction. *Discourse Studies*, 8, 85–93.
- Lieberman, P. (2000). *Human language and our reptilian brain: The subcortical bases of speech, syntax, and thought*. Cambridge: Harvard University Press.
- Lieberman, P. (2002). On the nature and evolution of the neural bases of human language. *Yearbook of Physical Anthropology*, 45, 36–62.
- Mahon, B. F. and Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *J. of Physiology*, 102, 59–70.
- Martin, A. (2007). The representation of object concepts in the brain. *Annual Review of Psychology*, 58, 25–45.
- Martin, A., Ungerleider, L. G., and Haxby, J. V. (2000). Category specificity and the brain. In M. S. Gazzaniga (Ed.), *The New Cognitive Neurosciences*. Cambridge: MIT Press.
- Matlock, T. (2004). Fictive motion as cognitive simulation. *Memory and Cognition*, 32, 1,389–1,400.
- Mayr, E. (1942/1982). *Systemics and the Origin of Species*. New York: Columbia University Press.
- Mead, G. H. (1934). *Mind, self, and society: From the standpoint of a social behaviorist*. Chicago: University of Chicago Press.
- Mellars, P. (2006). Why did modern human populations disperse from Africa ca. 60,000 years ago? *Proceedings of the National Academy of Sciences*, 103, 9,381–6.
- Meltzoff, A. N. (2007). ‘Like me’: A foundation for social cognition. *Developmental Science* 10, 126–34.
- Meltzoff, A. N., and Moore, M. K. (1977). Imitation of facial and manual gestures by human neonates. *Science*, 198, 75–8.
- Merleau-Ponty, M. (1962/1970). *Phenomenology of Perception*. New York: Routledge and Kegan Paul: The Humanities Press.
- Mithen, S. (1996). *The Prehistory of the Mind: The Cognitive Origins of Art and Science*. London: Thames and Hudson, Ltd. [see sp. in text]

AQ22

# PROOF

## Introduction 23

- Moll, J., and Schulkin, J. (2009). Social attachment and aversion: On the humble origins of human morality. *Neuroscience and Biobehavioral Reviews*, 33: 456–65.
- Moreno, J. M. (1995). *Deciding together*. Oxford: Oxford University Press.
- Murphy, G. L. (2002). *The Big Book of Concepts*. Cambridge: MIT Press.
- Noe, A. (2004). *Action in Perception*. Cambridge: MIT Press.
- Passingham, R. (2008). *What is Special about the Human Brain?* Oxford: Oxford University Press.
- AQ23 Peirce, C. S. (1878). Deduction, induction and hypothesis. *Popular Science and Monthly*, 13, 470–82.
- Peirce, C. S. (1898/1992). *Reasoning and the Logic of Things*. (K. L. Ketner, and H. Putnam, Eds). Cambridge: Harvard University Press.
- Perani, D., Cappa, S. F., Bettinardi, V., Bressi, S., Gorno-Tempini, Matarrese, M., and Fazio, F. (1995). Different neural systems for the recognition of animals and man-made tools. *Neuroreport*, 6, 1,636–41.
- Pinker, S. (1994). *The language instinct*. New York: William Morrow.
- Premack, D. (1990). The infant's theory of self-propelled objects. *Cognition*, 36, 1–16.
- Pulvermuller, F., Shtyrov, Y., and Ilmoniemi, R. (2005). Brain signatures of meaning in action word recognition. *J. of Cognitive Neuroscience*, 6, 884–92.
- Rizzolatti, G., and Luppino, G. (2001). The cortical motor system. *Neuron*, 31, 889–901.
- Rowlands, M. (2010). *The science of the mind*. Cambridge: MIT Press.
- Rozin, P. (1976). The evolution of intelligence and access to the cognitive unconscious. In J. Sprague, and A. N. Epstein (Eds), *Progress in psychobiology and physiological psychology*. New York: Academic Press.
- Rozin, P. (1998). Evolution and development of brains and cultures. In M. S. Gazzaniga, and J. S. Altman (Eds), *Brain and Mind: Evolutionary Perspectives*. Strasbourg, France: Human Frontiers Sciences Program.
- Sabini, J., and Schulkin, J. (1994). Biological realism and social constructivism. *Journal for the Theory of Social Behavior*, 24, 207–17.
- Schulkin, J. (2000). *Roots of social sensibility and neural function*. Cambridge: MIT Press.
- Schulkin, J. (2004). *Bodily sensibility: Intelligent action*. Oxford: Oxford University Press.
- Schulkin, J. (2007). *Effort: A neurobiological perspective on the will*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Schulkin, J. (2009). *Cognitive adaptation: A pragmatist perspective*. Cambridge: Cambridge University Press.
- Shapin, S. (1996). *The scientific revolution*. Chicago: University of Chicago Press.
- Simon, H. A. (1982). *Models of bounded rationality*. Cambridge, MA: MIT Press.
- Simon, H. A. (1962). The architecture of complexity. *Proc. Amer. Philos. Soc.*, 106, 470–3.
- Stereley, K. (2007). Social intelligence, human intelligence, and niche construction. *Philosophical Transactions of The Royal Society*, 362, 719–30.
- Stewart, J., Gepenne, O., and Di Polo, E. A. (2011) *Enaction: Towards a New Paradigm for Cognitive Science*. Cambridge: MIT Press.

# PROOF

## 24 *Introduction*

- Swanson, L. W. (2000). Cerebral hemisphere regulation of motivated behavior. *Brain Research*, 886, 113–64.
- Swanson, L. W. (2003). *Brain architecture*. Oxford: Oxford University Press.
- Tomasello, M., and Carpenter, M. (2007). Shared intentionality. *Developmental Science*, 10, 121–5.
- Tomasello, M., Carpenter, M., Call, J., Behne, T., and Moll, H. (2004). Understanding and sharing intentions: The origins of cultural recognition. *Behavioral and Brain Sciences*, 28, 675–735.
- AQ24 Ullman, M. T. (2004). Is Broca's area part of a basal ganglia thalamocortical circuit? *Cognition*, 92, 231–70.
- Varela, F. J., Thompson, E., and Rosch, E. (1991). *The Embodied Mind*. Cognitive Science and Human Experience series. Cambridge: MIT Press.
- von Holst, E. (1973). Relative coordination as a phenomenon and as a method of analysis of central nervous functions. In *The behavioral physiology of animals and man*. Selected papers of Eric von Holst. Coral Gables, Florida: University of Miami Press.
- Wheeler, M. (2005). *Reconstructing the Cognitive World*. Cambridge: MIT Press.
- Whitehead, A. N. (1927/1953). *Symbolism*. New York: MacMillan Company.
- Wilson, M., and Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychological Bulletin*, 131, 460–73.



## 1

## The Slow Process: A Hypothetical Cognitive Adaptation for Distributed Cognitive Networks

*Merlin Donald*

### Introduction

Human evolution is marked by the emergence of a special kind of social-cognitive process, unique to hominids: distributed cognition, performed in mind-sharing cultures. Human social groups are more cognitively complex than others, but, at the time of our emergence as a species, human social groups were not necessarily larger in population than their predecessors. The increased complexity of their cognitive system was inherent, not so much in group size, but in the nature of the cooperative, interactive social cognitive processes that apparently characterized species *Homo* from the outset.

Cooperative cognitive work emerged as groups of archaic hominids became more interactive and interdependent in their cognitive activities and operational rules. This included the transmission of tool-making and tool-using skills by imitation, cooperative hunting and migration, group fire-tending, and the construction of communal shelters. All these innovations required a degree of interpersonal coordination and communicative skill. The evolutionary trajectory of hominids was one of increasingly shared cognitive work – group decision-making, transmission of skill, sharing of knowledge, and division of cognitive labor. This trend led toward complex distributed social systems, which also served as the means to achieve distributed cognitive work. This led eventually to full language and symbolic thought.

In effect, the major evolutionary transformation of early humans was due to the emergence of distributed cognition. The prehistory of human cognition was thus a cultural scenario: major cognitive change

26 *Merlin Donald*

was most evident at the group level, in the improved coordination of effort, by which ancient humans were better able to achieve their cognitive goals.

Although the emergence of mind-sharing and distributed cognition may have been the most dramatic change marking human emergence, it remains true that these changes at the group level required corresponding evolutionary changes in the brain itself. Brain and culture were coevolving in a symbiosis, whereby natural selection was evidently favoring those with more socially adept brains. At the same time, human culture was being transformed into a flexible and powerful storehouse of the information needed to shape the social brain in epigenesis. The evolving hominid brain became increasingly dependent on cultural guidance for its full development, to the point where, in biologically modern humans, the brain cannot realize its design potential outside culture. In effect, we have brains designed to function optimally in a distributed network, with the network transmitting vital epigenetic information to the developing brain.

The latter was a critical evolutionary step. Human culture defines much about the human brain, especially the so-called higher order features of mind that are crucial to sharing mind. The human brain does not acquire language, symbolic skills, or any form of symbolic cognition without the pedagogical guidance of culture. Through its epigenetic impact, culture is a major determinant of how the brain self-organizes during development – both in its patterns of connectivity and in its large-scale functional architecture. And the brain has evolved the sensitivities needed to assimilate the vital information stored in cultural networks.

### **The impact of culture on physical brain development**

The most obvious example of culture's real physical impact on brain development is literacy skill. Literacy is a fairly recent historical change, with no precedent in archaic human cultures; the vast majority of the world's languages have never developed an indigenous writing system. Yet certain dominant modern cultures are not only literate, but also heavily dependent on mass literacy for much of their cognitive work. Mass literacy is spread only by imposing modifications on the developing nervous systems of large numbers of individuals. These modifications are imposed by "educational" systems: basically, systems of organized group pedagogy whose origins can be traced back to the beginnings of literate culture.

# PROOF

*The Slow Process* 27

The cognitive subroutines that enable a person to become literate consist of chains of deeply automatized responses to visual symbols. These are hierarchically organized in functional brain architectures that support specific subcomponents of reading and writing skills, which are typically learned by prolonged immersion in educational systems that are highly idiosyncratic and culture-specific. The algorithms of educational systems are generated and transmitted collectively, formed by the governing ideas of the cultural environment. Literacy training is not easy, takes a considerable amount of time, and is not even close to becoming a species-universal skill for biologically modern humans.

Automatization of complex, fast response systems is the key to acquiring literacy skill. Non-automatized responses, such as those of someone who is learning to read a new language, do not allow the reader or writer to concentrate on the meaning of what is written. Automatization of all the stages of literacy training – including word recognition, grammars, vocabulary expansion, and expressive skills – can be achieved only after very extensive practice, to the point of overlearning, in successive stages of competency. During the acquisition phase of such skills, continuous conscious monitoring and corrective feedback are necessary. Once the basic skill has been learned, so that the entire procedural system is automatic, conscious monitoring of the basic skill set is no longer needed, and the response of the system becomes mandatory: that is, the reader cannot avoid responding to a visually presented word as a word. At that point, words and sentences can no longer be treated by the brain merely as a series of lines and contrasts; the meaning literally “pops out” of the marks on the page. Yet no one claims that these popout experiences are innate; they are culturally arbitrary, and learned.

They are also one of the most important interfaces with the distributed systems of culture. And they are instantiated in physical changes to the brain, which have been imposed by means of extensive cultural programming.

The physical reality of the culturally imposed automatic brain systems underlying literacy skill can be seen clearly in certain cases of acquired dyslexia and dysgraphia. In such cases, injury to the brain of a literate person selectively destroys a particular cognitive subcomponent of the literacy system, without damaging other closely related brain systems, such as speech and symbolic thought. Literacy-related brain systems thus appear quasi-modular in their organization: they can suffer partial breakdown of certain components, while leaving others intact (Shallice, 1988). For example, one particular lesion might cause a patient to lose the ability to read, while retaining the ability to write, and another

# PROOF

28 *Merlin Donald*

patient might suffer the reverse. Specific lesions might even eliminate the ability to read irregularly spelled words, while the patient remains able to read words with regular spelling. These cases point to the existence of specificity of function in acquired brain architectures. This is incontrovertible evidence in support of the direct impact of culture on adult brain functional organization.

There are many other similar examples of cognitive skills that require extensive training, originate in culture, and depend upon acquired functional brain architectures: for instance, mathematical, musical, artistic, and athletic skills. But literacy skill stands as the clearest evidence that culture can impose radical cognitive reorganization on the brain.

## Cultures as clusters of distributed cognitive systems

What adaptive forces drove human cultures to invest so heavily in literacy education, and, consequently, in the epigenetic reprogramming of millions of brains at great cost?

Human cultures are unique in their cognitive nature: ideas and memories can be traded and shared among the members of a group. A useful perspective on this aspect of culture may be taken from computational modeling: culture can be compared in principle to distributed computational networks, in which many computers are interconnected in a network, which acquires properties lacking in the individual computers that constitute it. Membership in the network can make each individual computer look "smarter" than it appeared before joining the network. Specialization and division of labor can be coordinated in a network, and the cognitive power of the coordinated group system can far exceed the reach of any individual. One could perhaps point to the Manhattan Project as the supreme example of what technologically enhanced cooperative cognitive work (including decision-making) can achieve, when performed in distributed systems made up of specialized, symbolically coordinated components.

Distributed cognition is a useful paradigm in which to view the developing brain. From birth, the rapidly growing human brain is immersed in a massive distributed cognitive network: culture. The network "interface" of the brain to culture is a social one. It usually consists of unwitting "carriers" of the culture – parents, relatives, peers – who convey crucial information about where to direct attention, what to notice, and what to remember. The human infant's brain seeks such input from the start. One might say that it has evolved a specialized adaptation to

# PROOF



*The Slow Process* 29

search an early connection with cultural–cognitive networks; any serious failure in establishing this social–cognitive connection can result in delayed development, and in some cases, such as autism, in a permanent developmental disability. This early cultural bond is crucial; the human brain has evolved a dependency on culturally stored information for the realization of its design potential.

This dependency applies to the specific content of the knowledge stored in culture, but it applies especially to the process of gaining access to culture in the first place. The first priority of a developing human brain must be to acquire from culture the basic social and attentional tools that it needs to elaborate its cultural connection (Nelson, 1996; Tomasello, 1999). Having done this in early infancy, it will then be in a position to “download” a massive amount of specific cultural content, some of which is procedural, in the form of skill, including language skill, and some of which is semantic. Without completing that early phase of connection and sharing of mind, much of the information in culture will remain undetectable throughout life. Social–cognitive skills are enabling and empowering, in a capacity sense: they make possible and expand access to information stored in subtle and normally invisible cultural loci.



One of culture’s most important by-products, technology, has further extended these prototypically human symbolic capacities, by restructuring the distributed cognitive networks of culture and opening up new possibilities for both representing knowledge and remembering it. A typical modern cognitive–cultural distributed network links together many human brains with communications technology, images, books, papers, and computers. These kinds of distributed networks perform much of the cognitive work of modern society, from landing aircraft to predicting the weather and planning educational curricula (Hutchins, 1995). Individuals must be attuned to these networks to function effectively in our society. Decision-making occurs within tight network boundaries.



This raises a major scientific question: what are the specific domains in which the human brain attunes itself to culture? The major interface of the human brain and its cultures is undoubtedly a cognitive one: the uniquely cognitive nature of human cultures can only be explained in terms of a brain–culture symbiosis in the domain of cognition. Cognition can be appropriately singled out as the primary domain in which culture and brain interact. Human cognition constitutes a complex core of subcapacities and operations, interconnected by means of an equally complex array of algorithms, shaped by cultural forces



# PROOF

30 *Merlin Donald*

during development. This applies both to the individual brain and to the wider distributed systems of culture. The individual is transformed by immersion in a distributed system. In such systems, memory in particular is distributed in many locations, and access paths proliferate.

One property of distributed systems is the division of labor across individuals. In a distributed system, the individual brain no longer has to contain within itself all the skills and information needed for individual survival. Perceiving, remembering, recalling, searching, and attending are managed to a degree from outside, by means of various symbolic media. So are the specific learned algorithms of thought. As the division of cognitive labor in culture becomes more and more specialized, the adaptive task facing a young mind changes, and this has consequences for the deployment of the brain's resources. In particular, memory storage and retrieval are divided between brains and other media in the complex distributed systems of modern culture, as are many of the algorithms that drive thinking and problem-solving. Since this modifies the habitual use patterns involved in cognition, and brain activity and growth directly reflect its habitual use patterns, it is reasonable to postulate that concomitant brain processes, such as synaptic growth and regional localization, are also immediately affected. Unfortunately, although brain plasticity has been well documented in humans, there is not much direct empirical evidence from brain imaging studies on precisely how habitual, culturally imposed use patterns affect growth and development throughout the life span. We have only begun to collect empirical data on the neuropsychological impact of our close interaction with the external symbolic environment. By collecting more evidence, perhaps we will come to know more exactly how deep immersion in the distributed cognitive networks of culture affects the development of the nervous system.

One way to further refine the questions that need to be answered in this area is to observe brain–culture interaction over long periods. The pattern of emergence of cognitive change and cultural differentiation in human ancestors might prove helpful in conceptualizing how internal cognitive activity, in the brain, is interwoven with cognitive–cultural activity, in distributed networks. In turn, this might enable us to ask more telling questions of the brain.

## A model of human cognitive and cultural coevolution

The unique innovation of human beings in prehistory was the evolution of distributed cognition to a new level, indeed, to several new levels that had no precedent in other species. The human brain is adapted

# PROOF



The Slow Process 31

to the existence of cognizing, mind-sharing cultures that far exceed the individual's ability to store and transmit accumulated knowledge and skill. However, mind-sharing cultures could not have emerged by themselves, *de novo*. They are the product of a spiraling interaction between brain evolution and cultural change. The following is a brief review of a specific model of brain–mind–culture coevolution in hominids (Donald, 1991; 1993; 1995; 2001).

The methodology used to derive this model was inherently interdisciplinary, drawing from many fields that could provide relevant evidence. My basic technique was to test every hypothesis, whether it grew out of one single field of research or several, against evidence from all other relevant fields of research, and to reject any hypothesis that was incompatible with any solid fact, whatever its origin. This tends to produce robust theories, since accidental convergences from disparate fields of inquiry are highly unlikely to occur, and multiple convergences are even less likely.

There was one additional core postulate driving this model: brain–culture coevolution, with cultural–cognitive evolution leading eventually to such innovations as language. If brain and culture coevolved, the result should have produced a universal architecture of cognition – both on the individual and on the distributed levels – that is evident in all human cultures. Such a structure should endure, even in the modern context, because evolution is conservative, and systems that are working well do not tend to be replaced. The larger architecture of distributed cognitive cultural systems should be a relatively stable and universal structure. A large-scale cognitive–cultural hierarchy of mechanisms should form the basis for cognitive activity within the networks that support mind-sharing cultures.

A wide review of the evidence suggests that there are three hypothesized “stages” of cultural–cognitive change in hominid evolution, during which the nature of hominid culture gradually shifted from the marginally symbolic, to the protosymbolic, to the fully symbolic. This process was not conceived solely as a linear, gradualistic series of changes, but rather was characterized by several “punctuations” in an otherwise stable hominid survival strategy. There was an archaic preadaptation about 2 million years ago when *Homo* first emerged, followed by a much more recent cognitive shift, within the past 400,000 years, that was radical and relatively rapid, and culminated in the fully symbolic cultures of biologically modern humans.

The physical evidence favoring this two-stage model came initially from two principal sources: fossils and material culture. An analysis of the fossil remains of human ancestors reveals two periods where there



# PROOF

32 *Merlin Donald*

was a relatively rapid increase in hominid brain size, and a change in body shape toward the modern pattern: the period from approximately 2 Mya to 1.5 Mya, when the species *Homo* first appeared, and a second period from ~500 Kya to ~150 Kya, when the species *Homo sapiens* first appeared.

Without necessarily conceding that increased brain size or body shape tells us anything in detail about the hominid mind, they do allow some rough time markers, and a partial reconstruction of their way of life. Such reconstructions suggest that these were periods of significant cognitive challenge, with a concomitant change in the survival strategies of hominids. The material cultural record left behind by hominids agrees with this picture. There were major changes in the cultural record during, and following, these two periods. The changes included changes in such things as tool-making, fire-making and fire-tending, diet, hunting skill, migration patterns, and the location and construction of home bases and shelters. Cultural and anatomical changes have not always coincided, and there is much debate about such details as the number of hominid subspecies, but the standard story of hominid emergence has not changed fundamentally during the last two decades.

There are compelling neural and cognitive considerations that greatly enrich this picture. Comparative anatomical evidence is an important clue here. Hominid evolution follows a trajectory from Miocene apes to modern humans. The starting and end points of brain anatomy are well known. Major differences between ape and human anatomy have been subjected to more detailed study, using advanced techniques, during the past decade, and the picture that emerges does not permit as much theoretical leeway as some might assume this field allows.

The cognitive networks that permeate all human cultures evolved in three stages, each of which added a new kind of representational “layer” to human culture, and each of which had its own evolutionary rationale. These networks dominate the brain and mind in epigenesis, and impose a hierarchical structure on higher, or symbolic, cognition. Such networks might be labeled, for convenience, as “cognitive–cultural networks,” or CCNs. They have a significant influence on the developing brain of the child, through the mediation of parents and community. CCNs coevolved with changes in various brain structures, and cannot exist without the cerebral apparatus that allows the young brain to assimilate these representational systems. On the other hand, it appears that very little detail is specified in the genes at this level. Increasingly, as a result of human evolution, it is the interaction between a highly plastic genetic potential and cultural reality on the ground in any given

# PROOF

generation that generates the actual cognitive organization of the individual brain.

Table 1.1 illustrates the key points of this evolutionary theory of human cognitive origins. It begins in Miocene primates with cognitive capabilities that are assumed to be roughly similar to those of modern apes. These capabilities are labeled as “episodic.” Four successive stages of hominid cognitive evolution are proposed in this scenario, labeled episodic, mimetic, mythic, and theoretic. Note that hominid cognitive evolution has here been captured in three cultural stages, because the most radical innovation in the hominid line is distributed cognition, culminating in a system of language and symbolic communication that has cultural origins. The scenario is thus: first generate cognizing cultures of a protosymbolic nature, then let these become more complex, until they spontaneously “combust” into systems of symbolic convention, and, eventually, into full-fledged language.

This proposal will seem unfamiliar to many cognitive neuroscientists, but the “stages” of human cognitive–cultural evolution should not

Table 1.1 Three stages in the emergence of human CCNs, starting with the “episodic” cognitive–cultural style of primates

Stage	Species/period	Novel forms of representation	Manifest change	Cognitive governance
Episodic	Primate	Complex episodic event-perceptions	Improved self-awareness and event-sensitivity	Episodic and reactive; limited voluntary expressive morphology
Mimetic (first transition)	Early hominids, peaking in <i>Homo erectus</i> ; 4–0.4 Mya	Nonverbal action-modeling	Revolution in skill, gesture (including vocal), nonverbal communication, shared attention	Mimetic; increased variability or custom, cultural “archetypes”
Mythic (second transition)	Sapient humans, peaking in <i>H. sapiens</i> ; 0.5 Mya – present	Linguistic modeling	High-speed phonology, oral language, oral social record	Lexical invention, narrative thought, mythic framework of governance
Theoretic (third transition)	Recent sapient cultures	Extensive external symbolization, both verbal and nonverbal	Formalisms, large-scale theoretic artifacts and massive external memory storage	Institutionalized paradigmatic thought and invention

34 *Merlin Donald*

seem too unfamiliar, because they were established on rigorous cognitive criteria: each putative stage involved a novel form of memory representation, and a new style of cognitive governance at the top of the distributed cognitive system that was, quite literally, governing. Each new stage – mimetic, mythic, and theoretic – marked the genesis of a new medium, or domain, of memory representation in the distributed system, or CCN, and in the individual brain. The latter effect was an epigenetic change due to “deep enculturation.” Each CCN domain postulated in this model has a complex internal hierarchical structure that is dictated by the properties of the shared memory systems available to hominids at that stage. The superordinate descriptive labels – episodic, mimetic, mythic, and theoretic – capture the top, or governing, level of representation within each domain.

### One additional point

This is a “cascade” model inasmuch as it assumes a conservative process that retains previous gains. As hominids moved through this sequence of cognitive adaptations, they retained each previous adaptation, and it continued to perform its original cognitive work perfectly well. New levels of representation evolved to perform a different kind of cognitive work for the species. Mimetic cognition incorporated, and extended, prior gains at the episodic level; and mythic, or narrative-based, cognition was scaffolded on top of a mimetic, or gestural, mode of thought and communication. The final step, the so-called theoretic stage, evolved slowly out of the classic mythic–mimetic thought strategies of traditional human cultural networks, retaining the latter within it. It was a combined product of extensive experience with sophisticated literacy skill and symbolic technology, resulting in the institutionalized application of analytic thought strategies to government, science, and education.

The first two hominid transitions – from episodic to mimetic, and from mimetic to mythic – were mediated largely by neurobiological change, while the third transition, to the theoretic mode, was heavily dependent on changes in external, nonbiological, or artificial memory technology. The fully modern mind retains all of these cognitive structures, both in the individual and in the distributed networks that govern cognitive activity in modern humans.

Each of these stages was marked by complex modifications in hominid survival strategies that undoubtedly involved many different changes in skeletal anatomy, brain anatomy, emotional responsiveness, intelligence,

# PROOF



*The Slow Process* 35

memory, social organization, reproductive strategies, and temperament, among many other factors. Cognitive evolution could not have taken place in a vacuum, and major changes in cognition undoubtedly had implications for many survival-related variables, including diet, intraspecific and interspecific aggression, heat dissipation, metabolic energy, disease resistance, physical size, sexual dimorphism, and so on. The cognitive stages listed above were derived in that very wide theoretical context. But the prime driving force behind these changes was a cognitive one.

The reasons for labeling the primate cultures of the Miocene epoch as "episodic" have been spelled out in various previous publications (Donald, 1991; 1993; 2001). The theory begins with the assumption that the early hominid brain, like its primate, and most probably australopithecine, predecessors, lacked language or any capacity for generating explicit symbolic representation in the wild. The archaic hominid brain, like most others in the primate line, shared the same basic design features that humans share with all primate brains. This means that the earliest predecessors of hominids would have been very clever social animals, with a remarkable ability to understand complex social relationships, but limited expressive skill. In other words, they could understand social episodes and scenarios, but had no way of expressing this knowledge to one another.

The cognitive capacity that supports episodic intelligence is best described as "event-representation." Events are the "atoms" of episodic cognition (Nelson, 1986). Social life consists of events, clustered in episodes; these define alliances, troupe membership, and power relationships. By this definition, primates have excellent event-representations, or ERs. They can remember specific events in an episodic manner; that is, they remember vivid details that are specific to a particular episode. For instance, after a fight with a rival, they remember the principal agents, outcomes, and future social implications of the fight. That kind of vivid, detailed event-memory in humans is usually called episodic memory, and it is anchored in concrete events. For this reason, the cognitive and cultural style of primates might be labeled "episodic."

The episodic mind-set of primates is nonsymbolic or presymbolic in its expressive or representational style. There is no evidence that primates think or communicate in symbols in their natural state. The episodic mind is concrete, analogical, episode-bound, and anchored firmly in the perceived present. It acts largely within the span of working memory, using perceived similarities between situations (and distinctions between them) as a means of choosing appropriate behavior.



36 *Merlin Donald*

Hominids, who shared an ancestor with chimpanzees about 6 million years ago, evolved beyond this mind-set at some point in their emergence. If we assume a Miocene starting point for hominids that was very close to the cognitive capacities of modern apes (A), and use biologically modern humans as the end point (B), the theoretical exercise becomes one of identifying the most probable sequence of events – neural, cognitive, and cultural – leading from A to B. The three transitions outlined in Table 1.1 constitute a coherent theory of the nature and approximate time course of the path from A to B.

### **Implications for theories of temporal integration in the social brain**

The highly social and interactive nature of human cognition has not yet been fully investigated in brain research. There are some major innovations underway in the direction of studying the “social brain,” but, for the most part, these consist of demonstrations of emotional connections with society. The transactional and distributed nature of social cognition itself is more difficult to study. But it must be studied, because social cognition and its consequence, mind-sharing cultures, are key to understanding the unique nature of the human mind.

The emergence of a complex interactive social–cognitive system represented a significant shift away from primate social life, and undoubtedly presented a major cognitive challenge for the evolving hominid brain. Perhaps the most significant challenge was in mastering the temporal dimension of social perception. Human social life unfolds in long, complex, multimodal, and interactive episodes and scenarios which establish such things as social structure, hierarchy, custom, group intentionality, and, ultimately, sophisticated interactive behaviors such as gesture, pedagogy, skilled rehearsal, and social cooperation. The mastery of such lengthy episodic experiences required the evolutionary improvement of a pre-existing primate capacity for temporal integration. Although we cannot say with certainty when it began, it seems certain that basic evolutionary improvements in human temporal integration were in place by 2–2.5 million years ago, when the distinctive distributed cognitive strategy of human society had begun to bear fruit in the form of cooperative hunting with stone tools.

The magnitude of this change can be seen in the limitations of even the most accomplished and socially intelligent of enculturated primates, when it comes to coping with the social complexity demanded by successful group cooperative work. Kanzi and Panbanisha, Savage-Rumbaugh’s

# PROOF



*The Slow Process* 37

star bonobos (Savage-Rumbaugh et al., 1993), can master many practical uses of tools and symbols, including even simple video games, and can understand speech sounds and elementary grammars much better than was thought possible a mere decade ago. But they are extraordinarily limited in their ability to follow, let alone master, such vital human social skills as multiagent communication, multiagent games, complex tools, and the complex nonverbal conventions that we know as social gestures.

Why is this? Some have suggested that these are “wild animals” whose instinctual responses cannot be suppressed. Yet they do successfully suppress many of their natural response tendencies, compared with their wild-reared conspecifics. Others have argued that they lack a special “language acquisition device,” or that they lack certain metacognitive skills, such as a capacity for perspective-taking or “theory of mind,” a hypothetical (and undoubtedly complex) capacity for understanding the minds of others. However, we have no good model of either of these postulated mechanisms, or even any convincing evidence that they exist as coherent brain subsystems.

Underneath these kinds of surface capabilities, there is the more fundamental cognitive challenge that these enculturated apes apparently cannot match: the comprehension of extended human social scenarios that engage several agents in complex interaction. Apes are excellent perceivers of social events, up to a certain level of complexity. But human social interactions exceed that capacity, and hominid evolution has evidently extended primate event-perceptual capabilities. Clearly, human beings are very good at perceiving their own social complexities. The popularity of manufactured entertainments, such as plays, novels, and films, that endlessly rework various social scenarios testifies to our obsession with complex social plots and narratives.

What is the cognitive element, missing in primates, that has enabled human beings to master so complex a social life? One possibility is that apes lack a capacity for the wide temporal integration that is necessary to cope with the intricate plots and subplots of human life. The continuous integration of new events into old scenarios, so common in human social cognition, allows the mind to oversee short-term events and episodes from a deeper background vantage point, while bracketing the fast-moving events in the foreground, and placing them in an accurate context. The standard model of cognition is very much dominated by the study of the sensorimotor foreground: processes such as perception, short-term memory, working memory, and attention are normally studied as reactions to what is immediately present at the sensory surface.



38 *Merlin Donald*

Little is known of the mechanisms that enable the longer-term integration of such events into longer ongoing social scenarios.

In reality, human social life, and many other aspects of human mental life in a more general sense, are lived in a slower, wider time frame or “intermediate time zone,” within which many events and episodes are grasped and understood in terms of their implications for social relationships and future behavior. The capacity to achieve temporal integration on this scale seems to be absent in apes. Human brain researchers should therefore be looking for a “slow process” in the brain that is uniquely human, and can operate over long time frames, such as many hours, while maintaining a long-standing bias, noting the place of every intervening event that occurs in the sensorimotor domain, and retaining long-term control over thought and behavior.

### The slow process: a hypothetical neural entity

In effect, the hypothesized “slow process” is a vastly extended working memory system that serves as the overseer of human mental life, and is the deepest layer of the mind. This is the intermediate-term governor of human mental life, the deep background process that shapes our cognitive agendas over the longer run, while maintaining oversight over the foreground of mental activity that occurs closer to the sensory surface. While its application in cognition is wide, its prime function is to enable the mind to comprehend and navigate the multifaceted social–cognitive world that human beings inhabit.

The existence of this capability presents a challenge to neuroscience, because there is no known neural process that can remain active for such long periods, and tolerate so many interruptions at the sensorimotor interface, while continuing to update its temporary “worldview” with new information. The slow process can track intricate events over long periods, yet it can also guide moment-to-moment thought and behavior, providing the contextual framework for forecasting and planning social action. The slow process involves memory; in fact, it is a form of extended working memory.

Half a century ago, Hebb (1949; 1963) proposed that there were only two kinds of memory record in the nervous system. One of these, short-term memory (STM) traces (later relabeled as working memory, or WM), which I call ST-WM traces, are electrochemical in origin, and constitute the active focus of activity in the brain at any given moment. The other kind of trace, or long-term memory (LTM), consists of structural changes, mostly in the form of altered synaptic connections. The former,

# PROOF



*The Slow Process* 39

ST-WM, are active, bound neural traces, more or less coextensive with awareness, that correspond essentially to “the feeling of what happens” (Damasio, 1999); that is, the stream of consciousness. The LTM records are normally inactive, and amount to a dead storage mechanism, not unlike, in principle, the dead memory records of, say, books or DVDs, inasmuch as they do not play a role in mental life unless activated. LTM traces are only effective in behavior when they are primed, and are more influential when retrieved into awareness, and converted into a fully activated trace in ST-WM.

In its basic form, this model has remained largely unchallenged and unaltered, except for some hotly debated details, such as whether some bound traces are too short to be classified as ST-WM. Hebb’s criterion for the existence of ST-WM was the delayed response, which cannot be demonstrated in many species that are undoubtedly capable of binding simple stimuli. Thus, a very short bound trace is different: a briefer neural trace that lacks a ST-WM mechanism to give it life beyond the immediate presence of stimulation from the environment. Simple binding and ST-WM thus exist in two different temporal ranges, with the second capable of sustaining its activity for many seconds, autonomously of external stimulation.



Neither the paradigms of binding studies nor those of ST-WM studies hint at the existence of a class of active neural traces that can last for hours on end, governing decisions and maintaining the general direction of behavior and thought. Yet this class of trace must exist, given the overwhelming evidence of autonomous sustained imagination, thought, and planning in human social life. There must necessarily be a third kind of neural process that corresponds in its time parameters to a broader period of temporal integration. This kind of trace cannot be as ephemeral as instantaneous binding or ST-WM; nor can it be as static as a permanent structural synaptic change. I have called this kind of longer neural trace “intermediate-term governance,” or ITG.



In its field of influence and time parameters, ITG corresponds more or less to the “supervisory system” postulated by Shallice (1988) and others, and identified especially with the prefrontal cortex. We have no good neural model of the activation or localization of such a long-lasting process. Researchers have identified many more transmitters and modulators than were known in Hebb’s era, and there are many potential candidates for a slow neural process with some of the properties needed to explain the existence of ITG. However, none of them seems a feasible candidate for the maintenance of something as complex and subtle as a very slow-moving social scenario or mental plan, running



# PROOF

40 *Merlin Donald*

in the deep background, enduring for many hours, and influencing a whole succession of actions and changes of strategy.

Human social life involves intricate strategic planning, with many moves and countermoves, and online adjustments and updates. The same applies to most social games, such as chess or soccer, which are designed to engage the ITG systems of the brain for long periods, toward some competitive end. The game may be interrupted by any number of trivial or subsidiary events, but the basic operating context imposed by the game will continue to dominate until it is complete. It is that deep background, the governing context, that I am describing here.

Hebb's model, modified in many details, has been widely applied in theories of attention, perception, and memory. Theories of attention and perception are still concerned with the formation and filtering of the short-term active trace. Theories of neural binding are concerned with both attention and perception in the shorter term (Singer, 1994; Crick and Koch, 1995). Researchers are also actively examining how the ST-WM system, aided by attention, can capture several bound neural traces that occur slightly apart in time, and integrate them into a single bound event-representation (Miltner et al., 1999; Rodriguez et al., 1999). In this sense, theories of binding are basic to theories of ST-WM traces as well as short-term binding.

But, for the moment, neither binding nor ST-WM theories can deal with an active slow process such as ITG.

This theoretical challenge is summarized in a wide temporal framework, which illustrates three levels of temporal integration achieved by the human central nervous system. The first operates in the very short time range, from fractions of a second to a few seconds at the most; that is, traditionally defined, short-term and local binding. The second operates in the ST-WM range of a few seconds or tens of seconds, and can integrate bound material from any modality into an active trace that exists for a few minutes at most. Unlike very short-term binding, it seems to be very limited in capacity, and closely tied to selective attention. The neural mechanisms of these two kinds of temporal integration appear to be somewhat distinct from one another. Moreover, the two mechanisms seem to have evolved at different times, in different species. Elementary binding evolved earlier, and appears to occur in many species. On the available behavioral evidence on delayed response, ST-WM capacity occurs in much fewer species, and appeared later in the evolutionary record (Donald, 2001, pp. 184–95).

There are theoretical gaps between theories of mechanism for these three temporal ranges of neural integration. The current evidence on



short-term electrical traces does not provide any feasible mechanism for stable, long-lasting active traces in human beings, despite the fact that they obviously exist. Longer-lasting neural activity (ITG) sets the ground rules and semantic foci to direct and control complex interactive social scenarios, such as conversations that last for many hours, and organized games of various kinds. It predominates in human life. It can be lost selectively in cases of damage to the prefrontal cortex in particular (Stuss and Benson, 1986). Yet, we have no good candidate for a theoretical mechanism for the trace aspects of ITG.

In summary, there is good reason to predict the eventual discovery of a slow temporal domain of brain activity, longer in duration than any existing demonstration of bound neuroelectric activity, and stable across many interruptions at the level of short-term binding and ST-WM activity. Such a mechanism must be able to maintain a stable bias in a specified neural network, for a long enough time, in the waking brain to account for human social cognition. It must also be able to maintain its activity through many interruptions in the faster-moving working memory foreground of cognition, standing in the deep background, and overseeing the traditional architecture that forms the core of most information-processing models of the brain.

From a review of human cultural and cognitive evolution, there is good reason to expect that this mechanism, whatever it may prove to be, is at the very heart of human social and cultural life, and close to the core of human nature.

This chapter is reproduced with permission from *The Journal of Physiology Paris*: Merlin, D. (2007). The slow process: a hypothetical cognitive adaptation for distributed cognitive networks. *Journal of Physiology Paris*, 101, 214–22.



## References

- Crick, F., and Koch, C. (1995). Are we aware of neural activity in primary visual cortex? *Nature*, 375, 121–3.
- Damasio, A. R. (1999). *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. New York, NY: Harcourt Brace & Company.
- Donald, M. (1991). *Origins of the Modern Mind: Three Stages in the Evolution of Culture and Cognition*. Cambridge, MA: Harvard University Press.
- Donald, M.W. (1993). Précis of origins of the modern mind with multiple reviews and author's response. *Behavioral and Brain Sciences*, 16, 737–91.
- Donald, M.W. (1995). The neurobiology of human consciousness: an evolutionary approach. *Neuropsychologia*, 33, 1,087–102.
- Donald, M. (2001). *A Mind So Rare: The Evolution of Human Consciousness*. New York, NY: W. W. Norton.



# PROOF

42 Merlin Donald

- Hebb, D. O. (1949). *The Organization of Behavior: A Neuropsychological Theory*. New York, NY: John Wiley & Sons.
- Hebb, D. O. (1963). The semiautonomous process: its nature and nurture. *American Psychologist*, 18, 16–27.
- Hutchins, E. (1995). *Cognition in the Wild*. Cambridge, MA: MIT Press.
- Miltner, W. H. R., Braun, C., Arnold, M., Witte, H., and Taub, E. (1999). Coherence of gamma-band EEG activity as a basis for associative learning. *Nature*, 397, 434–6.
- Nelson, K. (1986). *Event Knowledge: Structure and Function in Development*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Nelson, K. (1996). *Language in Cognitive Development: Emergence of the Mediated Mind*. New York, NY: Cambridge University Press.
- Rodriguez, E., George, N., Lachaux, J., Martinerie, J., Renault, B., and Varela, F. (1999). Perception's shadow: long-distance synchronization of human brain activity. *Nature* 397, 430–3.
- Savage-Rumbaugh, E. S., Murphy, J., Sevcik, R. A., Brakke, K. E., Williams, S. L., and Rumbaugh, D. M. (1993). *Language Comprehension in Ape and Child*. Monographs of the Society for Research in Child Development, Chicago, p. 58.
- Shallice, T. (1988). *From Neuropsychology to Mental Structure*. New York, NY: Cambridge University Press.
- Singer, W. (1994). The organization of sensory motor representations in the neocortex: a hypothesis based on temporal coding. In Umiltà, C., and Moscovitch, M. (Eds), *Attention and Performance XV: Conscious and Nonconscious Information Processing*. Cambridge, MA: MIT Press, pp. 77–107.
- Stuss, D., and Benson, F. (1986). *The Frontal Lobes*. New York, NY: Raven Press.
- Tomasello, M. (1999). *The Cultural Origins of Human Cognition*. Cambridge, MA: Harvard University Press.



## 2

## Social Cognition and Cortical Function: An Evolutionary Perspective

*Susanne Shultz and Robin I.M. Dunbar*

### Introduction

The human brain sets us apart from the rest of our primate relatives. Humans have large brains compared with other primate species. Specific areas of brain, especially within the neocortex, have recently undergone rapid expansion, with most of this increase occurring within the last 500,000 years. Although understanding the causes and consequences of this brain expansion has long been a preoccupation, we still have limited evidence that points to the exact forces that drove the increase in the human brain. Across primates, however, there is an increasing body of evidence which links brain architecture to social cognition. In this chapter, we will review the patterns of gross brain evolution in vertebrates as a whole, and in primates in particular. We follow this discussion with a critical evaluation of what total brain size means in terms of cognitive function. Finally, we will focus on the substructures in the brain associated with social tasks, and whether there is evidence for exceptional increases in these areas in the human brain.



### Evolutionary changes in brain architecture

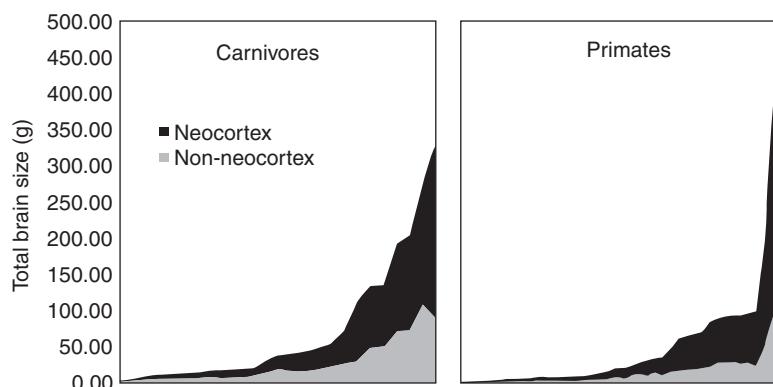
The vertebrate brain is segmented, with fore-, mid-, and hindbrain regions, and possesses an expanded and compartmentalized forebrain relative to other chordates (Shimeld and Holland, 2000). Reconstructing early vertebrate brains from fossil evidence is very difficult because the brain does not occupy the entire cranial cavity in fish, amphibians, and reptiles. However, fossil crania suggest that early vertebrate brains were small rudimentary structures similar to those found in the agnathans, or jawless vertebrates. Early mammals had brains that were



proportionately larger than contemporary reptile species, but, like modern insectivores, they had relatively small neocortices and large olfactory bulbs (Jerison, 1973).

Brain structure varies widely across extant vertebrates. The high level of variation in brain architecture between higher taxonomic levels is suggestive of mosaic (individual brain components evolving separately) rather than concerted (brain structures increasing in a scalable fashion) brain evolution (Brown, 2001; de Winter and Oxnard, 2001). Within taxa there tends to be more similarity and brain architecture and structures tend to be scalable with size (Clark et al., 2001). Even across groups there are some very predictable changes in architecture with increasing brain size. The forebrain, or telencephalon, is one of the areas that vary most between groups. Not only has the overall size of the telencephalon increased with brain size relative to the rest of the brain, so has the level of folding and gyration. The disproportionate increase in the telencephalon, and particularly the neocortex in mammals, means that the overall proportion of the brain that is composed of the neocortex varies dramatically across taxa (Figure 2.1). For example, in insectivores the neocortex is about 20 per cent of the overall brain size, whereas in hominoids (apes) the neocortex is nearly 80 per cent of the total brain (Clark et al., 2001). In contrast to the telencephalon, the proportionate size of the cerebellum is remarkably static across most mammals ( $\sim 0.13$ ).

It has long been assumed that brain size has increased over evolutionary time across all vertebrate taxa (Jerison, 1973). Although it is generally



*Figure 2.1* The relative size of the neocortex and non-neocortex across increasing brain size in carnivores and primates.



true that vertebrate brains have increased over time, we have recently shown that there is wide variation in the rate of brain increase and that this correlates with the relative brain size in living groups (Shultz and Dunbar, 2010). Many authors have speculated that reorganization, and not just expansion, is a critical factor in the functional increase in brain size across primate species. The ratio of white matter to gray matter has increased disproportionately (Rilling and Insel, 1998; Zhang and Sejnowski, 2000). Additionally, in mammals many structures, such as the neocortex and the dorsal lateral geniculate (LGN), have become highly laminated, which may allow efficient partitioning of information. Cortical minicolumns are vertically stacked neurons that represent a fundamental neural microcircuit (Chance et al., 2006); humans have a particularly expanded neuropil (interneuron) spacing between minicolumns, which suggests a reorganization of dendritic and synaptic connectivity. Such changes in the neural connectivity, parcellation, and substructuring potentially have profound consequences for cognitive performance. The specialization of function within structures, together with higher connectivity between brain regions, potentially leads to more efficient distributed cognition.

## How to compare brains

A seemingly obvious approach to comparing brains across taxa is to compare overall size, under the assumption that the larger the brain, the more powerful the “computer.” However, assuming a direct relationship between size and power is problematic for several reasons. Over computers as a whole, size does not directly relate to functional efficacy (Chitka and Niven, 2009). Within a comparable technology, however, it is more likely for a size/function relationship to hold (e.g., when comparing hard drive or RAM size, or dual versus single Pentium processor speed). Similarly, within closely related taxa in which brains are structured in similar ways it may be appropriate to compare brain volume directly. However, across taxa, larger brains are not necessarily more efficient than smaller brains. One reason why it is difficult to compare across divergent taxa is that brain size is strongly correlated with body size. Larger bodies may need more cognitive investment for controlling physiology and coordination rather than for higher cognitive function. To partial out the impact of body size on brain size, it has become a standard practice to control for body size when comparing brains across species. Relative encephalization (whether the observed brain size is larger or smaller than predicted) commonly is estimated



46 *Susanne Shultz and Robin I.M. Dunbar*

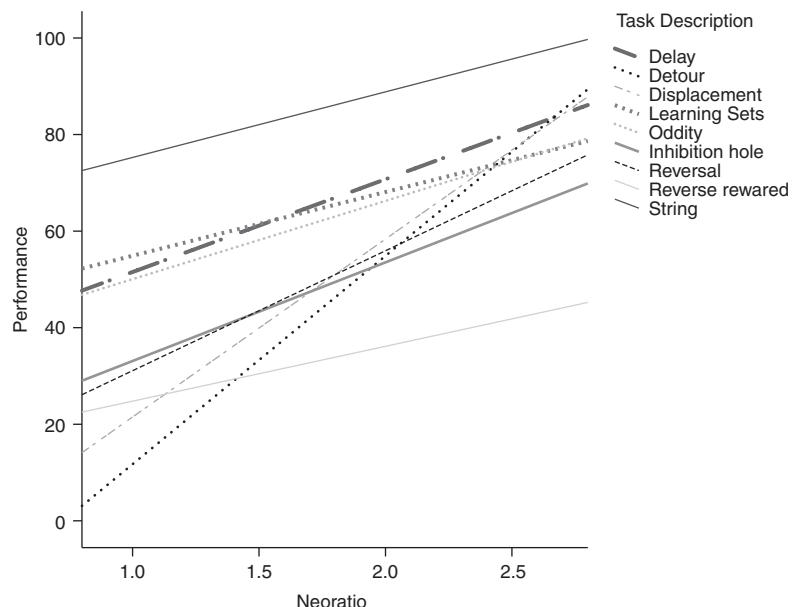
using deviations from the predicted allometric relationship between brain and body size. In sum, comparing the absolute size across brains is likely to be meaningful only when they have essentially the same structures and composition. Comparisons across divergent groups need to be carefully considered, and a “progressive” view of brain evolution, in which species are ranked, may be unhelpful (Deacon, 1990).

When we compare brains from different species, it is important to consider variation in the architecture as well as the size, especially when we wish to consider more fine-grained relationships between cognitive capacity and behavior. Not all parts of the brain are expected to be important in “higher level” cognitive processing. For example, the medulla/brain stem is associated with controlling metabolic and maintenance functions of the body and the cerebellum is implicated in physical coordination. In contrast, the parts of the brain most commonly implicated in learning, decision-making, discrimination, and memory are parts of the telencephalon, especially the hippocampus and neocortex. If specific areas of the brain are linked with higher cognition and behavioral complexity, then absolute or relative measures of these parts should be more enlightening than direct or relative measures of overall brain size. For example, neocortex volume or neocortex ratio (the ratio of the volume of the neocortex to the rest of the brain) can be used as a comparable measure of “executive brain” size (see Dunbar, 1998; Shultz and Dunbar, 2010).

### **Evolutionary explanations for brain size increase**

Brains are incredibly costly to develop and maintain. Adaptationists, who seek functional (fitness benefit) explanations for the existence of traits, argue that costly traits will not be maintained unless their adaptive benefit outweighs their costs. Given that brains are so costly, it seems intuitive that growing a large brain must offer some kind of cognitive advantage. Additionally, some vertebrate taxa have seen increases in brain size concurrent, or at least associated, with rapid speciation leading to adaptive radiations, such as within the passerines.

Despite the recognition that brain size is likely to have adaptive significance, there are surprisingly few attempts to relate measures of brain size with cognitive function or performance on cognitive tasks. One exception is the recent collation of various experimental cognitive paradigms administered across primate species during the twentieth century. Species performance on these tasks is positively associated with measures of brain size (Shultz and Dunbar, 2010, Figure 2.2). Thus,



**Figure 2.2** The relationship between neocortex ratio (neocortex/rest of brain) and cognitive performance on a standard battery of executive function tasks in primates (modified from Shultz and Dunbar, 2010).

there is some evidence for a cognitive benefit from having large brains. However, this does not provide an adaptive explanation for brain size increases. This section will present the major arguments for brain size evolution in vertebrates, starting with the nonadaptive “developmental” hypothesis, then the ecological and finally the social hypotheses for encephalization.

### Developmental

In the 1980s, comparative studies of brain size in birds and mammals identified numerous life history correlates with brain size. Species with larger brains tend to have delayed development, reach reproductive maturity later than small-brained species, have fewer offspring and typically have longer life spans (Clutton-Brock and Harvey, 1980). Many of these associations may be primarily driven by body size, as larger-bodied animals have larger brains and slower life histories. However, for a given body size, species that have large brains have higher relative

AQ1



basal metabolic rates (BMR) than species with smaller brains because brains are metabolically costly to maintain. It therefore appears that species with large brains need to be able to support them energetically. [Martin (1979)] extended this argument further by suggesting that available maternal energy and gestation period exert the strongest constraints on fetal, and hence adult, brain weight. However, life history variables may be either causes or consequences of large brains. If large brains are an adaptive response to environmental challenges, then it follows that the life history and physiology of large-brained species will need to be able to support the demands of growing and maintaining a large brain. Using a path analysis approach, we demonstrated that correlates between metabolism and life history are more likely to act as constraints on brain size rather than driving directional selection themselves (Dunbar and Shultz, 2007b).

### Ecological

Behavioral flexibility should allow individuals to respond to unpredictable environments and to use novel resources. There have been a number of efforts to quantify behavioral flexibility. The behavioral drive hypothesis (Wyles et al., 1983) suggests that novel behaviors spreading through a population can change (or potentially release) selection pressures. For primates, ecological hypotheses have centered on the need to track resources that vary over time and space. Frugivorous species have larger relative brain size than folivorous species (Allman et al., 1993; Harvey et al., 1980); larger-brained species also tend to have larger home ranges than smaller-brained species. This has been interpreted as an indication that the variable nature of fruit resources (over both space and time) has imposed a cognitive demand on individuals which need to track resource availability. However, there are several inconsistencies with this logic. First, there are a number of relatively small-brained frugivores that are sympatric with primates, and thus experience the same ecological pressures (e.g., forest duikers, flying lemurs, squirrels, viverrids). Thus, it does not seem essential for frugivorous species to be large-brained. Second, we partialled out the relative contribution of different traits to brain size and have shown that high-quality fruit diets are more likely necessary to support the high metabolic demands of large brains rather than being the driving selective pressure increasing brain size (Dunbar and Shultz, 2007a).



### Social hypotheses

The cognitive demands imposed by social complexity were first proposed as an explanation for why primates have unusually large brains for their body size (Humphrey, 1976; Jolly, 1966). This proposal has since become crystallized as the Social Brain Hypothesis (SBH) (Dunbar, 1992; 1998). The essence of the SBH is that the need to solve (ecological) problems in a social context, rather than in a demographic vacuum, imposes significant cognitive demands. Within a social environment, individual decisions must be responsive to the decisions made by other group members and the constraints these impose. For example, in order for a foraging group to maintain coherence and not fission, individuals must make choices that not only allow their energetic needs to be met, but also permit other group members to meet theirs. Byrne and Whiten (1988) developed the "Machiavellian Intelligence" hypothesis, which proposes that the cognitive demands of living in a social group may be more linked to cheating and detecting cheating than just monitoring relationships. Although Byrne and Whiten's hypothesis is not in opposition to Dunbar's social brain, it does put a slightly more malevolent twist on primate intelligence.

One widely cited prediction resulting from the social brain hypothesis is that social group size should correlate with brain size because the number of potential dyadic relationships (interpreted as one index of social complexity) is proportional to group size. Although strong support for this prediction has been found in primates (Barton, 1996; Dunbar, 1992; 1998), several recent analyses suggest that this prediction may be too simplistic: group size does not consistently correlate with brain size in some taxonomic groups (Beauchamp and Fernandez-Juricic, 2004; Shultz and Dunbar, 2007b), implying that the relationship between brain size and sociality may be more complex than previously supposed. Thus, the nature and stability of relationships may be more important than the number of aggregating individuals. Some animal groups, such as herding antelope, wildebeest migrations, or flocks of wading birds, are little more than aggregations and exhibit a fluid structure: individuals join and leave the group as their needs, or the environment, dictate. Other groups appear to be more structured, with membership relatively stable over time. It is this stable structure, in pair-bonds or foraging groups, which appears to be most related to large relative brain size in bats, birds, carnivores, and ungulates (Shultz and Dunbar, 2007a). Thus, primates are unusual in that group size itself is positively correlated with brain size.

AQ3





50 *Susanne Shultz and Robin I.M. Dunbar*

A “complex” social environment may be more the result of an individual’s role within the group and its relationships with other group members rather than the total number of individuals with whom it associates. If this is true, then the nature of interactions between individuals should be more important than the number of aggregated individuals. Indeed, alternative indices of social complexity such as deception rates (Byrne and Whiten, 1988), mating strategies (Pawlowski et al., 1998), grooming clique size (Kudo and Dunbar, 2001), and coalition rates (Dunbar and Shultz, 2007) all correlate with relative brain size in primates.

The social brain hypothesis can also be tested by comparing indices of sociality with evolutionary patterns of brain size increase. We have recently found that evolutionary changes in brain size are associated with sociality in extant species (Figure 2.3). Thus, not only are large brains associated with sociality in living species, but there is also a signature that brains have increased most in highly social taxa. Although we do not have a firm understanding of why sociality is cognitively demanding, we do have these consistent associations between brain size and sociality that strongly suggest a causal relationship.

### What are the cognitive costs of sociality?

One unresolved issue is what exactly constitutes the cognitive demands of bonded sociality. In a broad sense, both the ecological and social intelligence theories focus on an individual’s ability to make flexible adaptive decisions in the face of unpredictability. One approach to reconciling the ecological and social hypotheses may be to look at the problems that animals face in their environment. Basic ecological problems involve resource availability and distribution. However, individuals also have to contend with inter- and intraspecific problems, which

interact with and compound the initial ecological problems. A solitary animal needs to maintain its energy balance by finding enough food of appropriate nutritional quality. Individuals must make a series of complex decisions, including how to locate appropriate food, choosing dietary items, how to optimally use resource patches, and how to allocate time to various activities. However, individuals do not occur in isolation; they have to avoid predators, locate mates, and minimize competition over resources, and thus need to make strategic decisions about how to balance predation risk, time reproduction, and allocate resources. In stable social groups, then, individuals must make decisions in the context of what other group members are doing, which necessarily will add complexity to decision-making.



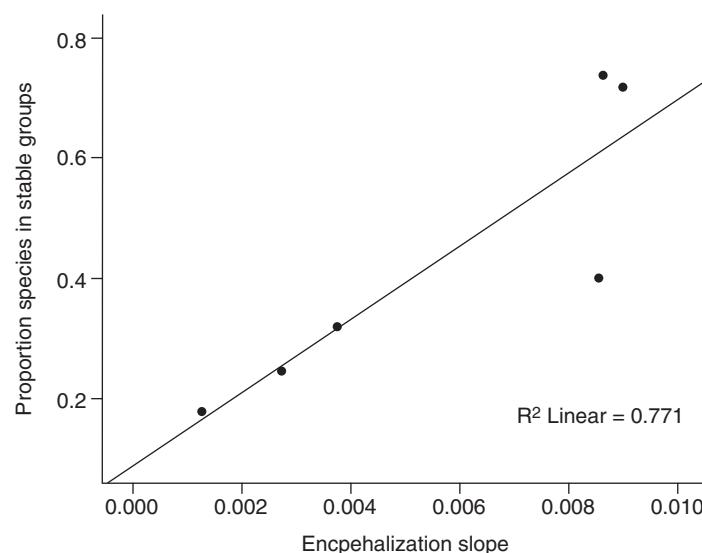
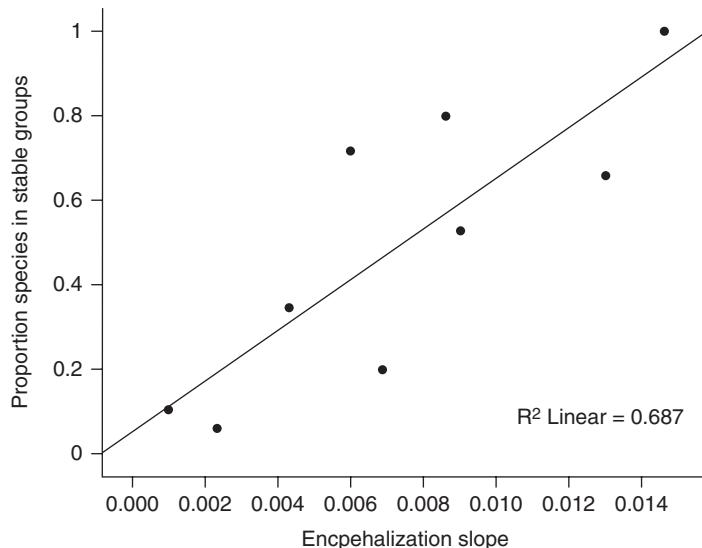


Figure 2.3 Brain size has increased more rapidly in highly social taxa than in less social ones.

52 Susanne Shultz and Robin I.M. Dunbar

This still does not tell us what types of social information individuals must process. The group size relationship in primates leads to a first principle assumption that individual recognition must be a key component of social intelligence. Primates are able to recognize a large number of conspecific faces, and it has even been suggested, somewhat controversially, that chimpanzees can recognize familial similarity in unknown individuals (Parr and de Waal, 1999). Although there is little doubt that individuals recognize other group members, it is less clear whether this itself is a cognitively difficult task. Sheep, and even fish, have been shown to be able to recognize a large number of conspecifics, and appear to have a remarkable ability to remember other individuals they have previously interacted with (Bshary et al., 2002; Kendrick et al., 2001). Thus, individual recognition *per se* may be a fairly rudimentary cognitive skill common to many social vertebrates. Recognizing intention and emotion in other individuals' faces is likely to be a much more complex cognitive task, and has been proposed as a major social function for cognition (Brothers, 1990).

Many primate species interact in cooperative contexts (Harcourt and De Waal, 1992). Game theory, and particularly prisoner's dilemma, has been invoked as a mechanism to stabilize cooperative interactions. However, to engage in strict reciprocal altruism, individuals must possess a memory of past interactions with each player. Additionally, it has been suggested that individuals must possess a concept of third-party relationships when deciding about intervening in conflict between two other players. Finally, interpreting intentional behavior of other individuals may be a fundamental cognitive advance; by predicting the behavior of others it may be possible to develop more effective counterstrategies. This has led to the proposal that theory of mind (or being able to represent others' mental states) was a key cognitive innovation in the hominoid lineage.

A related, but slightly different, take on the social intelligence hypotheses is the potential benefit of social learning. Individuals that can effectively exploit and utilize the knowledge of others can avoid a costly trial and error learning process, and adaptive behavior can spread quickly through populations. A special case of learning from conspecifics, imitation, allows fast and accurate transmission of information, but does not necessarily require higher social cognitive capacity such as theory of mind (Heyes, 1993). Populations of many animals show "cultural" behaviors, those which on the surface do not appear to be adaptive, but are consistently maintained in a population. A lot of interest has focused on cultural differences in chimpanzees; as culture is

AQ8



seen as a distinctive feature of human societies, the claim that culture is seen as cognitively advanced appears to be logical. Although it is not reasonable to state that there are no cultures in primates, interpreting population differences in a cultural context does not really highlight the cognitive demands or adaptive significance of these behaviors.

### Brain architecture and sociality

Although there remain crucial gaps between micro-level neurophysiology, neuro anatomy, and behavior [Dunbar, 2009; Dunbar and Shultz, 2007b; Healy and Rowe, 2007], there is a growing body of evidence which implicates specific regions with social cognitive tasks. The neocortex, rather than the rest of the brain, is most strongly associated with social characteristics in primates and other mammals. Changes in the neocortex underpinning the evolution of social cognition are further supported by the localization of structures within the neocortex, particularly regions of the frontal and temporal lobes, associated with social tasks. Within the frontal lobe there is extensive evidence for the anterior cingulate cortex (ACC), the ventro-medial prefrontal cortex (VMF) and orbitofrontal cortex (OFC) being involved in processing social tasks (Behrens et al., 2009; Brothers, 1990). In the temporal lobe the amygdala, superior temporal sulcus (STS), and temporal-parietal junction (TPJ) are most frequently associated with sociocognitive tasks. Finally, within the parietal lobe the precuneus has been proposed to have a crucial role in self-awareness and perspective-taking. Here we review the evidence for these regions being associated with social cognition and evaluate scaling patterns in primate brains to determine whether the regions implicated in social cognition have disproportionately increased with brain size.

AQ9

#### The frontal lobe

The frontal lobe is generally involved with executive functioning, including decision-making, episodic memory, emotional processing, and inhibition. As previously mentioned, within the frontal lobe there are a few key substructures with particular social relevance, most of which are located in the anterior, or prefrontal, cortex. The lateral and dorsal sections of the prefrontal cortex are primarily associated with secondary processing of sensory information and this region also has associations with the premotor cortex (Miller and Cohen, 2001). The orbital PFC is associated with tasks involving social, emotional behaviors, whereas the posterior frontal cortex is more implicated in executive



54 *Susanne Shultz and Robin I.M. Dunbar*

functioning tasks (Hecaen and Albert, 1978; O'Doherty et al., 2000; Rolls et al., 1994; Stuss and Benson, 1986). The ventro-medial prefrontal cortex (VMF) roughly corresponds to Brodmann's area 10. However, the exact boundaries of the region are not universally consistent: either the term is used to describe the area which incorporates the medial and ventral portion of the frontal lobe (including the orbital prefrontal cortex) or it specifically refers to the regions above the orbital prefrontal cortex.<sup>AQ10</sup> The VMF is believed to be involved with linking perceptual information with its emotional and social significance (Damasio et al., 1994). Damage to this region is associated with changes, sometimes profound, to personality, empathy, and socially appropriate behavior (Adolphs, 1999). The anterior cingulate cortex was traditionally considered part of the limbic system. However, the more recent view is that it is part of the ventral frontal cortex.<sup>AQ11</sup> The ACC lies below the prefrontal cortex, and is believed to regulate cognitive and emotional processing (Bush et al., 2000).

The insula, located deep within the fissure or junction between the temporal and frontal lobes, is tightly linked to the ACC via the frontoinsular cortex. The insula was first implicated in interoception and is therefore important for regulating motor and homeostatic control (Fink et al., 1997; Penfield and Faulk, 1955). However, there is increasing evidence that the insula, and in particular the anterior insular cortex, is involved with self-awareness and abstract corporeal representation (Craig, 2009). The anterior insular cortex (AIC) is activated in response to self-recognition (viewing own image: Devue et al., 2007). The AIC is activated during more general facial recognition, but this activity is strongest when viewing offspring or romantic attachments (Bartels and Zeki, 2004). Additionally, the insula is widely recognized as playing a role in music and speech processing (Ackermann and Riecker, 2004; Koelsch et al., 2006; Patel et al., 1997). Finally, insula activation correlates risk-related decision-making involving probabilistic rewards and uncertainty (Critchley et al., 2001; Elliott et al., 2000; Ernst et al., 2002; Grinband et al., 2006; Hsu et al., 2005; Huettel et al., 2005; 2006; Paulus et al., 2003). The anterior insula appears to be functionally involved in a wide range of emotional and perceptual processing, and hence may function as a cognitive centre for self-awareness and consciousness (Craig, 2009).

### **The temporal lobe**

A primary function of the temporal lobe is identification of objects (Goodale and Milner, 1992);<sup>AQ12</sup> however, several regions of the temporal



lobe are key to social cognition, particularly facial processing, language, and theory of mind. Superior temporal sulcus neurons respond to actions observed in other individuals: approach or retreat, head orientation and hand-object interaction as well as dynamic information from faces (Puce, 1988). Responding to or recognizing these behaviors potentially forms the basis for interpreting intentional acts of other individuals (Perrett et al., 1985; 1990). The superior temporal sulcus is activated in response to facial stimuli in both primates and humans (Baylis et al., 1987; Desimone et al., 1984). The fusiform gyrus in the temporal lobe appears to contribute disproportionately to the perception of faces, and in humans part of the fusiform gyrus, the fusiform face area, appears to be specifically devoted to facial recognition (Kanwisher et al., 1997). A number of recent studies have suggested that the temporal-parietal junction is particularly associated with representation of the mind of other individuals.

AQ13

AQ14

The amygdala has long been recognized to be important for processing emotional responses. The amygdala is involved in downstream processing of faces (activation occurs after the superior temporal sulcus) (Brothers, 1992) and lesions result in impairment in reading faces. Additionally, the amygdala is also responsive to movement, especially within a social context (Brothers et al., 1990). However, the amygdala cannot be viewed as a region specialized for processing socially relevant information as it is a structurally and functionally heterogeneous region (Swanson and Petrovich, 1998). For example, a large proportion of the amygdala is devoted to processing olfactory information and it also receives input from the temporal, prefrontal, insular, temporal, and olfactory cortical areas, from brainstem viscerosensory and nociceptive centers, and from parts of the caudal thalamus transmitting auditory and somatosensory information (Swanson and Petrovich, 1998). Activation between the hippocampus and amygdala during social interactions may point to a distributed network associated with episodic memory. More generally, it has been proposed that the amygdala has a role in processing emotionally arousing stimuli (Adolphs, 1999). Thus, although the amygdala appears to play a critical role in processing social information, it is involved in receiving a wide range of input, some, but not all, of which is germane to social interactions.



### The parietal lobe

Although a primary function of the parietal lobe is to integrate sensorimotor information (Goodale and Millner, 1992), within the parietal lobe there are regions that are specifically linked to social cognition.



# PROOF

56 *Susanne Shultz and Robin I.M. Dunbar*

The precuneus is implicated in a diversity of cognitive processing, such as visuospatial imagery, episodic memory retrieval, and first-person perspective-taking. The precuneus is more highly developed in humans and has complex columnar cortical organization and extensive links with the frontal lobes, including the ACC (Cavanna and Trimble, 2006). In particular, the right temporal-parietal junction (RTPJ) has been found to be activated in response to mentalizing tasks, and its function has been proposed to be highly specific to representation of mental states (Saxe and Kanwisher, 2003; Saxe and Powell, 2006; Saxe and Wexler, 2005). However, a parallel body of evidence implicates the RTPJ as key to tasks which require a redirection of attention (Mitchell, 2008).

In sum, a range of brain regions are believed to play key roles in processing social information. A number of these are specifically implicated in mentalizing and self-awareness tasks. The range of regions associated with mentalizing is supported by theory of mind studies. A recent review of 40 neuroimaging studies of Theory of Mind (ToM) (Carrington and Bailey, 2009) has identified that the brain regions most consistently activated by ToM tasks are the medial prefrontal and orbitofrontal cortex (93 per cent of studies), temporo-parietal junction (58 per cent of studies), anterior cingulate cortex (55 per cent) and superior temporal sulcus (50 per cent). The amygdala (13 per cent) and insula (13 per cent) were less commonly activated in these studies.

We have not discussed language as a key sociocognitive function, as there is limited evidence available about cross-species brain architecture which allows a comparison of the complexity of fully fledged human language with primate communication. However, language is clearly a key innovation in the human lineage, and a substantial body of work has identified key regions of the brain associated with language production and comprehension. The traditional view of language organization highlights the frontal lobe as the seat of language production and the (left) temporal lobe as the seat of language comprehension. Damage to the left frontal lobes can produce deficits in language comprehension and particularly production, including pathologies such as the classical Broca's aphasia (Broca, 1861; Alexander et al., 1989), whereas damage to the left temporal lobe causes deficits in comprehension, including Wernicke's aphasia (Wernicke, 1874). However, imaging studies have suggested that language is not as localized as the classical models implied, but, rather, involves a network of areas in the frontal, temporal, and parietal lobes (Binder et al., 1997).



### Neuron functionality

There is considerable evidence that single neurons play a pivotal role in individual, and especially face, recognition (Brothers, 1992). Single neurons that fire in response to complex information such as faces or landmarks are generally termed grandmother cells. More recent work has provided more evidence for the role of individual neurons in the medial temporal lobe in recognizing familiar faces (highly publicized as Jennifer Aniston neurons; Quiroga et al., 2005). However, this claim has been criticized on the grounds that the mere fact that some neurons associated with facial recognition appear to have high specificity does not mean that facial recognition is the responsibility of single neurons (sparseness). Rather, facial recognition neurons are likely to form a critical part of a wider distributed network. In a response to this issue, Quiroga et al. (2010) agree that types of representation vary across levels of sensory processing, and there may be a convergence from distributed to sparse representations. Higher-level processing will necessarily be more distributed than the initial stimulus response. In such a case there may be varying levels of specificity, such that some neurons fire to a general class of objects (e.g., landmarks or faces) with downstream neurons functioning on a more specific basis. Additionally, evidence from macaques suggests that there is high sensitivity not just to identity of other individuals but to their actions and orientation (Brothers, 1990).

A final class of neurons that deserves recognition are the mirror neurons, whose discovery has generated a great deal of excitement among those interested in social cognition. Mirror neurons are a group of neurons that are activated both when an individual is performing an action and when it observes another individual perform the same action (Gallese et al., 1996). Mirror neurons have been found in the premotor cortex (guiding movements with sensory feedback) and the supplementary motor cortex, as well as two regions in the parietal cortex (Gallese et al., 2002). The apparent dual role of these neurons has led to the inference that they play a role in “action understanding” (Gallese et al., 1996; Rizzolatti and Craighero, 2004; Rizzolatti et al., 2001). This suggestion has been widely taken up in the literature as a potential neural and cognitive foundation for understanding intentionality, being able to predict behavior of other individuals (Fogassi et al., 2005), and developing a theory of mind (Keysers and Gazzola, 2006). However, the overgeneralization of the role of mirror neurons has received considerable criticism (Hickok, 2008), as the ability to understand others’ actions does not appear to be limited to the mirror neuron system; nor

AQ17



does understanding appear to be impaired by damage to the mirror neurons. Mahon and Caramazza (2008) suggest that their role is primarily to link sensory and motor modalities without necessarily invoking a primary role in action understanding. Thus, the functional role of mirror neurons is not fully resolved and, although they are certainly activated in social contexts, they may not have a central role in higher cognitive function.

### Scaling in the social brain

To evaluate the relative changes in the regions associated with social cognition, we have used several complementary databases of primate histological (Stephan et al., 1981) and structural MRI (Schoenemann et al., 2005; Semendeferi et al., 2002) brain architecture. If the evolution of the primate brain has been driven by social cognition, then the regions of the brain associated with social tasks should scale superlinearly to nonsocial regions. Given the evidence previously presented, we predict that the frontal lobe, and especially the prefrontal area, should show the most pronounced increases. Secondary regions that process socially salient information, but are also implicated in a wider range of cognitive function, such as the amygdala, the temporal lobe, the hippocampus, the insula and, more generally, the parietal–occipital lobe should show less pronounced increases. Finally, regions of the brain that show less specific associations with social cognition (e.g., brain stem, cerebellum, visual regions including V1 and the LGN) should show the least pronounced increases. Unfortunately, the available databases do not subset the lobes well enough to allow a comparison of the specific components associated with social cognition, such as the TPJ, the precuneus, and the ACC.

As predicted, both the neocortex as a whole and the non-V1 cortex increase superlinearly as compared with other major brain regions using the Stephan et al. database of gross primate brain architecture (Figure 2.4a). However, even more intriguing are the results from the Semendeferi et al. database, which indicate that it is particularly the frontal lobe and insula which have undergone the most dramatic changes (Figure 2.4b). Finally, the Schoenemann data indicate that the prefrontal cortex in particular scales superlinearly (Figure 2.4c), although this database does not provide additional regions with which a direct comparison can be made. The nonexecutive parts of the brain in the Stephan and Semendeferi datasets (the core, the medulla, and the cerebellum) scale at the lower slopes. The components involved with,

AQ18



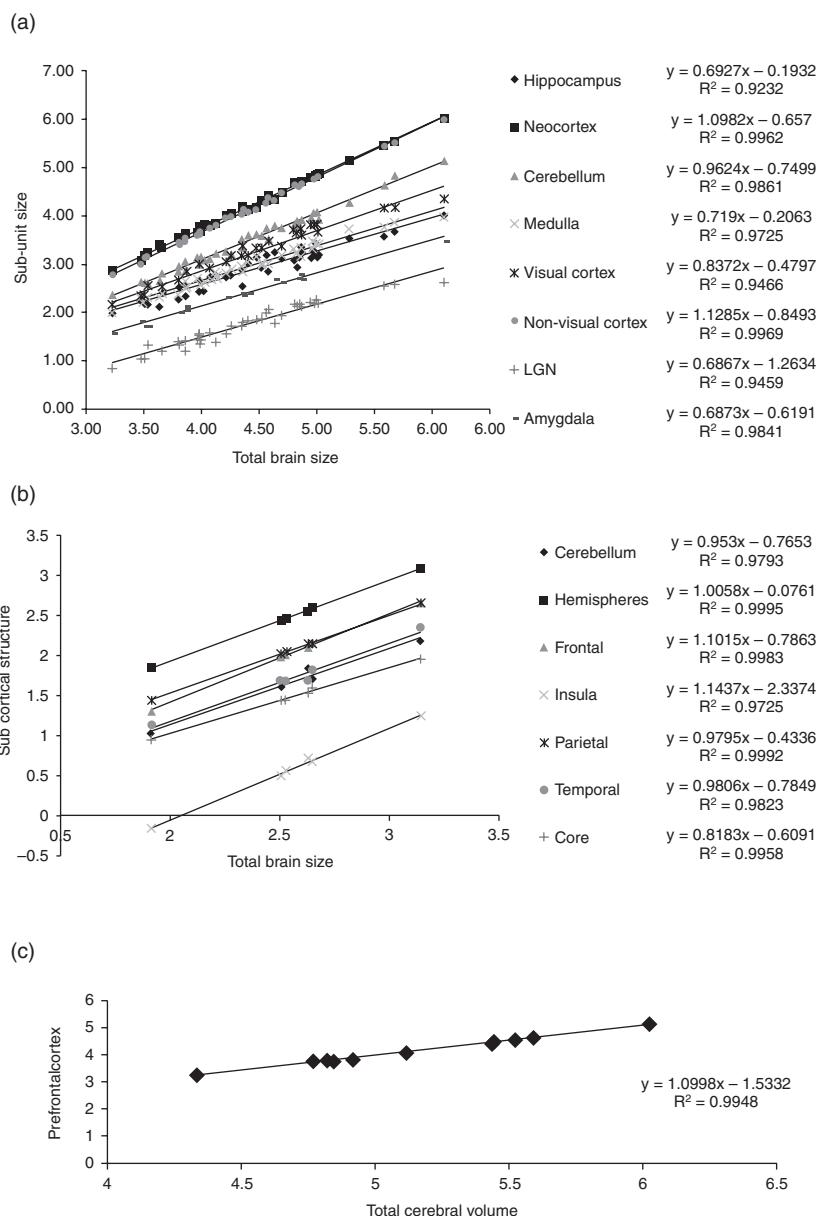
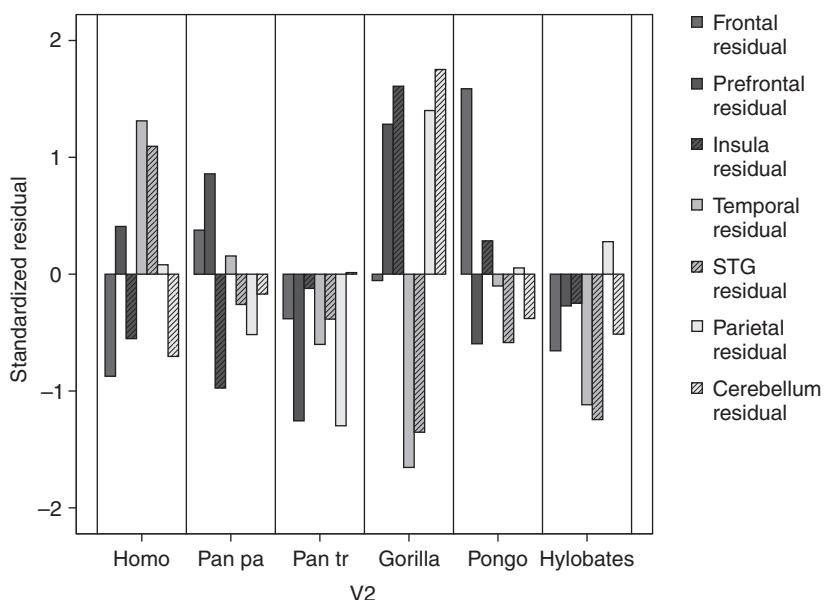


Figure 2.4 Scaling relationships in primate brains (a) Stephan et al., 1981; (b) Semendeferi et al., 2002; (c) Schoenemann et al., 2005). Slopes >1 indicate disproportionate increase in these brain regions.

but not dedicated to, social tasks (i.e., the parietal lobe, the visual cortex, the temporal lobes, and the amygdala and the hippocampus) also show lower slopes than the frontal executive regions. Semendeferi et al. (2002) argue that the human frontal lobe is not larger than expected based on scaling relationships, and that therefore there is little evidence that the human frontal lobe has disproportionately increased. However, their scaling relationships do clearly show that the slope of frontal lobe against the rest of the hemispheres is well above unity ( $B = 1.142$ ). Thus, an alternative interpretation of their data suggests that the frontal lobe overall *has* undergone a rapid increase relative to other brain structures.

We then evaluated the residuals from the scaling relationship in apes to determine where human brain regions deviate from expected sizes. Figure 2.5 demonstrates that human prefrontal and temporal lobes are larger than expected. Within the temporal lobe the STG is also larger than expected. The insula, which is believed to be important in emotional processing, appears to have undergone less expansion than



AQ19

Figure 2.5 Residuals from scaling relationships within apes. Positive residuals indicate brain measures larger than expected and negative residuals indicate regions that are smaller than expected.



expected. Additionally, the frontal lobe as a whole has increased less than expected. Although tentative, these results can be interpreted as suggesting that language and sociocognitive areas have increased disproportionately in humans. Expansion of nonsocial executive areas (frontal lobe) and emotional areas (insula), which both show superlinear scaling across available species datasets, has slowed. One intriguing interpretation of this might be that executive decision-making and socioemotional processing are fundamental ape cognitive developments and that these structures have been co-opted and linked in the prefrontal lobe to result in modern human social cognition.

### Social cognition modularity and domain general cognition

Finally, we must consider the question of whether the brain functions as a series of subprocessors or as a holistic unit. Proponents of modularity posit that the brain contains domain-specific units that are specialized to deal with particular cognitive functions. The opposite perspective, known as domain general cognition, is that the brain flexibly consolidates information across regions. This line of research is no doubt fascinating, as it provides evidence for how the brain recognizes familiar or salient objects. However, how the brain processes behavioral or episodic information, especially in social contexts, cannot be distilled down to single-neuron analyses. How does the knowledge that there are specialized subunits in the brain link to ideas of modularity and distributed cognition? Given that there appear to be particular parts of the brain that are consistently associated with interpreting social information, yet the brain is highly connected and multiple regions are invoked, perhaps the best consensus is that the brain works at both the modular and the holistic level.

Two functional networks have been proposed based on resting state activation patterns. First, an “executive control” network includes the dorsolateral prefrontal cortex (DFC) and parietal areas. Second, an emotional “salience” network includes the anterior insular cortex and the anterior cingulate cortex, the amygdala and the hypothalamus (Seeley et al., 2007). Small regions in the AIC and the medial prefrontal cortex (MPFC) are included in both networks, suggesting a strong link between emotion and cognition (Craig, 2009). The ACC has frequently been associated in functional emotional and attentional neural networks (Luu et al., 2000; Vogt et al., 1992). A key neuronal development allowing efficient communication may be the spindle cells called von





62 Susanne Shultz and Robin I.M. Dunbar

Economy neurons that are found in the ACC, dorsolateral prefrontal cortex, and frontoinsular cortex. Craig (2009) also pointed out that von Economo neurons are found only in species that can pass the mirror test (the ability to recognize oneself in a mirror, putatively a criterion of self-awareness and/or advanced cognition: de Waal, 2008). Finally, the superlinear scaling of white matter to gray matter suggests that interconnectivity is a key aspect of cognition: this has been interpreted as implying that the need to integrate neural processing units across widely distributed regions of the brain is a major constraint.

### Human evolution and brain size

Although humans have brains that are much bigger than those of other primates, early hominids had brains that were little bigger than those of modern apes. Therefore, there has been much speculation about what has driven the increase in brain size in later species. Hominids such as *Australopithecus afarensis* evolved bipedal locomotion, with no concurrent increase in brain size. In fact, it was over a million years later that there was a marked increase in brain size. So what factors could have driven the dramatic jump in brain size seen in the genus *Homo*? It is unlikely to have been driven by tool use, as evidence for more sophisticated, manufactured tools appears much later in the fossil record. Ecological arguments also do not seem to hold; the seasonality of environments where hominids were found does not tightly correlate with brain size changes. [Dunbar and Aiello \(1993\)](#) suggested that increases in group size, similar to the relationship shown in primates, could explain the increased demand on cognitive ability. Moving from a forest or woodland environment into one where predation risks are high could provide an initial pressure to increase group size. Thus, savannah-living primate species are found in larger groups than their forest counterparts (Shultz, *in press*). Coordination, relationship maintenance, and information transfer in such large groups would not only be cognitively demanding, but may also have been the impetus driving the evolution of language.

AQ20



### References

- Ackermann, H. and Riecker, A. (2004). The contribution of the insula to motor aspects of speech production: a review and a hypothesis. *Brain Lang*, 89, 320–8.
- Adolphs, R. (1999). Social cognition and the human brain. *Trends in Cognitive Science*, 3, 469–79.



# PROOF



- Aiello, L. and Dunbar, R. (1993). Neocortex size, group size and the evolution of language. *CurrAnthropol*, 34, 184–93.
- Alexander, M.P., Benson, D.F., and Stuss, D.T. (1989). Frontal lobes and language. *Brain and Language*, 37, 656–91.
- Allman, J., McLaughlin, T., and Hakeem, A. (1993). Brain weight and life-span in primate species. *PNAS*, 90, 118–22.
- Bartels, A. and Zeki, S. (2004). The neural correlates of maternal and romantic love. *Neuroimage*, 21, 1,155–66.
- Barton, R. A. (1996). Neocortex size and behavioural ecology in primates. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 263, 173–7.
- Baylis, G.C., Rolls, E.T. and Leonard, C.M. (1987). Functional subdivisions of the temporal lobe neocortex. *Journal of Neuroscience*, 7, 330–42.
- Beauchamp, G. and Fernández-Juricic, E. (2004). Is there a relationship between forebrain size and group size in birds? *Evolutionary Ecology Research*, 6, 833–42.
- Behrens, T., Hunt, L., and Rushworth, M. (2009). The Computation of Social Behavior. *Science*, 324, 1,160–4.
- Binder, J.R., Frost, J.A., Hammeke, T.A., Cox, R.W., Rao, S.M., and Prieto, T. (1997). Human brain language areas identified by functional magnetic resonance imaging. *Journal of Neuroscience*, 17, 353–62.
- Broca, P.P. (1861). Loss of speech, chronic softening and partial destruction of the anterior left lobe of the brain. *Bulletin de la Société Anthropologique*, 2, 235–8.
- Brothers, L. (1990). The social brain. *Concepts Neurosci*, 1, 27–51.
- Brothers, L. (1992). Perception of social acts in primates: cognition and neurobiology. *Seminars in Neuroscience*, 4, 409–14.
- Brothers, L., Ringa, B., and Klinga, A. (1990). Response of neurons in the macaque amygdala to complex social stimuli. *Behavioral Brain Research*, 41, 199–213.
- Brown, W. M. (2001). Natural selection of mammalian brain components. *Trends in Ecology and Evolution*, 16, 471–3.
- Bshary, R., Winkler, W., and Fricke, H. (2002). Fish cognition: a primate's eye view. *Animal Cognition*, 5, 1–13.
- Bush, G., Luu, P., and Posner, M.I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences*, 4, 215–22.
- Byrne, R. W., and Whiten, A. (Eds) (1988). *Machiavellian Intelligence*. Oxford: Oxford University Press.
- Carrington, S. J., and Bailey, A. J. (2009). Are there Theory of Mind regions in the brain? A review of the neuroimaging literature. *Human Brain Mapping*, 30, no. 2, 2, 313–35.
- Cavanna, A.E. and Trimble, M.R. (2006). The precuneus: a review of its functional anatomy and behavioral correlates. *Brain*, 129, 564–83.
- Chance, S. A., Casanova, M. F., et al. (2006). Minicolumnar structure. *Neuroscience*, 143, no. 4, 1,041–50 33.
- Chittka, L., and Niven, J. (2009). Are Bigger Brains Better? *Current Biology*, 19, R995–R1,008.
- Clark, D.A., Mitra, P.P., and Wang, S.S.H. (2001). Scalable architecture in mammalian brains. *Nature*, 411, 189–93.
- Craig, A.D. (2009). How do you feel – now? The anterior insula and human awareness. *Nature Reviews Neuroscience*, 10, 59–70.

AQ21

AQ22



# PROOF

64 Susanne Shultz and Robin I.M. Dunbar

Critchley, H.D., Melmed, R. N., Featherstone, E., Mathias, C. J., and Dolan, R. J.(2001). Brain activity during biofeedback relaxation: A functional neuroimaging investigation. *Brain* 124,1,003–12.

Damasio, H., Grabowski, T., Frank, R., Galaburda, A.M., and Damasio, A.R. (1994). The return of Phineas Gage: clues about the brain from the skull of a famous patient. *Science*, 264, 1,102–5.

Deacon, T. W. (1990). Fallacies of progression in theories of brain-size evolution. *International Journal of Primatology*, 11,193–236.

Desimone, R., Albright, T.D., Gross, C.G., and Bruce, C.J. (1984). Stimulus-selective properties of inferior temporal neurons in the macaque. *Journal of Neurosciences*, 4, 2,051–62.

AQ23 Devue, C. et al. (2007). Here I am: the cortical correlates of visual self-recognition. *Brain Res*, 1,143, 169–82.

de Waal, F. B. (2008). The thief in the mirror. *PLoS Biol*, 6, e201.

AQ24 de Winter and Oxnard (2001). Evolutionary radiations and convergences in the structural organization of mammalian brains. *Nature*, 409, 710–14.

Dunbar, R. I. M. (1992). Neocortex Size as a constraint on group-size in primates. *Journal of Human Evolution*, 22, 469–93.

Dunbar, R. I. M. (1998). The social brain hypothesis. *Evol Anthropol*, 6, 178–90.

Dunbar, R.I.M. and Shultz, S. (2007a). Understanding primate brain evolution. *Philosophical Transactions of the Royal Society B*, 29, 649–58.

Dunbar, R. I. M., and Shultz, S. (2007b). Evolution in the social brain. *Science*, 317,1,344–7.

Elliott, R., Friston, K. J., and Dolan, R. J. (2000). Dissociable neural responses in human reward systems. *J Neurosci*, 20, 6,159–65.

Ernst, M., Bolla, K., Mouratidis, M., Contoreggi, C., Matochik, J. A., Kurian, V., Cadet, J. L., Kimes, A. S., and London, E. D. (2002). Decision-making in a risk-taking task: a PET study. *Neuropsychopharmacology*, 26, 682–91.

Fink, G. R., Frackowiak, R. S., Pietrzek, U., and Passingham, R. E. (1997). Multiple nonprimary motor areas in the human cortex. *J Neurophysiol* 77,2,164–74.

Fogassi, L., Ferrari, P.F., Gesierich, B., Rozzi, S., Chersi, F., and Rizzolatti, G. (2005). Parietal Lobe: From Action Organization to Intention Understanding. *Science*, 308, 662–7.

Gallese, V., Fadiga, L., Fogassi, L., and Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, 119, 593–609.

AQ25 Gallese, V., Fadiga, L., Fogassi, L., and Rizzolatti, G. (2002). Action representation and the inferior parietal lobule. In W. Prinz, and B. Hommel (Eds),*Common Mechanisms in Perception and Action: Attention and Performance*, Vol. XIX. Oxford: Oxford University Press, pp. 247–66.

Grinband, J., Hirsch, J., and Ferrerra, V. P. (2006). A neural representation of categorization uncertainty in the human brain. *Neuron*, 49,757–63.

Harcourt, A.H. and de Waal, F.B.M. (1992). *Coalitions and alliances in humans and other animals*. Oxford: Oxford University Press.

Harvey, P.H., Clutton-Brock, T.H. and Mace, G.M. (1980). Brain size and ecology in small mammals and primates. *PNAS*, 77,4,387–9.

Healy, S.D., and Rowe, C. (2007).A critique of comparative studies of brain size. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 274, 453–64.

# PROOF



- Hecaen, H. and Albert, M. L. (1978). *Human Neuropsychology*. New York: J. Wiley.
- Hickok, G. (2008). Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *Journal of Cognitive Neuroscience*, 21, 1,229–43.
- Hsu, M., Bhatt, M., Adolphs, R., Tranel, D., and Camerer, C.F. (2005). Neural systems responding to degrees of uncertainty in human decision-making. *Science*, 310, 1,680–3.
- Huettel, S. A., Song, A. W., and McCarthy, G. (2005). Decisions under uncertainty: probabilistic context influences activation of prefrontal and parietal cortices. *J Neurosci*, 25, 3,304–11.
- Huettel, S. A., Stowe, C. J., Gordon, E. M., Warner, B. T., and Platt, M. L. (2006). Neural signatures of economic preferences for risk and ambiguity. *Neuron*, 49, 765–75.
- Humphrey, N.K. (1976). The social function of intellect. In P.P.G. Bateson and R.A. Hinde(Eds),*Growing points in ethology*. Cambridge, UK: Cambridge University Press, pp. 303–17.
- AQ26 Jerison, H. J. (1973). *Evolution of the Brain and Intelligence*. Academic Press.
- Jolly, A. (1966). Lemur social behavior and primate intelligence. *Science*, 29, 501–6.
- Kanwisher, N., McDermott, J., and Chun, M.M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *JNeurosci*, 17, 4,302–11.
- AQ27 Kendrick, K.M., da Costa, A.P., Leigh, A.E., Hinton, M.R., and Peirce, J.W. (2001). Sheep don't forget a face. *Nature*, 414, 165–6.
- Keysers, C., and Gazzola, V. (2006). Towards a unifying neural theory of social cognition. *Prog Brain Res*, 156,379–401.
- Koelsch, S., Fritz, T., v. Crimon, D. Y., Muller, K., and Friederici, A. D. (2006). Investigating emotion with music: an fMRI study. *Hum Brain Mapp*, 27, 239–50.
- Kudo, H., and Dunbar, R. I. M. (2001). Neocortex size and social network size. *AnimBehav*, 62, 711.
- AQ28 Luu, P. et al. (2000). Medial frontal cortex in action monitoring. *JNeurosci*, 20, 464–9.
- Mahon, B. Z., and Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology (Paris)*, 102, 59–70.
- Martin, R. D. (1981). Relative brain size and basal metabolic-rate in terrestrial vertebrates. *Nature*, 293, 57–60.
- Miller, E.K. and Cohen, J.D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, 24,167–202.
- O'Doherty, J., Kringelbach, M.L., Rolls, E.T., Hornak, J., and Andrews, C.(2001). Abstract reward and punishment representations in the human orbitofrontal cortex. *Nature Neuroscience*, 4, 95–102.
- Parr, L.A. and De Waal, F.B.M. (1999). Visual kin recognition in chimpanzees. *Nature*, 399, 647–8.
- Paulus, M. P., Rogalsky, C., and Simmons, A. (2003). Increased activation in the right insula during risk-taking decision making is related to harm avoidance and neuroticism. *NeuroImage*, 19, 1,439 –48.



# PROOF

66 Susanne Shultz and Robin I.M. Dunbar

Pawlowski, B., Lowen, C. B., and Dunbar, R. I. M. (1998). Neocortex size, social skills and mating success in primates. *Behaviour* 135, 357–68.

Penfield, W., and Faulk, M. E. (1955). The insula; further observations on its function. *Brain*, 78, no. 4, 445–70.

Perrett, D. I., Smith, P. A., Potter, D. D., Mistlin, A. J., Head, A. S., Milner, A. D., and Jeeves, M. A. (1985). Visual cells in the temporal cortex sensitive to face view and gaze direction. *Proc R Soc Lond B Biol Sci*, 223, no. 1,232, 293–317.

Perrett, D.I., Mistlin, A.J., Harries, M.H., and Chitty, A.J. (1990). *Understanding the visual appearance and consequence of hand action. Vision and action: the control of grasping*. Ablex: Norwood, New Jersey.

AQ29 Platel, H. et al. (1997). The structural components of music perception. A functional anatomical study. *Brain*, 120, 229–43.

AQ30 Puce, A. et al.(1998). Temporal cortex activation in humans viewing eye and mouth movements. *JNeurosci*, 18, 2,188–99.

Quiroga, R.Q. and Kreiman, G.(2010). Postscript: about grandmother cells and Jennifer Aniston neurons. *Psychological Review*, 117, 297–9.

Quiroga, R.Q., Reddy, L., Kreiman, G., Koch, C., and Fried, I. (2005). Invariant visual representation by single neurons in the human brain. *Nature*, 435, 1,102–7.

Rilling, J. K. and Insel, T. R. (1998). Evolution of the cerebellum in primates: differences in relative volume among monkeys, apes and humans. *Brain, Behavior & Evolution*, 52, 308–14.

Rizzolatti, G., and Craighero, L. (2004). The Mirror-Neuron System. *Annual Rev Neurosci*, 27, 169–92.

Rizzolatti, G., Fogassi, L., and Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nat Rev Neurosci*, 2, 661–70.

Rolls, E. T., Hornak, J., Wade, D., and McGrath, J. (1994). Emotion-related learning in patients with social and emotional changes associated with frontal lobe damage. *J NeurolNeurosurgPsychiat*, 57, 1, 518–24.

AQ31 Schoenemann, P.T., Sheehan, M.J., and Glotzer, L. D. (2005). Prefrontal white matter volume is disproportionately larger in humans than in other primates. *Nature Neuroscience*, 8, 242–52.

Seeley, W.W., Menon, V., Schatzberg, A.F., Keller, J., Glover, G.H., Kenna, H., Reiss, A.L., and Greicius, M.D. (2007). Dissociable intrinsic connectivity networks work for salience processing and executive control. *JNeurosci*, 27, 2,349–56.

Semendeferi, K., Damasio, H., and Frank, R. (1997). The evolution of the frontal lobes: a volumetric analysis based on three-dimensional reconstructions of magnetic resonance scans of human and ape brains. *Journal of Human Evolution*, 32, 375–88.

Shimeld, S.M. and Holland, P.W.H. (2000). Vertebrate innovations. *PNAS*, 97, 4,449–52.

Shultz, S. (in press). Primate social evolution: bonding and dispersal patterns revisited. *International Journal of Primatology*.

Shultz, S., and Dunbar, R. I. M. (2007). The evolution of the social brain: anthropoid primates contrast with other vertebrates. *Proceedings of the Royal Society, London B*, 274,2,429–36.

# PROOF



Shultz, S. and Dunbar, R. I. M. (2010). Species differences in executive function correlate with brain size across non-human primates. *Journal of Comparative Psychology*, 124, 252–60.

Stephan, H., Frahm, H., and Baron, G. (1981). New and revised data on volumes of brain structures in insectivores and primates. *Folia Primatologica*, 35, 1–29.

Stuss, D.T., and Benson, D.F. (1986). *The frontal lobes*. New York: Raven Press.

Swanson, L.W. and Petrovich, G.D. (1998). What is the amygdala? *Trends in Neurosciences*, 21, 323–31.

AQ32 Vogt, B.A. et al. (1992). Functional heterogeneity in cingulate cortex: the anterior executive and posterior evaluative regions. *Cereb Cortex*, 2, 435–43.

Wernicke, C. (1874). *Der Aphasische Symptomencomplex*. Breslau: Cohn and Weigert.

Wyles, J.S., Kunkelt, J.G., and Wilson, A.C. (1983). Birds, behavior, and anatomical evolution: rates of evolution/nongenetic propagation of new habits/brain size. *PNAS*, 80, 4, 394–7.

Zhang, K., and Sejnowski, T.J. (2000). A universal scaling law between gray matter and white matter of cerebral cortex. *PNAS*, 97, 5, 621–6.





## 3

### *Homo heuristicus* and the Bias – Variance Dilemma

*Henry Brighton and Gerd Gigerenzer*

#### Introduction

*Homo heuristicus* makes inferences in uncertain environments using simple heuristics that ignore information (Gigerenzer and Brighton, 2009). Traditionally, heuristics are seen as second-best solutions which reduce effort at the expense of accuracy, and lead to systematic errors. The prevailing assumption is that, to understand the ability of humans and other animals to cope with uncertainty, one should investigate cognitive models that optimize. We introduced the term *Homo heuristicus* to highlight several reasons why this assumption can be misleading, and argue that heuristics play a critical role in explaining the ability of organisms to make accurate inferences from limited observations of an uncertain and potentially changing environment. In this chapter we use examples to sketch the theoretical basis for this assertion, and examine the progress made in the development of *Homo heuristicus* as a model of human decision-making.



#### What are heuristics and why study them?

Although frustratingly little detail is known about the mechanisms used by organisms to make inferences, illuminating insights into animal decision-making do exist. The ant species *Leptothorax albipennis* inhabit flat rock crevices, and select their nest site from a range of alternatives by estimating the relative area of potential nest using a heuristic. First, an ant will run around a potential nest on an irregular path for a fixed period of time while laying down a pheromone trail, and then leave. Later, the ant returns and runs around on a different irregular path, and then estimates the size of the site using the frequency with



# PROOF



Homo heuristicus and the Bias – Variance Dilemma 69

which it reencounters the old trail. This heuristic is remarkably precise: nests half the area of others yielded reencounter frequencies 1.96 times greater (Mugford et al., 2001). Peahens also use a heuristic when choosing among potential mates. Rather than examining all the available peacocks, a peahen will investigate only three or four, and then choose the one with the largest number of eyespots (Petrie and Halliday, 1994). These two examples illustrate how organisms solve problems in an uncertain world using heuristics (see Hutchinson and Gigerenzer, 2005, for further examples).

What is a heuristic? The examples above highlight two hallmarks of heuristics: one-reason decision-making and limited search. Specifically, peahens could in principle integrate many features which describe potential mates, such as their size, dominance within the group, or volume of mating call. Instead, they use just one feature. Peahens could also exhaustively examine all the potential mates, but they instead consider just a handful. In the same way, ants tend to revisit potential nest sites a couple of times, rather than repeatedly approximating reencounter frequency. More generally, heuristics are best seen in contrast to optimization processes, which attempt to gather all available information, integrate this information, and derive predictions based on potentially complex models of the underlying processes which govern the observations. Heuristics, in contrast, are examples of what Herbert Simon (e.g., 1955; 1991) referred to as *satisficing* processes (a Northumbrian word for “satisfying”). Satisficing is the process of seeking a good-enough solution rather than seeking an optimal solution. For example, when selecting a good alternative from a series of options encountered sequentially, a satisficer might set an aspiration level, choose the first option that meets or exceeds this aspiration level, and then terminate search. Although heuristics tend to consume fewer processing resources as a result of ignoring information, this does not mean that they are less accurate than processes which consume more processing resources. Before considering these issues in greater detail, it is worth spelling out the relationship between optimal solutions, optimizing processes, and uncertainty.

### Optimal solutions vs. the process of optimization

Consider a tin can manufacturer attempting to reduce costs by minimizing the surface area of the cans it produces. To package 12 ounces of soup, the manufacturer has calculated the height and width of the can which minimizes the amount of tin used. No other design uses less tin to package the same soup. This is an example of an optimal solution



70 *Henry Brighton and Gerd Gigerenzer*

to a problem: from the space of candidate solutions, the optimal solution is one which cannot be improved on. In this example, the relationship between the variables in question and the solution is certain. Solid geometry, coupled with the ability to take precise measurements, ensures a close fit between our model and the real world. The term “optimal” refers to such a state, but the term “optimization” refers to the process of searching for the optimal state. If we lacked knowledge of solid geometry, iteratively fine-tuning the tin can dimensions until the surface area is at a minimum would be a process of optimization. This process assumes that we can measure the effect of our actions, and this measure serves as a proxy for performance. If we had knowledge of solid geometry, then we could derive the optimal solution directly. Both are examples of optimization.

Broadly speaking, optimization is any process which explicitly attempts to maximize some criterion assumed to be monotonically related to performance. As a consequence, optimization methods have a tendency to assume that more computation leads to greater precision. Optimization models are common in the study of cognition. Neural network models of cognition, for instance, usually rely on algorithms which attempt to minimize the error of the network by iteratively fine-tuning the strength of its synaptic weights (Rumelhart et al., 1986). While optimization is certainly a valid, widely practiced, and successful approach, it is not the only approach. Following Simon, we will use the term satisficing to refer to processes which do not optimize. Heuristics are precise specifications of these non-optimizing solutions (Gigerenzer et al., 1999).

### Beyond optimization

Optimization methods are both widely studied and successful. Why, then, study heuristics? First of all, we may observe an organism satisficing, and focus on uncovering a precise understanding of the heuristic mechanism used by the organism. These findings lead naturally to the question of how ubiquitous satisficing mechanisms are in the natural world, and spur the search for underlying functional and information processing principles explaining why and how organisms might satisfice. For instance, the most common justification for heuristic processing is reduced resource consumption. Organisms consume resources such as time and energy when processing information, and, because these resources are limited, it is reasonable to assume a trade-off exists between accuracy and effort. Heuristics, according to this view, allow the decision-maker to make accurate inferences without expending the additional resources required to make the best inferences. This

# PROOF



Homo heuristicus and the Bias – Variance Dilemma 71

justification is widely accepted, and rooted in the commonly assumed principle of the *accuracy–effort trade-off*:

**Accuracy–effort trade-off:** information and computation cost time and effort; therefore minds rely on simple heuristics that are less accurate than strategies that use more information and computation.

Much of the following discussion will center on more fundamental reasons for examining heuristics. First, some problems are computationally intractable, and require, of necessity, decision-makers to satisfice. For instance, no human or machine can implement the optimal chess-playing strategy, due to its explosively large search space. For such problems, the only option may be to use heuristics. Second, another reason to look beyond optimization is the inherent uncertainty of natural environments. Organisms regularly face problems for which the optimal solution – the underlying data generating distribution, for example – is unknown or unknowable. Optimization in a world which is uncertain in this way is still possible, but less appealing, since we must knowingly optimize a misspecified criterion. Issues such as these open the door to alternative, potentially more appropriate, and superior information-processing models. This possibility leads to another reason for studying heuristics, which is to identify and explain less-is-more effects:



**Less-is-more effects:** more information or computation can decrease accuracy; therefore, minds rely on simple heuristics in order to be more accurate than strategies that use more information and time.

The occurrence of less-is-more effects tells us that strategies which ignore information and limit search have the potential to better explain how organisms make accurate inferences in uncertain environments. Next, we will explore less-is-more effects, and work toward the study of the ecological rationality of heuristics, which examines in which environments a given strategy succeeds or fails, and why. These issues form part of a broader research program which aims at a systematic theory of heuristics that identifies their building blocks and the evolved capacities they exploit, and views the cognitive system as relying on an “adaptive toolbox” of heuristics.

## Less-is-more effects

The term heuristic is often used with a negative connotation, suggesting a second-best solution to a problem better addressed by a more



# PROOF

72 *Henry Brighton and Gerd Gigerenzer*

principled, optimizing solution. The use of heuristics by people, according to this view, explains a number of human reasoning errors. The “Heuristics and biases” approach of Daniel Kahneman, Amos Tversky, and their collaborators emphasized that heuristics are sometimes good and sometimes bad, but placed a heavy focus on experiments designed to show that people often violate laws of logic, probability, or some other standard of rationality (Tversky and Kahneman, 1974). The association between heuristics and shoddy mental software is rooted in three widespread misconceptions:

1. Heuristics are always second-best.
2. We use heuristics only because of our cognitive limitations.
3. More time, more information, and more computation would always be better.

These three beliefs assume that the accuracy–effort trade-off, described above, holds. The inaccuracy of this overly simplistic picture can be demonstrated by considering perhaps the most widely used statistical model: the linear model fitted using least squares. Use of linear regression has become automatic among sociologists, economists, and psychologists when making inferences about their observations. Linear regression estimates the optimal beta weights for the predictors. In the 1970s, researchers discovered that unit weights (-1 or 1), or even random weights, can predict almost as accurately as, and sometimes better than, multiple linear regression (Dawes, 1979; Dawes and Corrigan, 1974; Einhorn and Hogarth, 1975; Schmidt, 1971). These less-is-more phenomena came as a surprise to the scientific community. When Robin Dawes presented the results at professional conferences, distinguished attendees told him that they were “impossible,” his paper with Corrigan was first rejected and deemed “premature;” a sample of recent textbooks in econometrics revealed that none referred to the findings of Dawes and Corrigan (Hogarth, in press).

Before taking a closer look at less-is-more effects, what exactly does “more” refer to in the term “less-is-more?” For problems of inductive inference, the decision-maker’s task is to process a series of observations with a view to identifying patterns of regularity among these observations. Identifying systematic regularities allows the decision-maker to make accurate predictions about novel or future observations. Consider, for instance, a weather forecaster observing past temperature trends in an attempt to spot a pattern which can be used to make better predictions. Problems like these highlight a crucial distinction



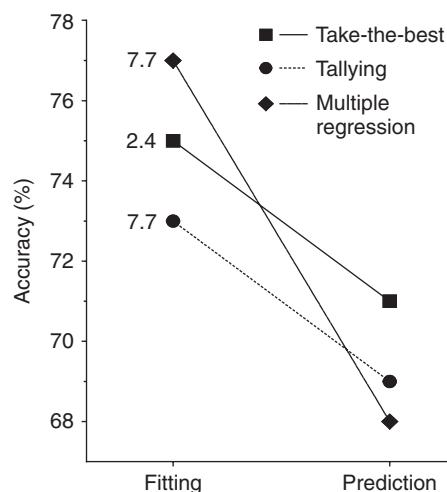
between two kinds of accuracy. First, *data fitting* is the ability to accurately describe already observed data. Experts find it alarmingly easy to explain, post hoc, why several hurricanes occurred in the previous year, or why the stock market crashed. *Data prediction* is much harder, and tells us to evaluate weather forecasters and financial experts by taking note of their past predictions, and checking how accurately they in fact predicted future events. Data prediction is the true test of a decision-maker's ability to make inductive inferences. In exactly the same way, an organism's inference mechanisms should help to accurately second-guess future events in its environment, rather than accurately describe past events. In short, when we refer to less-is-more effects in inductive inference, the "more" will refer to *predictive* accuracy, and the "less" will refer to various forms of ignoring information and limiting search.

#### Further less-is-more effects

Using predictive accuracy as the performance criterion, Czerlinski et al. (1999) conducted 20 studies comparing unit weighted regression (also known as tallying) and multiple regression. The models were compared using cross-validation, a process which repeatedly partitions the observations into one set used to estimate the model parameters, and another set used to measure the predictive accuracy of the fitted models (Stone, 1974). Specifically, Czerlinski et al. examined paired comparison tasks in which, for instance, the problem is to estimate which of two Chicago high schools will have a higher drop-out rate, based on cues such as writing score and proportion of Hispanic students. Ten of the 20 data sets were taken from a textbook on applied multiple regression (Weisberg, 1985). Averaged across all data sets, tallying achieved a higher predictive accuracy than multiple regression (Figure 3.1). Regression tended to overfit the data, as can be seen by the cross-over of lines: it had a higher fit than tallying, but a lower predictive accuracy.

These results illustrate that under certain circumstances tallying leads to higher predictive accuracy than multiple regression. They also illustrate that claims of universal superiority of one statistical model over another are rarely, if ever, true. Instead, they highlight the need to know in which environments simple tallying is more accurate than multiple regression, and in which environments it is not. This is the question of the *ecological rationality* of tallying. Early attempts to answer this question concluded that tallying succeeded when linear predictability of the criterion was moderate or small ( $R^2 \leq 0.5$ ), the ratio of objects to cues was 10 or smaller, and the cues were correlated





*Figure 3.1* Less-is-more effects. Both tallying and take-the-best predict more accurately than multiple regression, despite using less information and computation. Note that multiple regression excels in data fitting (“hindsight”), that is, fitting its parameters to data that is already known, but performs relatively poorly in prediction (“foresight,” as in cross-validation). Take-the-best is the most frugal, that is, it looks up, on average, only 2.4 cues when making inferences. In contrast, both multiple regression and tallying look up 7.7 cues on average. The results shown are averaged across 20 studies, including psychological, biological, sociological, and economic inference tasks (Czerlinski et al., 1999).

(Einhorn and Hogarth, 1975). The discovery that tallying can often match and even outperform complex calculations is important in understanding the nature of adaptive cognition. To what extent do findings such as these provide hints as to how the cognitive system makes inferences with the limited resources it has available? Note that the conditions under which tallying succeeds – low predictability of a criterion, small sample sizes relative to the number of available cues, and dependency between cues – is highly relevant for decision-making in natural environments.

Restricting attention to unit weights is one of many approaches to simplifying the basic linear model. Another simplification is to restrict attention to only a single cue when making decisions. The heuristic take-the-best uses this simplification when deciding between objects in a paired comparison task. First, take-the-best estimates the validity of each cue. Cue validity is the proportion of correct inferences the cue makes on previously observed comparisons. To make inferences



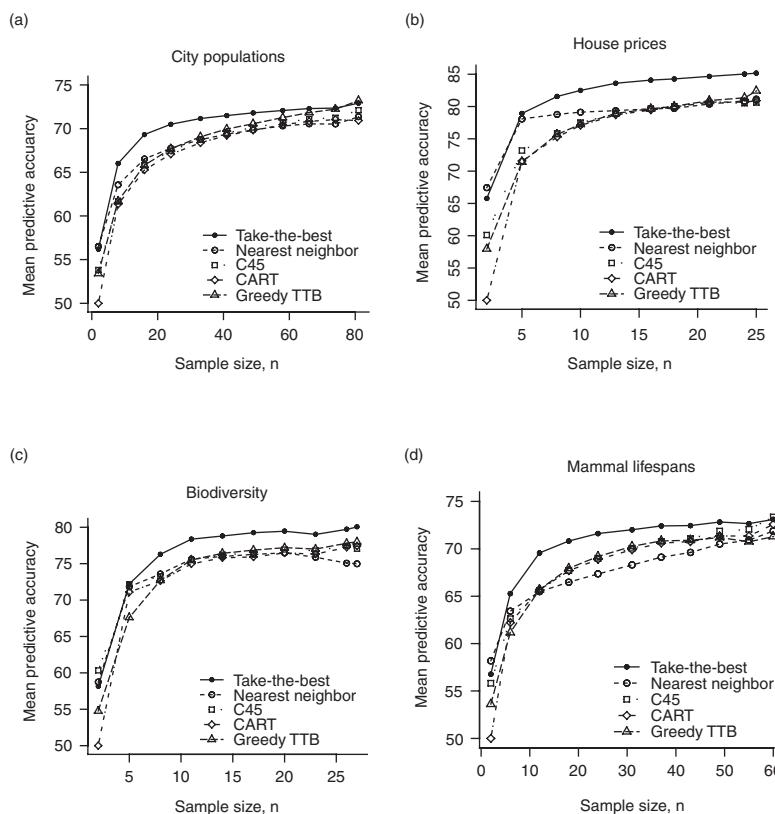
on novel, unobserved comparisons, take-the-best searches the cues in validity order sequentially until it finds a discriminating cue. A cue discriminates when the objects in question have different values for this cue, and therefore allow a decision to be made. Take-the-best always makes decisions on the basis of a single discriminating cue, although several cues may be examined before this cue is chosen. Figure 3.1 compares the performance of take-the-best with tallying and linear regression. Again, a less-is-more effect can be seen, but now the effect is more dramatic: take-the-best performs poorly in fitting the observations, but achieves a higher predictive accuracy than both tallying and linear regression.

Some researchers hypothesized that these less-is-more effects exploit the weakness of linear regression as a competitor. To address this point, we compared take-the-best with four additional inference strategies (Brighton, 2006). First, the *nearest neighbor classifier* is a nonlinear, widely used, and widely studied classification method which is known to perform well across many domains (Cover and Hart, 1967). Second, the tree induction algorithm *C4.5* is also widely studied and explicitly attempts to combat the problem of overfitting by pruning the decision trees it induces (Quinlan, 1993). Third, we compared take-the-best with another tree induction algorithm, *CART*, which tends to produce yet smaller trees than C4.5 (Breiman et al., 1994). Fourth, take-the-best will be compared with a slightly modified version of itself – *greedy take-the-best* – which commits the additional resources required to assess conditional dependencies between cues (Martignon and Hoffrage, 2002; Schmitt and Martignon, 2006). This variant of take-the-best orders cues by taking into account that cue validities, when estimated conditionally on the value of other cues, can differ [from the basic] validity estimates used by take-the-best. These estimates of conditional validity result in the induction of cue orders which differ from those selected by take-the-best.

AQ1

Figure 3.2 compares the performance of take-the-best with these four alternative strategies in four environments taken from the Czerlinski et al. (1999) study. In all four environments, take-the-best outperforms the alternative mechanisms for most, if not all, sample sizes. Two points are worth making here. First, the less-is-more effects shown in Figure 3.1 and Figure 3.2 are to a certain degree robust, rather than hinging on a potentially “straw man” comparison with linear regression. Second, and more generally, it is statistically obvious that in some environments, and against some competitors, take-the-best (like any other model of inductive inference) will perform poorly.





*Figure 3.2* The performance of take-the-best in comparison to three well-known learning algorithms (nearest neighbor classifier, C4.5, and CART) and the greedy version of take-the-best, which orders cues by conditional validity. Mean predictive accuracy in cross-validation is plotted as a function of the size of the training sample for the task of deciding (a) which of two German cities has the larger population; (b) which of two houses has the higher price; (c) which of two Galapagos islands has greater biodiversity; and (d) which of two mammals is likely to live longer. These environments are taken from the study by Czerlinski et al. (1999).

As we noted above, the real question is to understand when and why certain forms of processing perform well, and when and why they perform poorly. Of particular interest here is the question of when and why ignoring information leads to superior performance, as we saw above.



## Explaining less-is-more effects

Understanding *Homo heuristicus* requires understanding when and why less-is-more effects hold, and how an organism might exploit the existence of less-is-more effects as part of its design. The policies of simplifying weights and one-reason decision-making both reduce effort. They also increase accuracy, clearly demonstrating that the effort–accuracy trade-off is invalid as a general rule. Findings like these provide hints as to how organisms might make inferences in uncertain environments, where the task is to use limited observations to second-guess the predictive patterns underlying the observations. To understand the functioning of heuristics, we will again consider the task faced by a weather forecaster.

### Over- and underfitting

The temperature in London on a given day of the year is uncertain but follows a seasonal pattern. Using the year 2000 as an example, we have plotted London's mean daily temperature in Figure 3.3(a). On top of these observations we have plotted two polynomial models that attempt to capture a systematic pattern in London's temperatures. The first model is a degree-3 polynomial (a cubic equation with four parameters), and the second is a degree-12 polynomial (which has 13 parameters). Comparing these two models, we see that the degree-12 polynomial captures monthly fluctuations in temperature while the degree-3 polynomial captures a simpler pattern charting a rise in temperature that peaks in the summer, followed by a slightly sharper fall.

If the weather forecaster is only interested in describing the past as accurately as possible, then picking the polynomial model that fits the data with the least error is the best option. This criterion would prefer the degree-12 polynomial. More generally, to maximize goodness of fit, higher and higher-degree polynomials could be chosen, with each added degree improving the ability of the model to capture minor fluctuations in temperature over smaller and smaller timescales. If London's daily temperatures for all subsequent years were guaranteed to match precisely those measured in the year 2000 then this approach would be ideal, because what we have observed in the past will continue to be observed in the future, and by describing the past more accurately, as with a higher-degree polynomial, we will also describe the future more accurately. The future is certain in this hypothetical world. As soon as uncertainty enters the picture, using goodness of fit to judge this ability





78 *Henry Brighton and Gerd Gigerenzer*

is a dangerous practice which will often lead to faulty conclusions (Pitt et al., 2002; Roberts and Pashler, 2000).

To estimate how well different models predict the future, based on observations of the past, we can reframe the task by estimating model parameters using only a sample of the observations and then test how well such models predict novel instances of the problem. More specifically, if we observe the temperature on 50 randomly selected days in the year 2000 and then fit a series of polynomial models of varying degree to this sample, we can measure how accurately each model goes on to predict the temperature on those days we did not observe in the year 2000. As a function of the degree of the polynomial model, the mean error in performing this prediction task is plotted in Figure 3.3(b). The model with the lowest mean error (with respect to many such samples of size 50) is a degree-4 polynomial – which shows that more complexity is not better. In short, Figure 3.3(b) tells us that the error in fitting the observations decreases as a function of the degree of the polynomial, which means that the best-predicting model would not have been chosen if we had judged models merely by checking how well they fit the observations. The most predictive model is very close to the lower bound of complexity, rather than at some intermediate or high level.

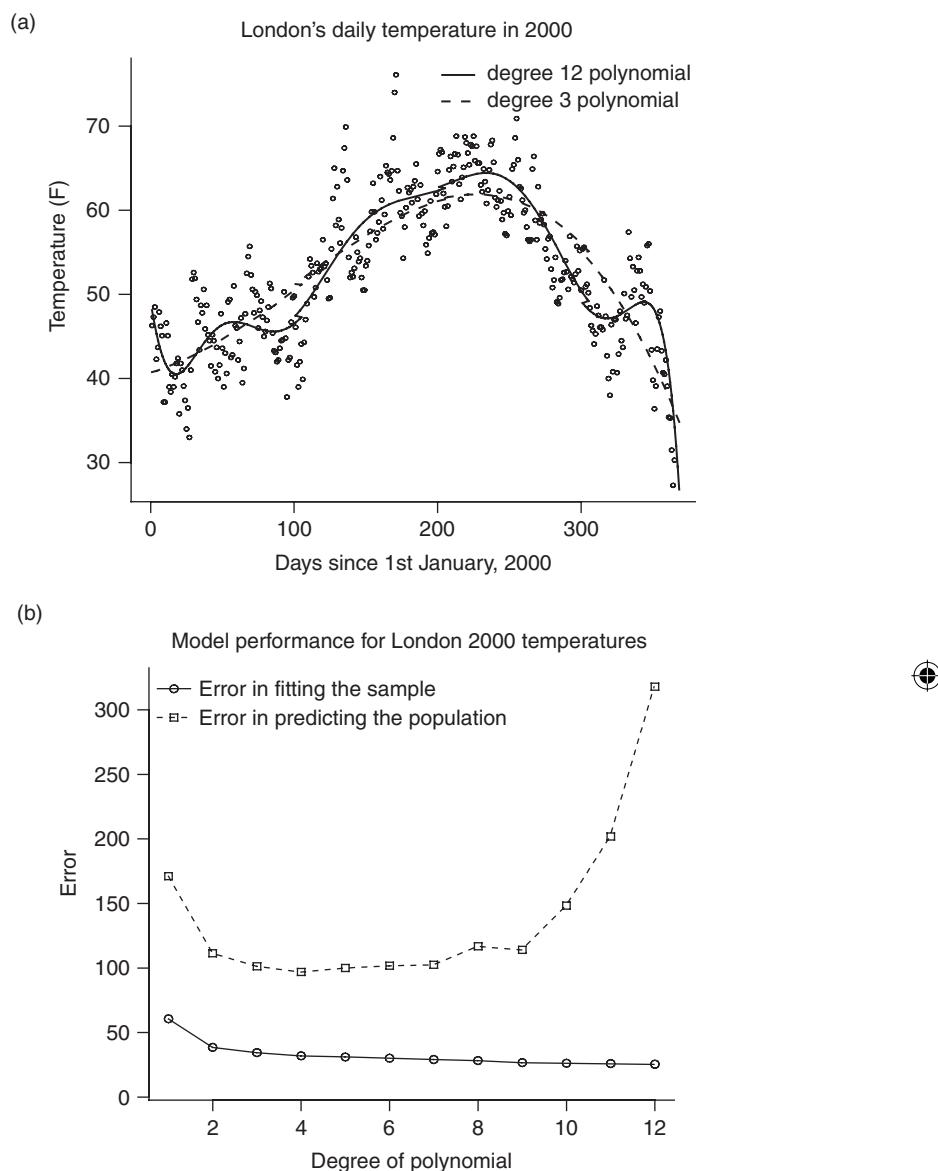
### The bias–variance dilemma

Understanding how properties of a decision-maker's learning algorithm interact with properties of its task environment is a crucial step toward understanding how organisms can deal with uncertainty and error. To understand this problem we will adopt the perspective of an omniscient observer and consider the bias–variance dilemma (Geman et al., 1992; Hastie et al., 2001), a statistical perspective on the problem of inductive inference that decomposes prediction error into three components: a bias component, a variance component, and a noise component. Total prediction error is the sum of the following three terms:

$$\text{Error} = (\text{bias})^2 + \text{variance} + \text{noise}$$

This decomposition clarifies different sources of error, and how they are related to the properties of the learning algorithm. To illustrate this relationship, we will revisit the daily temperature example but change the rules of the game. The “true” underlying function behind London’s mean daily temperatures is unknown. Nevertheless, we will put ourselves in the position of grand planner with full knowledge of the underlying function for the mean daily temperatures in some fictional





*Figure 3.3* Plot (a) shows London's mean daily temperature in 2000, along with two polynomial models fitted using the least squares method. The first is a degree-3 polynomial, the second a degree-12 polynomial. Plot (b) shows both the mean error in fitting samples of 30 observations and the mean prediction error of the same models, both as a function of degree of polynomial.

# PROOF

80 *Henry Brighton and Gerd Gigerenzer*

location. We denote this degree-3 polynomial function  $h(x)$  and define it as

$$h(x) = 37 + \frac{15}{365}x + \frac{120}{365}x^2 + \frac{130}{365}x^3, \text{ where } 0 \leq x \leq 364.$$

Figure 3.4(a) plots this underlying trend for each day of the year. We will also assume that, when  $h(x)$  is sampled, our observations suffer from normally distributed measurement error with  $\mu = 0$  and  $\sigma^2 = 4$ . A random sample of 30 observations of  $h(x)$  with this added error is shown on top of the underlying trend in Figure 3.4(a). If we now fit a degree- $p$  polynomial to this sample of observations, and measure its error in approximating the function  $h(x)$ , can we draw a conclusion about the ability of degree- $p$  polynomials to fit our “true” temperature function in general? Not really, because the sample we drew may be unrepresentative: it could result in a lucky application of our fitting procedure that perfectly models the underlying polynomial  $h(x)$ , or an unlucky one which results in high error. This single sample may not lead to a representative picture of the performance of degree- $p$  polynomials in general, after other samples are taken into account.

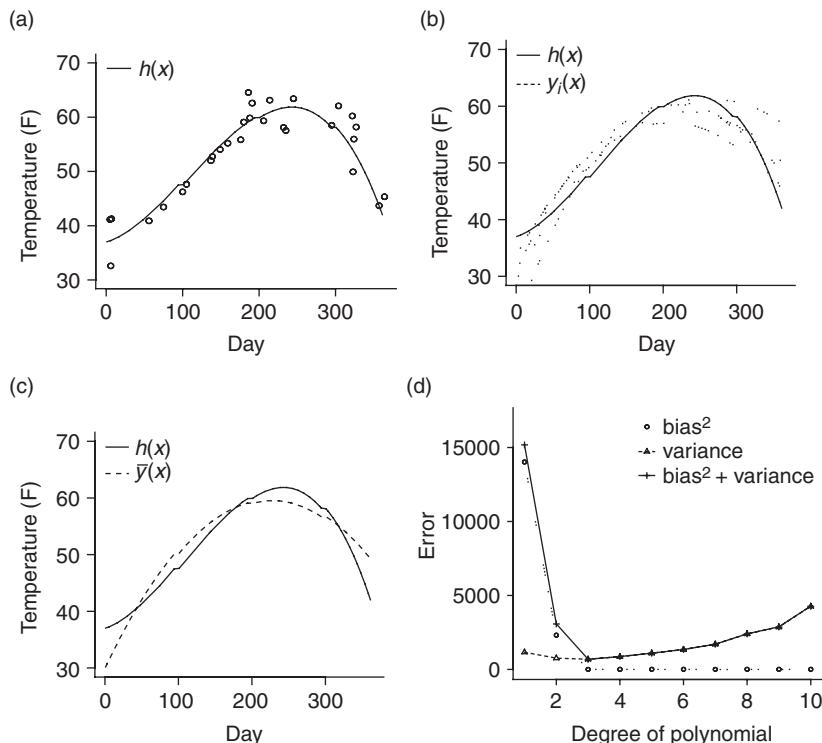
A more reliable test of a model is to measure its mean accuracy by taking  $k$  random samples of size  $n$ , fitting a degree- $p$  polynomial model to each one, and then considering this ensemble of models denoted by  $y_1(x), y_2(x), \dots, y_k(x)$ . Figure 3.4(b) shows five polynomials of degree 2 resulting from  $k = 5$  samples of  $n = 30$  observations of  $h(x)$ . From the perspective of the organism, these samples can be likened to separate encounters with the environment, and the fitted polynomials likened to the responses of the organism to these encounters. The question now is how well a given type of model – here polynomials of degree 2 – captures the underlying function  $h(x)$ , which we can estimate by seeing how well the induced models perform on average, given their individual encounters with data samples. First consider the function  $\bar{y}(x)$ , which for each  $x$  gives the mean response of the ensemble of  $k$  polynomials:

$$\bar{y}(x) = \frac{1}{k} \sum y_i(x).$$

The *bias* of the model is the sum squared difference between this mean function and the true underlying function. Our omniscience is important now, because to measure the difference we need to know the underlying function  $h(x)$ . More precisely, bias is given by

$$(bias)^2 = \sum \{\bar{y}(x^n) - h(x^n)\}^2.$$

# PROOF



*Figure 3.4* A fictional daily temperature function  $h(x)$  used to illustrate bias and variance. (a) Graph of  $h(x)$  and a sample of 30 points with added noise. (b) Five polynomials of degree-2,  $y_i(x)$  for  $1 \leq i \leq 5$ , fitted to five further samples. (c) Mean of these five functions,  $\bar{y}(x)$ . Bias is the squared difference between  $h(x)$  and  $\bar{y}(x)$ . Variance is the sum of the squared difference between each function  $y_i(x)$  and  $\bar{y}(x)$ , measuring how much the induced functions vary about their mean. Plot (d) shows, as a function of degree of polynomial, the mean error in predicting the temperature on those days not in the observed samples, after fitting polynomials to samples of 30 noisy observations. This error is decomposed into bias and variance, also plotted as function of degree of polynomial.

Figure 3.4(c) shows the  $\bar{y}(x)$  arising from the five polynomials shown in figure 3.4(b). Assuming  $k = 5$  is sufficient to provide us with a good estimate of  $\bar{y}(x)$ , this plot tells us that the model is biased, since it differs from  $h(x)$ . Zero bias is achieved if our average function is precisely the true function. Bias usually occurs when the model we use to explain the observations lacks the appropriate functional form to represent the true underlying function. In the absence of knowledge about the underlying function, bias can be reduced by making the space of models considered

# PROOF

82 *Henry Brighton and Gerd Gigerenzer*

by the learning algorithm sufficiently rich. However, this policy can simply replace one problem with another. Although the mean function response of the ensemble may capture the true underlying function without error, the individual models that contribute to this mean may each incur high error. That is, zero mean error can mask high error of the individual models. This source of error, which arises from the sensitivity of the learning algorithm to the contents of individual samples, is termed *variance*. Variance is the mean squared difference between each induced model function and the mean function:

$$\text{variance} = \sum_n \frac{1}{k} \sum \{y_i(x^n) - \bar{y}(x^n)\}^2.$$

When variance increases as we consider more complex models, we say that these models are overfitting the data. The two properties of bias and variance reveal that the inductive inference of models involves a fundamental trade-off. We can aim to use a general purpose learning algorithm, such as a feed-forward neural network, that employs a wide and rich space of potential models, which more or less guarantees low bias. But, when we have a limited number of observations, the flexibility of the model space can incur a cost in high variance, since the learning algorithm is likely to induce a model that captures unsystematic variation. To combat high variance we can place restrictions on the model space and thereby limit the sensitivity of the learning algorithm to the vagaries of samples. But these restrictions run counter to the objective of general purpose inference, since they will necessarily cause an increase in bias for some problems.

This is the bias-variance dilemma. The balancing act required to achieve both low variance and low bias is clear in figure 3.4(d), which decomposes the error arising from polynomials from degree 1 (a straight line) to degree 10 at predicting our temperature function  $h(x)$  from samples of size 30. For each polynomial degree we have plotted the bias (squared) of this type of model, its variance, and their sum. The polynomial degree that minimizes the total error is, not surprisingly, 3, because  $h(x)$  is a degree-3 polynomial. Polynomial models of less than degree 3 suffer from bias, since they lack the ability to capture the underlying pattern. Polynomials of degree 3 or more have zero bias, as we would expect. But for polynomials of degree 4 or more the problem of overfitting arises, and their variance begins to increase due to their excess complexity. None of the models achieve zero error. This is due to the observation error we added when sampling, which corresponds to the noise term in the bias-variance decomposition.



### Heuristics and the bias–variance dilemma

Bias and variance provide an insightful framework for examining the inference tasks faced by organisms. In particular, how do organisms control bias and variance, and keep them within acceptable limits? More specifically, how can the functioning of heuristics be understood in these terms? Recall Figure 3.1 and Figure 3.2, which showed how take-the-best outperformed several alternative inference strategies. Now we will analyze take-the-best by performing a bias–variance decomposition of its error. We will also decompose the error of its greedy counterpart, which we described above. As seen in Figure 3.2, the performance of the neural, exemplar, and decision tree models tend to be very similar to each other in paired comparison tasks, which in turn are very similar to the performance of the greedy version of take-the-best. Consequently, the performance of the greedy version of take-the-best provides a good proxy for the behavior of a number of alternative models of inductive inference.

Two artificially constructed environments will be used to compare the strategies. Both environments, therefore, will be governed by a known underlying functional relationship between the cues and criterion. Knowing these functional relationships will allow us to perform a bias–variance decomposition of the prediction error of the two strategies. The first environment is an instance of the class of *binary environments*, where the validity of the cues follows a noncompensatory pattern, and all cues are uncorrelated. An environment has a noncompensatory pattern when the validity of the cues decays rapidly as a function of their rank in the cue order. Noncompensatory environments are one example of a class of environments for which we have analytic results showing that take-the-best is unbiased and likely to perform well (Katsikopoulos and Martignon, 2006; Martignon and Hoffrage, 2002). The second environment used in our comparison, however, is an instance of the class of *Guttman environments*, inspired by the Guttman scale (Guttman, 1944), in which all the cues are maximally correlated with criterion, and all have a maximum validity of 1. Formal definitions and illustrations of both these environments are provided in Appendix 1 of Gigerenzer and Brighton (2009).

Figure 3.5(a–d) plots, for both of these environments, the prediction error achieved by take-the-best and its greedy counterpart. The performance of each model is shown separately in order to clearly distinguish the bias and variance components of error, which, when added together, comprise the total prediction error. Three findings are



# PROOF

84 *Henry Brighton and Gerd Gigerenzer*

revealed. First, in the binary environment, take-the-best performs worse than its greedy counterpart. This result illustrates that analytic results detailing when take-the-best is unbiased will not necessarily help to explain when take-the-best performs well. Second, in the Guttman environment, take-the-best outperforms its greedy counterpart. This result illustrates that proving that another strategy achieves a better fit than take-the-best is something quite different from proving that the strategy also achieves a higher predictive accuracy. Third, and perhaps most importantly, Figure 3.5 reveals that both of these behaviors are driven by the variance component of error, and the relative ability of the two strategies to keep variance within acceptable limits. Bias plays almost no role in explaining the difference in performance between the models, and the less-is-more effect we demonstrated in Figure 3.2 can also be explained by the relative ability of the models to control variance. In short, this comparison tells us that take-the-best bets on the fact that ignoring dependencies between cues is likely to result in low variance. Model comparisons in natural environments show that this bet is often a good one. But, as this comparison has revealed, the bet can also fail, even when take-the-best has zero bias.

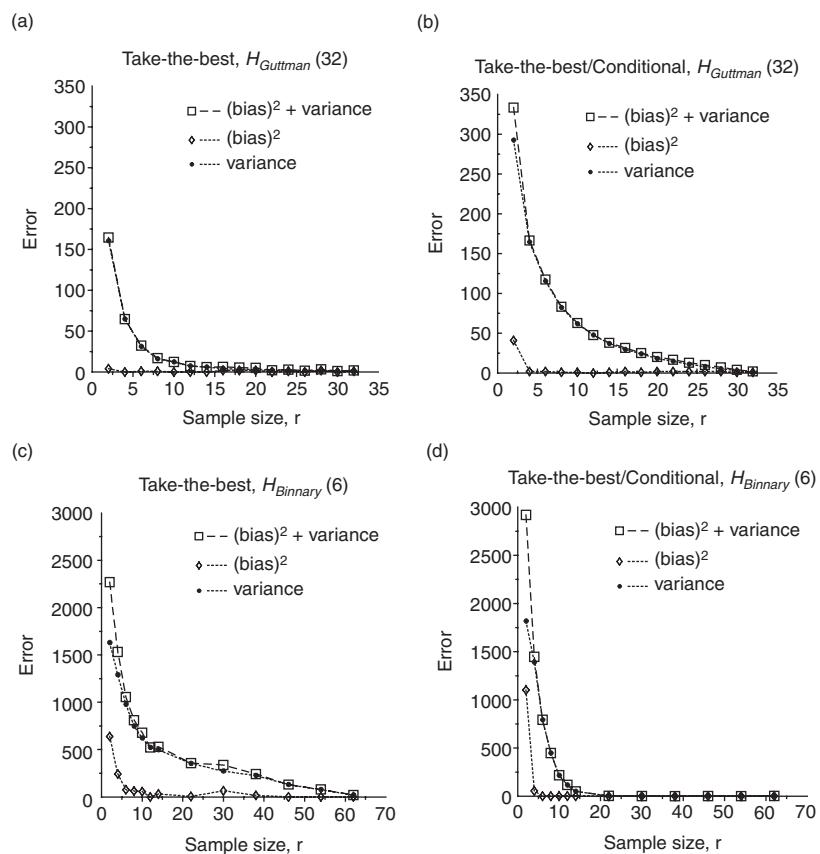
At this point, it is important to note that the concepts of bias and variance have allowed us to move beyond simply labeling the behavior of an induction algorithm as “overfitting the data,” or “suffering from excess complexity,” because the relative ability of two algorithms to avoid these pathologies will always depend on the amount of data available. First of all, from the perspective of bias, take-the-best offers no advantage over the alternative methods we have considered, because practically all models of inductive inference are capable of capturing the same systematic patterns in data as take-the-best. Consequently, if a heuristic like take-the-best is to outperform an alternative method, it must do so by incurring less variance. Second, the variance component of error is always an interaction between characteristics of the inference strategy, the structure of the environment, and the number of observations available. Thus, saying that a heuristic works because it avoids overfitting the data is only a shorthand explanation for what is often a more complex interaction between the heuristic, the environment, and the sample size.

### Bias, variance, and cognition

Organisms experience a limited number of observations, and cannot be expected to know, or possess the ability to model, underlying environmental regularities without error. The former constraint tells us that an organism’s inference mechanisms must control the variance

AQ2

# PROOF



*Figure 3.5* An illustration of the role played by variance in the performance of take-the-best. Plots (a) and (b) illustrate that, in Guttman environments, take-the-best outperforms the greedy variant of this heuristic, which orders cues by conditional validity. The performance difference is due to variance alone. Plots (c) and (d) illustrate that variance also explains why take-the-best is outperformed in binary environments. In both cases, take-the-best is unbiased and the relative performance of the models is explained almost entirely by variance.

component of error in order to make accurate inferences, while the latter constraint tells us that bias is inevitable. Taken together, considerations of bias and variance suggest that the apparent ability of humans and other animals to generalize accurately from limited observations of an uncertain environment is likely to hinge on the use of biased inference mechanisms which excel at limiting variance. Our analysis of



86 *Henry Brighton and Gerd Gigerenzer*

take-the-best has shown that heuristics work for precisely this reason. The other models considered opt to perform some form of optimization over a rich collection of potential hypotheses. Yet take-the-best can make more accurate inferences by being specialized (and therefore biased), and limiting variance by ignoring information.

This observation is the basis for our assertion that “biased minds make better inferences” (Gigerenzer and Brighton, 2009, p. 107). More generally, several established statistical models use misspecified models in order to reduce variance. Ridge regression is one well-known example, and works by limiting, or “squashing,” the fitted parameters’ values in a linear model (Hoerl and Kennard, 2000). By studying less-is-more effects, we are essentially exploring the question of how ignoring information and limiting the use of computation resources can introduce bias, but offset this bias with a greater reduction in variance. Heuristics are one way of exploring this question. The study of ecological rationality asks in which environments these heuristic tricks work, and in which will they fail. The notion of *Homo heuristicus* asserts that the cognitive system relies on such tricks in order to make accurate inferences in uncertain environments, while at the same time using limited processing resources.

### ***Homo Heuristicus* relies on an adaptive toolbox**

Let us return to our original proposition, that *Homo heuristicus* accurately captures fundamental aspects of human decision-making. We set the scene for this claim using examples of less-is-more effects, and explained them using the bias-variance dilemma. Throughout this discussion we have contrasted heuristics with optimization processes, and argued that, performing by ignoring information, limiting search, and relying on potentially impoverished representations of the world, an organism can make accurate inferences from limited observations. In our research group, this functional approach has been developed in parallel to a significant body of empirical work investigating human use of heuristics. Although this discussion has centered on functional concerns, both approaches are essential to developing a more detailed understanding of what we term the *adaptive toolbox*, which is a metaphor used to conceptualize the stock processing strategies available to the organism.

Table 3.1 lists 10 heuristics we view as being in the adaptive toolbox of humans, along with some of the surprising findings they have led to. Some of these heuristics address the same task, but perform well in different kinds of environment. For example, tallying and take-the-best



# PROOF

Homo heuristicus and the Bias – Variance Dilemma 87

*Table 3.1* Ten well-studied heuristics for which there is evidence that they are in the adaptive toolbox of humans. Each heuristic can be used to solve problems in social and nonsocial environments. See the references given for more information regarding their ecological rationality, and the surprising predictions they entail.

Heuristic	Definition <sup>1</sup>	Ecologically rational if:	Surprising findings (examples)
Recognition heuristic (Goldstein and Gigerenzer, 2002)  AQ3	If one of two alternatives is recognized, infer that it has the higher value on the criterion.	Recognition validity >0.5	Less-is-more effect if systematic forgetting can be beneficial (Schooler and Hertwig, 2005).
Fluency heuristic (Jacoby and Dallas, 1981)	If both alternatives are recognized but one is recognized faster, infer that it has the higher value on the criterion.	Fluency validity >0.5	Less-is-more effect; systematic forgetting can be beneficial (Schooler and Hertwig, 2005)
Take-the-best (Gigerenzer and Goldstein, 1996)	To infer which of two alternatives has the higher value: (1) search through cues in order of validity, (2) stop search as soon as a cue discriminates, and (3) choose the alternative this cue favors.	Cue validities are heavily skewed (Katsikopoulos and Martignon, 2006; Martignon and Hoffrage, 2002).	Often predicts more accurately than multiple regression (Czerlinski et al., 1999); neural networks, exemplar models, and decision tree algorithms (Brighton, 2006).
Tallying (unit-weight linear model, Dawes, 1979)  AQ4	To estimate a criterion, do not estimate weights but simply count the number of positive cues.	Cue validities vary little, low redundancy (Hogarth and Karelaia, 2005; 2006).	Often predict with equal or greater accuracy than multiple regression (Czerlinski et al., 1999).
Satisficing (Simon, 1955; Todd and Miller, 1999)	Search through alternatives and choose the first one that exceeds your aspiration level.	Number of alternatives decreases rapidly over time, such as in seasonal mating pools (Dudey and Todd, 2002).	Aspiration levels can lead to significantly better choices than chance, even if they are arbitrary (e.g., the secretary problem, see Gilbert and Mosteller, 1966; the envelope problem, see Bruss, 2000).

Continued

# PROOF

88 *Henry Brighton and Gerd Gigerenzer*

Table 3.1 Continued

Heuristic	Definition <sup>1</sup>	Ecologically rational if:	Surprising findings (examples)
1/ $N$ ; equality heuristic (DeMiguel et al., 2009)	Allocate resources equally to each of $N$ alternatives.	High unpredictability, small learning sample, large $N$ .	Can outperform optimal asset allocation portfolios.
Default heuristic (Johnson and Goldstein, 2003)	If there is a default, do nothing.	Values of those who set defaults match those of the decision-maker, when the consequences of a choice are hard to foresee.	Explains why mass mailing has little effect on organ donor registration; predicts behavior when trait and preference theories fail.
Tit-for-tat (Axelrod, 1984)	Cooperate first and then imitate your partner's last behavior.	The other players also play tit-for-tat; the rules of the game allow defection or cooperation but not divorce.	Can lead to a higher payoff than optimization (backward induction).
Imitate the majority (Boyd and Richerson, 2005)	Consider the majority of people in your peer group and imitate their behavior.	Environment is stable or only changes slowly; info search is costly or time-consuming.	A driving force in bonding, group identification, and moral behavior.
Imitate the successful (Boyd and Richerson, 2005)	Consider the most successful person and imitate his or her behavior.	Individual learning is slow; information search is costly or time-consuming.	A driving force in cultural evolution.

<sup>1</sup>For formal definitions, see references.

both describe how decision-makers make paired comparisons, but their predictive relative accuracy will depend on the statistical properties of the environment. Other heuristics address different tasks, but share common design features. For example, an organism relying on the default heuristic makes no decision if the environment already supplies a default option. An organism relying on the recognition heuristic opts for a recognized alternative when the other is unrecognized. Both heuristics implement one-reason decision-making, which means that they focus on one source of information, rather than integrating several potentially relevant additional sources of information, when making a decision. Notice, also, that none of the heuristics in Table 3.1



implement the process of optimization. For a more thorough discussion of the heuristics in Table 3.1, we refer the reader to Gigerenzer and Brighton (2009).

Several fundamental questions remain. For example, how do organisms select between strategies in the adaptive toolbox, and how can we understand, more generally, the functional relationship between cognitive mechanisms and the structure of the environment? These two questions are closely related, and center on the basic question of understanding the adaptive relationship between an organism and its environment. Such issues are clearly not specific to the study of heuristics, but the study of heuristics has proven a productive way of exploring them. Faced with the problem of designing an organism capable of functioning in a certain environment, it would be functional to equip the organism with a rich and accurate model of its environment, such that the consequences of its actions could be predicted. Such certainty allows the organism “to look before you leap.” In highly uncertain environments, one must face the unavoidable conclusion that accurate models of the world are beyond reach, observations will be limited, and error is inevitable. The study of heuristics is the study of simple strategies which respond to this problem, and the notion of *Homo heuristicus* proposes that these heuristics play a fundamental role in how the cognitive systems of humans and other animals respond so successfully to environmental uncertainty.



## References

- Axelrod, R. (1984). *The evolution of cooperation*. New York: Basic Books.
- Boyd, R., and Richerson, P. J. (2005). *The origin and evolution of cultures*. New York: Oxford University Press.
- Breiman, L., Friedman, J. H., Olshen, R. A., and Stone, P. J. (1994). *Classification and regression trees*. Belmont, CA: Wadsworth International Group.
- Brighton, H. (2006). Robust inference with simple cognitive models. In C. Lebriere, and R. Wray (Eds), *AAAI Spring Symposium: Cognitive Science Principles Meet AI-Hard Problems*. Menlo Park, CA: American Association for Artificial Intelligence, pp. 17–22.
- Bruss, F. T. (2000). Der Ungewissheit ein Schnippchen schlagen. *Spektrum der Wissenschaft*, 6, 106.
- Cover, T., and Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13, 21–7.
- Czerlinski, J., Gigerenzer, G., and Goldstein, D. G. (1999). How good are simple heuristics? In G. Gigerenzer, P. M. Todd, and the ABC Research Group, *Simple heuristics that make us smart*. New York: Oxford University Press, pp. 97–118.
- Dawes, R. M. (1979). The robust beauty of improper linear models in decision making. *American Psychologist*, 34, 571–82.



# PROOF

90 Henry Brighton and Gerd Gigerenzer

- Dawes, R. M., and Corrigan, B. (1974). Linear models in decision making. *Psychological Bulletin, 81*, 95–106.
- DeMiguel, V., Garlappi, L., and Uppal, R. (2009). Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *Review of Financial Studies, 22*, 1, 915–53.
- Dudey, T., and Todd, P. M. (2002). Making good decisions with minimal information: Simultaneous and sequential choice. *Journal of Bioeconomics, 3*, 195–215.
- Einhorn, H. J., and Hogarth, R. M. (1975). Unit weighting schemes for decision making. *Organizational Behavior and Human Performance, 13*, 171–92.
- Geman, S., Bienenstock, E., and Doursat, R. (1992). Neural networks and the bias/variance dilemma. *Neural Computation, 4*, 1–58.
- Gigerenzer, G., and Brighton, H. (2009). *Homo heuristicus*: Why biased minds make better inferences. *Topics in Cognitive Science, 1*, 107–43.
- Gigerenzer, G., and Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review, 103*, 650–69.
- Gigerenzer, G., Todd, P. M., and the ABC Research Group (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.
- Gilbert, J. P., and Mosteller, F. (1966). Recognizing the maximum of a sequence. *American Statistical Association Journal, 61*, 35–73.
- Goldstein, D. G., and Gigerenzer, G. (2002). Models of ecological rationality: The recognition heuristic. *Psychological Review, 109*, 75–90.
- Guttman, L. (1944). A basis for scaling qualitative data. *American Sociological Review, 9*, 139–50.
- Hastie, T., Tibshirani, R., and Friedman, J. (2001). *The elements of statistical learning: Data mining, inference, and prediction*. New York: Springer.
- Hoerl, A. E., and Kennard, R. W. (2000). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics, 42*, 80–6.
- Hogarth, R. M. (in press). When simple is hard to accept. In P. M. Todd, G. Gigerenzer, and the ABC Research Group, *Ecological rationality: Intelligence in the world*. Oxford, UK: Oxford University Press.
- Hogarth, R. M., and Karelaia, N. (2005). Ignoring information in binary choice with continuous variables: When is less “more”? *Journal of Mathematical Psychology, 49*, 115–24.
- AQ5 Hogarth, R. M., and Karelaia, N. (2006). “Take-the-best” and other simple strategies: Why and when the work “well” with binary cues. *Theory and Decision, 61*, 205–49.
- Hutchinson, J. M. C., and Gigerenzer, G. (2005). Simple heuristics and rules of thumb: Where psychologists and behavioural biologists might meet. *Behavioural Processes, 69*, 97–124.
- Katsikopoulos, K. V., and Martignon, L. (2006). Naive heuristics for paired comparisons: Some results on their relative accuracy. *Journal of Mathematical Psychology, 50*, 488–94.
- Jacoby, L. L., and Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General, 110*, 306–40.
- Johnson, E. J., and Goldstein, D. G. (2003). Do defaults save lives? *Science, 302*, 1,338–9.
- Martignon, L., and Hoffrage, U. (2002). Fast, frugal, and fit: Simple heuristics for paired comparisons. *Theory and Decision, 52*, 29–71.

# PROOF

Homo heuristicus and the Bias – Variance Dilemma 91

- Mugford, S. T., Mallon, E. B., and Franks, N. R. (2001). The accuracy of Buffon's needle: A rule of thumb used by ants to estimate area. *Behavioral Ecology*, 12, 655–8.
- Petrie, M., and Halliday, T. (1994). Experimental and natural changes in the peacock's (*Pavo cristatus*) train can affect mating success. *Behavioral Ecology and Sociobiology*, 35, 213–7.
- Pitt, M. A., Myung, I. J., and Zhang, S. (2002). Toward a method of selecting among computational models of cognition. *Psychological Review*, 109, 472–91.
- Quinlan, J. R. (1993). *C4.5: Programs for machine learning*. San Mateo, CA: Morgan Kaufmann.
- Roberts, S., and Pashler, H. (2000). How persuasive is a good fit? A comment on theory testing. *Psychological Review*, 107, 358–67.
- Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning internal representations by error propagation. In D. E. Rumelhart, and J. L. McClelland (Eds), *Parallel distributed processing: Explorations in the microstructure of cognition (Vol. 1: Foundations)*. Cambridge, MA: MIT Press, pp. 318–62.
- Schmidt, F. L. (1971). The relative efficiency of regression and simple unit weighting predictor weights in applied differential psychology. *Educational and Psychological Measurement*, 31, 699–714.
- Schmitt, M., and Martignon, L. (2006). On the complexity of learning lexicographic strategies. *Journal of Machine Learning Research*, 7, 55–83.
- Schooler, L. J., and Hertwig, R. (2005). How forgetting aids heuristic inference. *Psychological Review*, 112, 610–28.
- Simon, H. A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics*, 69, 99–118.
- Simon, H. A. (1991). *Models of my life*. New York: Basic Books.
- Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions. *Journal of the Royal Statistical Society B*, 36, 111–47.
- Todd, P. M., and Miller, G. F. (1999). From pride and prejudice to persuasion: Realistic heuristics for mate search. In G. Gigerenzer, P. M. Todd, and the ABC Research Group, *Simple heuristics that make us smart*. New York: Oxford University Press, pp. 287–308.
- Tversky, A., and Kahneman, D. (1974). Judgement under uncertainty: Heuristics and biases. *Science*, 185, 1,124–31.
- Weisberg, S. (1985). *Applied linear regression*. New York: Wiley.



## 4

### Action, Embodied Meaning, and Thought

*Mark Johnson*

#### Introduction

Human perception, experience, consciousness, feeling, meaning, thought, and action all require a functioning human brain operating in and through a live body that is in ongoing engagement with environments that are at once physical, interpersonal, and cultural. This embodied perspective demands an explanation of how all of the wondrous aspects of human mind – from our ability to have unified, intelligible experience all the way up to our most stunning achievements of theoretical understanding, imaginative thought, and artistic creativity – can emerge from our bodily capacities. I want to examine how the intricate intertwining of perception and action might provide the basis for our so-called “higher” acts of cognition and communication. In other words, I will explore how important parts of our abstract conceptualization and reasoning appropriate structures and processes of our most basic sensory-motor operations.

The proper starting point for an account of embodied meaning and thought is the acknowledgment that perception emerges in the context of action, that is, in the interaction of an organism with its environment. In an important early essay, “The Reflex Arc Concept in Psychology” (1896), John Dewey argued that perception – his example was visual perception – occurs as part of an action and not merely as a passive receptive event that later gives rise to responsive action:

Upon analysis, we find that we begin not with a sensory stimulus, but with a sensorimotor co-ordination, the optical-ocular, and that in a certain sense it is the movement which is primary, and the sensation which is secondary, the movement of body, head and eye muscles

# PROOF



*Action, Embodied Meaning, and Thought* 93

determining the quality of what is experienced. In other words, the real beginning is with the act of seeing; it is looking, and not a sensation of light. The sensory quale gives the value of the act, just as the movement furnishes its mechanism and control, but both sensation and movement lie inside, not outside, the act. (Dewey, 1896, p. 359)

I suppose that what in Dewey's day was a novel and disruptive claim is today a commonplace in perceptual psychology and cognitive neuroscience; namely, that we are active beings exploring our environment and that the character and directedness of our explorations creates a circuit of action–perception coordination. Perception and action are not two independent functions, but rather are aspects of a series of ongoing organism–environment interactions. As Alva Noe argues,

According to this sensorimotor, enactive, or actionist approach, seeing is not something that happens in us. It is not something that happens to us or in our brains. It is something we do. It is an activity of exploring the world making use of our practical familiarity with the ways in which our own movement drives and modulates our sensory encounter with the world. Seeing is a kind of skillful activity.

(Noe, 2009, p. 60)

This ongoing circuit of our skillful coping activity operates through energy transfer and transformation between organism and environment. Vittorio Gallese describes this organic transformation process as follows:

If we analyse at the *physical level of description* the relationship between biological agents and ‘the world outside’, we will find living organisms processing the different epiphanies of energy they are exposed to: electromagnetic, mechanical, chemical energy. Energy interacts with living organisms. It is only by virtue of this interaction that energy can be specified in terms of the ‘stimuli’ (visual, auditory, somatosensory, etc.) to which every organism is exposed. The result of the interaction between energy and living organisms is that the energy, now ‘stimulus’, is translated, or better, transduced into a *common information code*. The receptors of the different sensory modalities are the agents of the transduction process: they convert the different types of energies resulting from organisms-world interactions into the common code of action potentials. Action potentials express the electro-chemical excitability of cells, and constitute

# PROOF

94 *Mark Johnson*

the code used by the billions of neurons that comprise the central nervous system to ‘communicate’ with each other. (Gallese, 2003, p. 1,232)

Here we have the elementary basis for the organic transformations we know as cognition:

1. Perception arises in the context of an organism’s action, that is, its directed engagement with aspects of its environment.
2. This engagement is an interaction of energies (of the organism and its environment).
3. For the organism, the energy it receives is a “stimulus” that gets transduced into action potentials for the firing of nerve cells.
4. For the most part, there is no language-like structure involved here, but only the “code” of action potentials by which systems of neuronal assemblies “communicate” with one another.

However obvious this embodied cognition orientation might perhaps be for most cognitive neuroscientists, its implications for contemporary philosophy of mind and language strike me as far-reaching and profound, because (1) it ties perception inextricably to directed action within an organism–environment process, (2) it grounds meaning in sensory–motor processes (or so I shall argue), and (3) it challenges the Representational Theory of Mind. By a “representational theory” I mean the idea that human thought consists of a series of functional operations on language-like symbols “in the mind” that can be used to represent external states (Fodor, 1975; 1983). As Fodor puts it: “Mental representations (including, paradigmatically, thinking) are computations, that is, they are operations defined on the syntax of mental representations, and they are reliably truth preserving in indefinitely many cases” (Fodor, 2000, p. 19). I will argue that the Representational Theory of Mind is not just unnecessary to account for human meaning and thought; rather, it is also false and misleading, insofar as it claims that “the mind” thinks in a language of thought with quasi-linguistic symbols, as Fodor (1975) claims.

It is *not* my primary purpose in this essay to mount a direct sustained critique of the Representational Theory of Mind; rather, I am more concerned with sketching the outline of an alternative view of meaning and concepts. I would only note that Lakoff and I (Lakoff and Johnson, 1999) have analyzed the Language As Thought metaphor to show its fateful influence on early cognitive science and much of contemporary



philosophy of mind and language. We then show how it leads us to overlook all of the embodied processes that go into human meaning-making and thought, which is not intrinsically language-like, even though some of our highest achievements of thought often employ language. (For extended critiques of the representational view, see, for example, Barsalou, 1999; Churchland, 1986; 2002; Johnson, 2007.)

AQ1

The chief alternative to various discredited representational theories of mind would be some form of embodied view of meaning and concepts. The version of embodied cognition I want to explore here is what Gallese and Lakoff (2005) have called an *interactionist, multimodal, simulation theory of meaning and thought*. The interactionist perspective focuses on “the information processing carried out by the brain of an organism in the larger frame of the interactions between the organism and the environment it is acting upon.... The brain, a brain wired to a body that constantly interacts with the world is, at the same time, the vehicle of information *and* part of its content, the latter being conceived as a way to model organism-environment interactions” (Gallese, 2003, p. 1,233). A theory of this sort has a teleological dimension insofar as it recognizes that the locus of human meaning and thought is the directed activity of an embodied creature whose surroundings supply various “affordances” (Gibson, 1979) for present and future engagement with its world: “The energetic signals resulting from the organism-environment interactions are transduced and processed in the way they are, in respect of their content, because of the *relevance* (see Sperber, 2000) of this content for the possibility of establishing appropriate links between animal behavior and environment” (Gallese, 2003, p. 1,233). On this view, then, meaning and cognition are situated, embodied, and shaped by values that emerge for an organism with a certain type of body orienting and moving itself within specific types of environments.



AQ2

### Multimodal neuronal clusters

From this interactional/enactive perspective (Varela et al., 1999), I want to focus primarily on the controversial claim that sensory-motor systems are multimodal, instead of being only a set of modular units connected by supramodal links. To explain and illustrate the multimodal character of sensory-motor systems, Gallese and Lakoff (2005) have used recent research on the perception, performance, and understanding of acts of grasping. The most detailed studies have so far been done with monkeys, but Gallese and Lakoff argue that there are strong parallels between sensory-motor processing in monkeys and humans,



# PROOF

96 Mark Johnson

enough to support a strong analogy between the monkey and human brain regions under examination. More recent studies with humans are beginning to confirm the presence of similar structures and processes in humans.

To explain the notion of multimodality, consider what goes on in an act of grasping some small object in peripersonal space, that is, in the area of space that can be reached by your body parts, such as your head, hands, arms, and feet. In any act of grasping, perception and motor activity have to be intricately coordinated within a physical setting. A number of motor synergies, such as the ability to straighten each finger on the grasping hand and then to bend each finger around the object grasped, have to be sequenced into a coordinated, smooth motor action of closing the hand in just the right way and with just the right amount of force applied to some object located in peripersonal space. Grasping a baseball bat with the whole hand and palm is distinctively different from picking up a coin with your thumb and forefinger. The different types of grasping actions can thus be defined by identifying sets of action *parameters* and then giving specific values to those parameters. Typical parameters would include neural clusters such as the action performed (grasping small object), its direction (movement toward object), its amplitude (the “size” of the motion), its force (strength of motion), and so on. A specific type of grasping action – say grasping an egg versus grasping a softball – gets defined by specifying the values of the relevant parameters of the action, such as which parts of the hand need to be activated, in which direction the hand must move, and how much force is required in the gripping action. Grasping, of course, is not just an action of the hand, so further parameterizations are required for prior positioning of the trunk of your body, proper bending and twisting of the arm, and the eye–hand coordination that guides the hand to the location of the ball. The entire story of this sequencing of motor synergies is a small wonder that involves a grand orchestration of bodily perception and action. Each of the motor synergies must happen in just the right manner, at just the right moment of time. Too much gripping force and you crush the egg, and too little force and you drop it. Although these parameterizations of neuronal clusters operate automatically, mostly beneath the level of consciousness, we are sometimes consciously aware of specific parameters, such as when we feel the force with which we are gripping a ball or are proprioceptively aware how wide we have opened our hand in reaching for the ball.

To say that an action, such as grasping an egg, is *multimodal* is to say that “(1) it is neurally enacted using neural substrates used for both action and perception, and (2) that the modalities of action and perception are

# PROOF



*Action, Embodied Meaning, and Thought* 97

integrated at the level of the sensory–motor system itself and not via higher association areas” (Gallese and Lakoff, 2005, p. 459).

Multimodality is contrasted with the idea of “supramodality,” according to which an action such as grasping would be seen as requiring an association area in the brain that is independent of any sensory or motor areas but which is capable of integrating outputs from both the perceptual and motor systems. This is not to deny the existence of some supramodal areas (Damasio, 2010), but only to argue that much of our sensory–motor activity is done via multimodal structures.

Philosophically, there is a great deal at stake in the debate between supramodal and multimodal accounts. The supramodal view is compatible with a disembodied view of concepts that claims widespread modularity of systems and posits conceptual structure computed in brain areas other than those responsible for sensory–motor processing. The *multimodal view*, by contrast, argues that our conceptual knowledge is embodied through and through, because the multimodal sensory–motor system actually supplies the semantic content and inferential structure of our concepts. *In other words, the very same sensory–motor structures that make possible our perceptions and actions are the basis for our capacity to conceptualize those processes.* Multimodal structures are an example of what Goldman and de Vignemont (2009) call “bodily formats,” in which “mental representations in various bodily formats or codes have an important causal role in cognition” (p. 156); however, Goldman and de Vignemont do not embrace the strong Gallese–Lakoff claim that multimodal neural clusters are the basis for all our perceptual concepts.

In this essay, I am not denying the existence of supramodal systems. Indeed, I will identify some, and Damasio (2010) offers a theory of what he calls “convergence zones.” However, I am claiming that the existence of multimodal systems is sufficient to undermine strong versions of the Representational Theory of Mind.

Support for the multimodality hypothesis comes from studies of perception and action in monkeys and humans. It has been shown that sensory–motor processing is carried out via certain *functional clusters* of neurons, each of which constitutes a specific cortical network organized so as to carry out a specific function for the organism. Gallese and Lakoff describe three major types of functional clusters in three parietal premotor cortical networks in monkeys:

1. *Spatial Position Locators* – The F4-VIP neural cluster makes it possible for the monkey to be consciously aware of, and to interact purposefully with, objects in peripersonal space. Neurons in this cluster are



# PROOF

98 Mark Johnson

activated, for instance, when the monkey is aware of the spatial location of an object that it is about to reach for. Consequently, the visual presentation of an object within reach, or of a sound associated with that object in the same spatial location, appears to trigger an *action simulation* for interacting with that object. At the human level we have hundreds or thousands of these action simulation plans that permit us to go about our daily business of living. Moreover, these plans can be activated even when we don't actually perform an action, in which case we are said to be *simulating* that specific action, such as imagining ourselves reaching out toward a ball 30 degrees to the right of center and within our grasping range.

2. *Canonical Neurons* – The F5ab-AIP cluster plays a key role in actual and simulated action. When you see an object before you that you want to pick up and ingest, you automatically and nonconsciously prepare yourself for the requisite motor program for engaging that object. Gallese and his colleagues at Parma found that area F5 contains what they call “canonical neurons” (Rizzolatti and Fadiga, 1998) that fire not just when a monkey grasps an object, but also when the monkey merely sees an object that it might grasp in a certain way. The neural “representation” is not merely a visual experience, but rather a visuomotor activation: “Because most grasping actions are executed under visual guidance, a relationship has to be established between the features of 3D visual objects and the specific motor specifications they might engender if the animal is aiming at them. The appearance of a graspable object in the visual space will retrieve immediately the appropriate ‘motor representation’ of the intended type of hand-object relation” (Gallese, 2003, p. 1,236). Whereas the functional cluster described in (1) above focuses on the *location* of an object, canonical neuron clusters focus on the *shape* of an object as that relates to the capacity to grasp and manipulate that particular object.
3. *Mirror Neurons* – The third and most widely celebrated type of evidence for multimodality comes from the so-called “mirror neurons.” These are neurons in the F5c-PF cluster that fire both in the performance of a purposeful, goal-directed activity, such as grasping or manipulating an object, and also when the monkey merely sees another individual perform the same type of motor action. Mirror neurons are highly specific. They fire only relative to purposeful, goal-directed activity. They fire when a specific grasping action is either performed or observed, but they do not fire when merely observing the use of a tool, instead of the hand, to perform the action. The excitement over mirror neurons and their importance for neural simulation has engendered numerous studies of their operation in humans, whereby



it has been shown that analogous clusters operate when humans perform specific actions, see them performed by others, imagine performing them, and dream about performing them.

For our purposes, these three types of functional clusters are important because (1) they support the multimodality hypothesis, and (2) they give evidence for a simulation theory of meaning and conceptual understanding. The first type of neuronal cluster ties perception of spatial location to possible action plans. The second, canonical neurons, are clearly multimodal, because they fire both when the individual interacts with an object (e.g., grasps it) and when he or she sees the object that could be interacted with (e.g., grasped). The third type, mirror neurons, reveals perceptual and motor multimodality as well as the existence of simulation. You see someone reach for and grasp a banana and you automatically understand that action as it is simulated via mirror neurons, as though you were reaching for and grasping that banana. All three of these systems make possible the simulation of a purposive action without actually carrying that action out.

The moral here is an important one: "Within the operation logic of such neural networks, a series of physical entities, 3D objects, are identified, differentiated and conceptualized, not in relation to their mere physical appearance, but in relation to the effect of the interaction with an acting agent" (Gallese, 2003, p. 1,236).



## Embodied concepts

The most promising and potentially revolutionary aspect of the *interactionist, multimodal, simulation view* is the way it provides an account of concepts as truly embodied. According to the traditional Representationalist View, all concepts are essentially disembodied, in the sense that concepts are alleged to be modality-neutral representations that are processed in different brain areas than those responsible for perception and bodily movement. Even though our present neuroscience cannot definitively disprove the disembodied view of concepts, it is at least possible to suggest the general outline of an alternative embodied, multimodal, neural theory of concepts. The Gallese/Lakoff thesis is a bold one, namely, that "*the job done by what have been called 'concepts' can be accomplished by schemas characterized by parameters and their values.*" Such a schema, from a neural perspective, consists of a network of functional clusters" (Gallese and Lakoff, 2005, p. 466).

AQ3

What is most radical about this embodied view of concepts is that it eliminates the need to postulate a vast set of neuronal clusters that are



# PROOF

100 *Mark Johnson*

alleged to be above, beyond, and utterly independent of sensory-motor networks, and that are supposed to do the work of conceptualization and reasoning. Instead, our concrete concepts are computed via sensory-motor functional clusters. If it is true, this model is elegant and parsimonious from an evolutionary perspective, because it piggy-backs conception onto sensory-motor processes.

## An embodied concept – the *Grasp* schema

My discussion so far has been highly theoretical and somewhat abstract. It is useful to give a concrete example of what it might mean, from an interactionist, multimodal, simulation perspective, to say that we have a concrete concept such as *grasp*. What is given below is a cognitive model for *grasp* that uses the notion of parameters and their values to specify the dimensions of a progressive action that moves purposively from an initial condition, through a series of intermediate stages, to a final state. This *grasp* schema is symbolically described in terms of names for functional neuronal clusters. Though the schema is represented conceptually and propositionally, the schema itself is a complex cluster of neuronal networks, and there is nothing symbolic, in the traditional sense, about it. Crudely put, this schema presents our concept of *grasp*, and the key idea is that this concept is executed neurally using our multimodal sensory-motor systems. This example, adapted from Gallese and Lakoff (2005), is meant to give the general idea of how one might begin to go about explaining concrete concepts of bodily perception and action.

### The *grasp* schema

1. The role parameters: agent, object, object location, and the action itself.
2. The phase parameters: initial condition, starting phase, central phase, purpose condition, ending phase, final state.
3. The manner parameter.
4. The parameter values (and constraints on them).

The various parameters can be described as follows.

**Agent:** An individual.

**Object:** A physical entity with parameters: size, shape, mass, degree of fragility, and so on.

**Initial condition:**<sup>1</sup> Object Location: Within peripersonal space.

# PROOF



Action, Embodied Meaning, and Thought 101

**Starting phase:** Reaching, with direction: Toward object location; opening effector.

**Central phase:** Closing effector, with force: A function of fragility and mass.

**Purpose Condition:** Effector encloses object, with manner (a grip determined by parameter values and situational conditions).

**Final state:** Agent in-control-of-object. (Gallese and Lakoff, 2005, p. 467)

Although this description of the *grasp* schema may seem simplistic and overly abstract, each part of it is meant to be neurally realizable, both via computational neural models and by actual human neural systems. In the Central Phase and Purpose Conditions, for instance, closing an effector in a certain way and with a certain force is meant to be accomplished by coordinated and sequenced motor synergies.

Another important fact about this concepts-as-schemas notion is that schemas are not typically defined by necessary and sufficient conditions in the way assumed by classical theories of concepts. Functional clusters can be activated to various degrees, rather than in an all-or-nothing fashion. They can be activated with various dimensions or phases left out. And, as we will see shortly, they can be extended metaphorically in the constitution of abstract concepts.

What I am suggesting about the analysis of the concept *grasp* is that it can be generalized as a way of explaining *all* our concrete concepts of physical objects and physical actions. I have presented some evidence of the multimodal nature of our sensory-motor system, which consists of connected functional neuronal clusters. I then used the *grasp* example to show how a sensory-motor *system* for grasping can actually characterize a sensory-motor *concept* of grasping. Understanding a sensory-motor concept is a process of sensory-motor *simulation*. The elegance and parsimony of this account consist in the way it eliminates the need to posit two entirely independent systems, one for sensory-motor processing and a different one for forming sensory-motor concepts. The work of conceptualization is carried out via a simulation within neuronal clusters for perception and bodily movement.



## On the way to abstract concepts

The embodied, multimodal, simulation story might seem to make sense for concrete concepts, that is, for concepts of discrete physical objects and events, sense perceptions, and bodily movements. But what about



# PROOF

102 *Mark Johnson*

our vast panoply of abstract concepts? How are we to explain them as embodied? Obviously, this is going to be a complex process, but I can at least sketch three of the components that are necessary for such an account: image schemas, COGS, and conceptual metaphors.

### *Image schemas*

My key hypothesis is that abstract conceptualization and reasoning operate via the recruitment of the semantics and inferential structures of our bodily experience. Although I have not argued this point so far, it can be shown that our sensory-motor schemas have their own internal logics. There are simple corporeal and spatial logics of various bodily actions. Consider, for example, the image schematic structure of the motion of some object along a path from a starting point to a destination. The structure of this motion is what is known as the SOURCE-PATH-GOAL image schema (Johnson, 1987; Lakoff, 1987; Lakoff and Johnson, 1999), and it includes at least the following:

- A starting point (source)
- An ending point (destination or goal)
- An object that moves (the trajectory)
- A path connecting the starting point to the destination
- The speed with which the object moves
- Possible obstacles or hindrances to motion
- The location of the moving object at a given time
- The manner of motion of the object (e.g., smooth, jerky, bouncy)

Simply by virtue of the fact that we have bodies with certain specific perceptual and motor capacities, along with the fact that we interact routinely with certain types of structured environments, each of us will automatically acquire the ability to experience and enact scores of such image-schematic patterns, such as VERTICALITY, BALANCE, ITERATION, RIGHT-LEFT, CENTER-PERIPHERY, COMPULSIVE FORCE, CONTAINMENT, DEGREE OF INTENSITY (SCALARITY), and on and on (Cienki, 1997; Gibbs, 2006; Hampe, 2005; Johnson, 1987).

Each specific image schema has its own internal bodily or spatial logic, which we learn automatically from interacting with objects and moving our bodies in space. Within a situation structured by the SOURCE-PATH-GOAL schema, for instance, we are able to make inferences of the following sort, based on the logic of the schema:

- Two objects starting from a source point and moving at the same speed along the same path will reach the goal simultaneously.

# PROOF



*Action, Embodied Meaning, and Thought* 103

- If two objects start out from the same source point and move along the same path toward a destination, and if one of them moves faster than the other, then the faster object will arrive at the destination before the slower one.
- If an object has moved along a path to the halfway point toward the goal, then it has covered the intermediate points along the path up to that point.

Such logical relations are so simple as perhaps to seem trivial, but they are logical relations nonetheless, and we reason (make inferences) with them about our perceptual and motor experience all the time.

The structure of typical image schemas and their internal logics has been studied for cultures and languages around the world (Hampe, 2005), and I cannot pursue this further here. However, it should be noted that the neural basis of image schemas is beginning to be explored, and structured connectionist neural models have been constructed that can process certain image-schematic structure (Dodge and Lakoff, 2005; Regier, 1996).

## Cogs



Srini Narayanan (1997) has developed structured connectionist neural models of motor actions and our concepts of those actions. He has modeled premotor systems, motor systems, and premotor-motor connections in which the premotor system coordinates and sequences the specific motor synergies. Narayanan's models recognize the various possible stages or phases of motor actions: (1) initial state, (2) starting phase transition, (3) precentral state, (4) central phase transition (instantaneous, prolonged, or ongoing), (5) postcentral state, (6) ending phase transition, and (7) final state. His model also has structures for assessing progress to goal, reiterating an action sequence, deciding to terminate the action, and so forth. The model can sequence actions, run them in parallel, or embed one action within another. Narayanan gave the name "executing schemas" (X-schemas) to these premotor structures of actions.



One of the more important features of Narayanan's model is that it is capable of representing what linguists call the "aspect" of actions – the manner in which they are done (e.g., one time only, repeatedly, prolonged over a span of time, etc.). All languages the world over must have ways of representing the necessary aspectual dimensions of actions. Although in retrospect this may seem obvious, Narayanan saw that, since his models of sensory-motor actions could represent all of the key features of aspect, it followed that his model might be able to represent



# PROOF

104 *Mark Johnson*

the structure of *any* action, whether a spatial motion, a general change of state, or an abstract mental process. In other words, once again, the sensory–motor system manifests all the pattern and inferential structure necessary for any kind of action, even a “non-physical” or “mental” action.

The premotor cortex is a *secondary (supramodal) area* of sorts, since it structures information going to both sensors and effectors, even though it does not have direct connections to sensory–motor areas. Think of *action structure* as an abstract concept, which Narayanan, Feldman (2006), Lakoff and others have called a *cog*. *Cog* structures can be characterized as follows (Gallese and Lakoff, 2005):

- They are simulated in secondary neural areas that are not directly connected to primary sensory or motor areas.
- If the connections to primary areas are inhibited, then the simulation can be run “abstractly,” without any specific motor action resulting.
- The simulation involves inferences that are computed based on the structure of the *cog*.
- Such simulations can characterize concepts in the grammar of a natural language.
- These concepts apply generally, and not just to sensory–motor events, so they can be used for abstract conceptualization and reasoning.
- *Cogs* will typically have image-schematic structure as part of their logic.

The *cogs* hypothesis is important insofar as it gives us a way of accounting for certain abstract structures that are necessary in natural languages around the world. The neural modeling is only beginning, but there is promising work on what Damasio (1999) has called “convergence zones” that appear to be likely candidates for the neural computation of these structures.

### *Conceptual metaphor*

A third important piece of the puzzle about abstract concepts is the notion of conceptual metaphor. Metaphor draws on, and blends with, all of the elements of embodied meaning we have surveyed so far: multimodal sensory–motor concepts, image schemas, executing schemas for action structures, and other cogs for the structure of grammatical constructions. A conceptual metaphor is a cross-domain conceptual mapping. The mapping is asymmetrical and directional, typically (but

# PROOF



not always) from a sensory-motor source domain onto a nonphysical or abstract domain. For example, consider the basic metaphor PURPOSES ARE DESTINATIONS. The source domain is motion in physical space and the target domain is any form of purposeful action. The mapping is as follows:

### The Purposes Are Destinations Metaphor

Source Domain		Target Domain
[Motion in Space]	>>>>	[Purposeful Action]
Starting Location S	>>>>	Initial State (Having a Purpose)
Ending Location E	>>>>	Final State (Purpose Achieved)
Destination	>>>>	Purpose
Movement from S to E	>>>>	Progress Toward Achievement of Purpose
Distance Moved at Time T	>>>>	Stage of Progress Toward Purpose
Obstacles to Motion	>>>>	Impediments to Action
Lack of Direction	>>>>	Lack of Purpose

Although for notational purposes we give the metaphor a propositional name (in this case, PURPOSES ARE DESTINATIONS), the metaphor is the underlying conceptual mapping, and that mapping is realized neurally as a complex neural binding, typically from sensory-motor areas of the brain to areas responsible for understanding and reasoning about nonphysical or abstract objects or processes. We are lulled into believing the classical view that metaphor is linguistic – a mere matter of words – primarily because the conceptual mapping gives rise to polysemous linguistic expressions that have meanings pertaining both to physical, sensory-motor experiences and also to abstract notions. For instance, the polysemous term *arrived* in “She has finally arrived” could have both a spatial sense (as in “Sally finally arrived at her destination after a long bus trip”) or a metaphorical sense (as in “Sally finally arrived at her goal of gaining the respect of her coworkers”). Thus, the PURPOSES ARE DESTINATIONS metaphor shows up in English in expressions such as: “She started out to get her doctoral degree,” “She wandered off the track along the way,” “He stood in the way of her progress,” “Her friends helped her get moving again along the right path,” “She was able to refocus on where she was going and how to get there,” “She finally reached her original destination – she got her PhD.” Notice that the conceptual metaphor is revealed not just in language, but in all types of symbolic expressions. For example, we often represent progress to some abstract goal visually as a point moving along a line toward a destination. We might diagram progress as an arrow point that gets progressively closer and closer to

106 *Mark Johnson*

some destination–goal. Or, you might hold up your thumb and forefinger a quarter of an inch or so apart to indicate that you were *soooooo* close to winning the championship. David McNeill (1992; 2005) has pioneered the empirical study of metaphor in spontaneous gestures, and his methods have been used for cross-cultural studies (e.g., Cienki, 1998; Kovecses, 2010).

Another important aspect of conceptual metaphor is its role in reasoning and inference. We utilize our knowledge of the source domain, which is typically structured via image schemas, to reason about the target. As we just noted above, the closer you are spatially to your destination, the sooner you will arrive there. Applied metaphorically, via the conceptual mapping above, to progress toward any abstract goal, the “closer” you are to it, the sooner you will achieve it. We know that, if someone blocks our path to our destination, it will frustrate our journey. So, if the “North Koreans *put a roadblock* in the peace talk process,” they are frustrating efforts to achieve nuclear nonproliferation. This reasoning may seem simple, almost trivial, but it is not, because it shows how the image-schematic logic of the source domain can be utilized in drawing target domain inferences, via the mapping.

Joseph Grady (1997) has studied the experiential origins of basic conceptual metaphors in our mundane bodily experience, pursuing the hypothesis that these metaphors are learned simply because we have bodies of a certain makeup that routinely interact with environments in structured ways. Grady incorporated Christopher Johnson’s research on the acquisition of metaphorical competence in children, which revealed three stages of development that lead to the understanding and use of metaphors. In the first stage, a child uses a perceptual term like *see* only in its “literal” visual sense. In other words, that term was used only in the source domain (here, the domain of visual experience). In the second stage (the *conflation stage*), the term is used in cases where both the source and target domains are active together, such as using *see* when there is visual experience (source domain) that is also correlated with knowing something (target domain), as in “See, mommy’s home.” In this case the child uses a source domain term with a meaning that is also appropriate to the target domain semantics and grammar. In the third stage (the *differentiation stage*), the term is used metaphorically, such as when there is no actual visual experience (e.g., “I *see* your point”).

Building on Johnson’s theory of the development of metaphoric competence, Grady hypothesized that an ordinary person interacting with his or her environment would acquire scores of metaphors



# PROOF



Action, Embodied Meaning, and Thought 107

simply through experiential correlations, that is, cases in which there are two conceptual domains coactivated. In neural terms, there are two functional clusters firing together, and, via the Hebbian neurons-that-fire-together-wire-together rule, these clusters are neurally coactivated and longer-term connections (potentiations) are established between them. Grady called these bindings across different domains (or functional neuronal clusters) *primary metaphors*. In the case of the PURPOSES ARE DESTINATIONS primary metaphor, every child and adult has many experiences each day in which some object's motion along a path is correlated with the progressive achievement of a purpose. For instance, the baby sees its coveted pacifier on the floor under the table. Its crawling motion toward the pacifier is correlated with progress toward the baby's satisfaction of its intended goal of getting that pacifier in its mouth and sucking. Later, this experiential correlation (realized as a neural coactivation with neural simulation of progress toward goal) is reactivated in the use of the language of motion in a context of purposive action.

Grady thus hypothesizes that in this mundane way, without conscious reflection, and based principally on bodily engagement with our environment, we learn large numbers of experientially grounded primary metaphors, each with its own inference patterns. The TEMPORAL CHANGE IS PHYSICAL MOTION metaphor is grounded in our experience of seeing the movement of an object through space as correlated with our sense of the passing of time, giving rise to expressions such as "The hours *flew by*," "The time *is approaching* when there will be no clean water," and "The time *has arrived* to take action." The STATES ARE LOCATIONS metaphor is based on the experienced correlation of being at some location and being in a certain condition tied to that location, as in "She's really *in trouble now*," "The water *went from hot to cold*," and "We're *out of danger now*."

The large-scale systematic mappings that [Lakoff and I (1980; 1999) have called conceptual metaphors are defined by sets of submappings, each of which either is a primary metaphor or else has primary metaphors as part of its submapping. For example, the vast systematic LOCATION EVENT-STRUCTURE metaphor has as submappings such primary metaphors as STATES ARE LOCATIONS, CHANGE OF STATE IS MOTION, CAUSATION IS FORCED MOTION, CAUSES ARE FORCES, HINDRANCES ARE OBSTACLES TO MOTION, ACTIONS ARE SELF-PROPELLED MOTIONS, PURPOSES ARE DESTINATIONS, and several others. There is large and growing literature analyzing the workings of conventional conceptual metaphors in language groups and cultures around the world (see, for example, any

AQ4



108 *Mark Johnson*

issue of the journal *Metaphor and Symbol*). There is also an emerging field of attention on what Charles Forceville has called “multimodal” metaphors, in which the source and target domains are in two different perceptual or experiential modes (Forceville and Urios-Aparisi, 2009). For example, Forceville analyzes a commercial in which a French green bean and a mini-corn cob “stand” together in front of a frozen vegetable pouch while we hear the strains of Mendelsohn’s Wedding March playing in the background. Such cases provide excellent evidence for the reality of metaphor as conceptual – as constitutive of our thinking – and not just matters of words.

With the topic of perception and action as our central focus, I want to conclude with some examples of cases in which our conception of mind and thought is metaphorically understood in terms of processes of perception and bodily movement or manipulation of objects. Lakoff and I (1999) analyzed five of the major metaphors for thinking that appear in English. In each case the source domains are drawn, as we would predict, from types of bodily action: perceiving (especially seeing), object manipulation, moving one’s body through space, eating, and reproducing. Here are partial mappings for four of the metaphors we found, followed by the kinds of expressions that arise from the metaphors. What we found so striking was the vast polysemy evidence for the existence of these metaphors and the manner in which the details of the mapping generate a distinctive logic for thinking about the nature and operations of thinking itself. Here are the five major metaphors, parts of their constitutive mappings, and examples of the polysemy of the terms we employ in our understanding of acts of thinking.

### The Thinking Is Moving Metaphor

The Mind Is A Body  
Thinking Is Moving  
Ideas Are Locations  
Reason Is A Force  
Rational Thought Is Motion (that is direct, deliberate, step-by-step, and in accord with the force of reason)  
Being Unable To Think Is Being Unable To Move  
A Line of Thought Is A Path  
Thinking About X Is Moving In The Area Around X  
Communicating Is Guiding (along a path)  
Understanding Is Following  
Rethinking Is Going Over The Path Again

# PROOF



Action, Embodied Meaning, and Thought 109

**Examples:** My mind *strayed from* the topic. Sarah's mind *wandered* all over the place. He's always *going off on flights of fancy*. Show me how you *reached that conclusion*. Take me *step by step* through your argument. I don't see *how to move from* that assumption *to* this conclusion. I'm *stuck!* I can't *go any further in this line of reasoning*. Don't *skip any steps* in your proof. Mike's *going in circles* and never *gets to* his point. *Slow down!* You're *going too fast* for me. I can't *keep up with you*. Could you *run over* (*go over*) *that again?* *Where are you going* with this?

## The Thinking Is Perceiving Metaphor (Vision Version)

The Mind Is A Body  
Thinking Is Perceiving (Seeing)  
Ideas Are Things Perceived (Seen)  
Knowing Is Seeing  
Communication Is Showing  
Becoming Aware Is Noticing  
An Aid To Knowing Is A Light Source  
Capacity To Know Is Being Able To See  
Impediments To Knowledge Are Impediments To Vision  
Knowing From A "Perspective" Is Seeing From A Point of View  
Explaining In Detail Is Drawing A Picture  
Directing Attention Is Pointing  
Paying Attention Is Looking At



**Examples:** I *see* what you're saying. I *see* your point. The politician attempted to *cover up* the facts. He *pulled the wool over their eyes*. He *put up a smokescreen* with his arguments. Your explanation is *unclear/murky/opaque*. Jayne was *looking for* a solution, but she only *discovered* more problems. We were left *in the dark*, hunting around *blindly*. That was the most *enlightening* account of relativity theory I've ever heard. New facts have *come to light*. You'd *see* the truth if you didn't *have blinders on* – it's as *clear/plain* as the nose on your face. He *pointed out* the best solution. Do I have to *draw you a picture*? We couldn't *see* the forest for the trees.

Note: there are obviously other systematic metaphors based on the other sensory modes of perception (touching, hearing, tasting) that would be subcases of the generic THINKING IS PERCEIVING metaphor.

## The Thinking Is Object Manipulation Metaphor

The Mind Is A Body  
Ideas Are Objects  
Thinking Is Object Manipulation



# PROOF

110 *Mark Johnson*

Communication Is Sending  
Understanding Is Grasping  
Memory Is A Storehouse  
Remembering Is Retrieval (Recall)  
The Structure Of An Idea Is The Structure Of An Object  
Analyzing Ideas Is Taking Apart Objects

**Examples:** It's a hard idea *to grasp*. Let's *play with* that idea awhile – *toss it around* a bit. I've got the argument *firmly in mind*. I *get it*. She *gave* me the best idea for a paper topic. The negotiators *exchanged ideas* throughout the night. His view *came across* to me. Our teacher's always trying to *cram* our heads *full of ideas*. He's *putting* dangerous ideas *into* their young minds. His theories *sail way over my head*. That's a *slippery* concept. The speaker was *throwing* too many ideas *at* me all at once. What Sally revealed *threw me a curve*. Let's *take apart* his *theory* and *break it into* its key components. Complex ideas have to be carefully *crafted, shaped, and reshaped*. Jack *turned the idea over* in his mind, examining every aspect. He *held up* my ideas to scrutiny, putting every one of them *under the microscope*.

## The Thinking Is Preparing And Eating Food Metaphor

A Well-functioning Mind Is A Healthy Body  
Ideas Are Food  
Acquiring Ideas Is Eating  
Interest In Ideas Is Appetite For Food  
Good Ideas Are Healthy Foods  
Helpful Ideas Are Nutritious Foods  
Bad Ideas Are Harmful Foods  
Disturbing Ideas Are Disgusting Food  
Interesting Ideas Are Appetizing Food  
Uninteresting Ideas Are Flavorless Foods  
Testing An Idea Is Smelling Or Tasting  
Considering Is Chewing  
Accepting Is Swallowing  
Fully Comprehending Is Digesting  
Preparing Ideas To Be Understood Is Food Preparation  
Communication Is Feeding  
Substantial Ideas Are Meat

AQ5

**Examples:** Derek has a real *thirst* for knowledge, a *huge appetite* for learning, and an *insatiable* curiosity. You don't expect me to *swallow*

# PROOF



Action, Embodied Meaning, and Thought 111

that *garbage*, do you? I'll have to *chew on that* for a bit. The whole idea *smells fishy*. He's known for putting out a bunch of *raw facts, half-baked ideas*, and *warmed-over theories*. Let's *put that on the back burner* for a while and *let it simmer*. What've you *cooked up* for me now? *What a rotten idea – disgusting, unsavory*, and enough to *make you puke*, if you have any intelligence at all. That's pure *bullshit*. You're not *shittin' me*, are you? We're gonna have to *sugar-coat it*, or even *force it down her throat*, to *make it palatable* to her. Do you ever feel like you have to *spoon-feed* your students? His scheme *left a bad taste in my mouth*. There's *too much here for me to digest*. Where's *the beef* in your theory? The really *meaty* issue is sustainability. Let's just *chew the fat*. Finally, something you can really *sink your teeth into!* Now that's *food for thought*! Our philosophy teacher just wants us to *regurgitate* what she gives us in lecture – just *spit it back out*.

There are other major metaphors for thinking, such as the planting, tending, and harvesting of plants, or the insemination, gestation, and giving birth to ideas. All of them are grounded on bodily source domains of human action that supply the semantics and logical inferences enacted in the target domain. It is no accident that the source domains are bodily, physical acts of perception and movement. We recruit the meaning and structure of our bodily experience for the purposes of abstract conceptualization and reasoning. George Lakoff and I have proposed that virtually all of our abstract concepts are defined by systematic conceptual metaphors composed from primary metaphors. Examine any abstract concept from any field of human activity and you do not have to look far to find metaphors – built up from primary metaphors – defining the way we think and reason with that concept. We appropriate the logic of our source domain reasoning for our reasoning in the target domain. You may have some vague and highly abstract literal sense of the meaning of a concept, but the details of the semantics and the specific inferences are generated by the metaphors that characterize the concept, not by some abstract literal core of meaning. The metaphorical constitution of our key concepts has been studied in nearly every field and discipline you can imagine, such as causal theories (Lakoff and Johnson, 1999), metaphysics (Johnson, 2008), logic and mathematics (Lakoff and Nunez, 2000), theory of knowledge (Lakoff and Johnson, 1999), morality (Johnson 1993; Lakoff and Johnson, 1999), law (Bjerre, 2005; Winter, 2001), science (Magnani and Nersessian, 2002), psychology (Gibbs, 2006; Johnson and Fernandez-Duque; Kovecses, 2000), music (Spitzer, 2004; Zbikowski, 2002), advertising (Forceville, 1996), and on and on.

AQ6



# PROOF

112 *Mark Johnson*

The metaphorical constitution of our most important abstract concepts – the ones that lie at the heart of our most significant ways of understanding and making sense of our world – is not a shortcoming to be bemoaned. Rather, it is a fact about how creatures like us are able to use the resources of our bodily encounters with our world in order to make sense of things and to gain insight.

If the experiential grounding of conceptual metaphor is a correct hypothesis, then one would expect that semantic priming of a specific source domain would facilitate the processing of a metaphor with that source domain, insofar as the priming would activate a simulation of source domain structure and inferences. There is now ample evidence of such priming effects (see Gibbs, 1994; 2006). For example, Boroditsky and Ramscar (2002) studied the priming effects of two different conceptual metaphors for temporal change. According to the MOVING TIMES metaphor, times are objects moving with various speeds toward and then past a stationary observer. Thus, we say things like “The day is fast approaching when we will leave for China,” “The time for action has arrived,” “That sordid event is past us now,” “Tuesday follows Monday, but it comes before Wednesday.” The second spatial motion metaphor for temporal change is the MOVING OBSERVER metaphor, in which times are locations on a landscape and the observer moves toward and beyond various time-locations. This second metaphor gives rise to expressions like “We’re fast approaching Thanksgiving,” “We’re coming up on Christmas,” “It’s a long way to Memorial day,” “What’s up ahead for us,” and so on. Lakoff and I noticed that an expression like “Let’s move the meeting ahead two days” can have two different meanings, and can generate two different outcomes, depending on whether one is understanding it via the MOVING TIMES or the MOVING OBSERVER metaphor. If the original meeting was scheduled for Wednesday, then according to the MOVING TIMES metaphor the meeting is moved *ahead* (earlier) to Monday, while with the MOVING OBSERVER metaphor the meeting would be moved *ahead* (later) to Friday. Boroditsky hypothesized that she could predict which interpretation would be favored, based on which source domain structure (hence, which metaphor) was primed. So, if the subject saw a video of a person seated in an office chair pulling a second chair toward her with a rope, the subject was more likely to interpret (via the MOVING TIMES metaphor) “move the meeting *ahead* two days” to mean move it to Monday, while a subject who saw the person seated in the chair pull herself across the floor by a rope tied to a fixed object was more likely to select the move-to-Friday interpretation via priming

# PROOF



Action, Embodied Meaning, and Thought 113

of the MOVING OBSERVER metaphor. These and other similar experiments have given evidence of priming effects for conceptual metaphors, based on the experiential structure of the source domains for the metaphors.

## Conclusion – embodied, multimodal meaning

I have been pursuing the Deweyan hypothesis that all perception occurs in relation to ongoing action with and within an environment, and I have argued that large parts of human meaning and thought are based on this perception–action connection. The interanimation of perception and action is evidenced by the multimodal character of both our perceptual experience and our perceptual concepts. Following Gallese and Lakoff (2005), I argued that a good portion of our conceptual knowledge is embodied via sensory–motor structures. In other words, our sensory–motor system not only makes it possible for us to act in the world, but it also provides both the content and structure for concrete concepts, based on perceptual and motor simulations. The sensory–motor system of the brain appears to be multimodal, rather than strictly modular, and language exploits this multimodality. Image schemas, cogs, and conceptual metaphors are three of the chief dimensions of our conceptual system by which we can recruit aspects of sensory–motor processing for “higher level” cognition, thereby moving from concrete to abstract concepts. At these levels of semantic and inferential structure, supramodal neural systems can come into play.

If the general outlines of the picture of conceptualization that I have sketched are cognitively realistic, this would be an evolutionarily elegant and neurally parsimonious picture of the appropriation of bodily experience and meaning for the highest reaches of human thought. Dewey put this well in *Experience and Nature* when he wrote:

Since mind cannot evolve except where there is an organized process in which the fulfillments of the past are conserved and employed, it is not surprising that mind when it evolves should be mindful of the past and future, and that it should use the structures which are biological adaptations of organism and environment as its own and its only organs. In ultimate analysis the mystery that mind should use a body, or that a body should have a mind, is like the mystery that a man cultivating plants should use the soil; or that the soil which grows plants at all should grow those adapted to its own physico-chemical properties and relations. (Dewey, 1925, p. 277)

AQ7



# PROOF

114 *Mark Johnson*

And Dewey then concludes:

Since both the inanimate and the human environment are involved in the functions of life, it is inevitable, if these functions evolve to the point of thinking and if thinking is naturally serial with biological functions, that it will have as the material of thought, even of its erratic imaginings, the events and connections of this environment. (Dewey, 1925, p. 279)

Thinking is a form of human activity that changes the patterns of our ongoing, developing experience. Thinking is in and of the world. As a highly complex mode of action, it is grounded in our capacities for perceiving and acting, which are themselves intertwined.

## Note

1. The “::” notation indicates the content of a phase.

## References

- Barsalou, L. (1999). Perceptual Symbol Systems. *Behavioral and Brain Sciences*, 22, 577–660.
- Bjerre, C. (2005). Mental Capacity as Metaphor. *International Journal for the Semiotics of Law*, 18, 101–40.
- Boroditsky, L. and Ramscar, M. (2002). The Roles of Body and Mind in Abstract Thought. *Psychological Science*, 13, no. 2, 185–9.
- Churchland, P. (1986). *Neurophilosophy: Toward a Unified Science of the Mind/Brain*. Cambridge, MA: MIT Press.
- Churchland, P. (2002). *Brain-Wise: Studies in Neurophilosophy*. Cambridge, MA: MIT Press.
- Cienki, A. (1997). Some Properties and Groupings of Image Schemas. In M. Verspoor, K. D. Lee, and E. Sweetser (Eds), *Lexical and Syntactical Constructions and the Construction of Meaning*. Amsterdam: John Benjamins, pp. 3–15.
- Cienki, A. (1998). Metaphoric Gestures and Some of Their Relations to Verbal Metaphoric Expressions. In J. P. Koenig (Ed.), *Discourse and Cognition: Bridging the Gap*. Stanford: CSLI, pp. 189–204.
- Damasio, A. (1999). *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. New York: Harcourt, Brace and Company.
- Damasio, A. (2010). *Self Comes to Mind: Constructing the Conscious Brain*. New York: Pantheon Books.
- Dewey, J. (1896). The Reflex Arc Concept in Psychology. *Psychological Review*, 3, 357–70.
- Dewey, J. (1925/1981). *Experience and Nature*. Vol. 1 of *The Later Works, 1925–1953*. (J. A. Boydston, Ed.). Carbondale: Southern Illinois University Press.

# PROOF



Action, Embodied Meaning, and Thought 115

- Dodge, E. and Lakoff, G. (2005). Image Schemas: From Linguistic Analysis to Neural Grounding. In Beate Hampe (Ed.), *From Perception to Meaning: Image Schemas and Cognitive Linguistics*. Berlin: Mouton de Gruyter, pp. 57–91.
- Feldman, J. (2006). *From Molecule to Metaphor: A Neural Theory of Language*. Cambridge, MA: MIT Press.
- Fodor, J. (1975). *The Language of Thought*. New York: Thomas Y. Crowell Co.
- Fodor, J. (1983). *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. Cambridge, MA: MIT Press.
- Fodor, J. (2000). *The Mind Doesn't Work That Way: The Scope and Limits of Computational Psychology*. Cambridge, MA: MIT Press.
- Forceville, C. (1996). *Pictorial Metaphor in Advertising*. London: Routledge.
- Forceville, C. and Urios-Aparisi, E. (Eds) (2009). *Multimodal Metaphor*. Berlin: Mouton de Gruyter.
- Gallese, V. (2003). A Neuroscientific Grasp of Concepts: From Control to Representation. *Philosophical Transactions of the Royal Society of London*, 358, 1,231–40.
- Gallese, V. and Lakoff, G. (2005). The Brain's Concepts: The Role of the Sensory-motor System in Conceptual Knowledge. *Cognitive Neuropsychology*, 22, 455–79.
- Gibbs, R. (1994). *The Poetics of Mind: Figurative Thought, Language, and Understanding*. Cambridge: Cambridge University Press.
- Gibbs, R. (2006). *Embodiment and Cognitive Science*. Cambridge: Cambridge University Press.
- Gibson, J.J. (1979). *The Senses Considered as Perceptual Systems*. Boston: Houghton-Mifflin.
- Goldman, A. and de Vignemont, F. (2009). Is Social Cognition Embodied? *Trends in Cognitive Science* 13, No. 4, 154–9.
- Grady, J. (1997). *Foundations of Meaning: Primary Metaphors and Primary Scenes*. PhD dissertation, University of California, Berkeley.
- Hampe, B. (Ed.) (2005). *From Perception to Meaning: Image Schemas in Cognitive Linguistics*. Berlin: Mouton de Gruyter.
- Johnson, M. (1987). *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason*. Chicago: University of Chicago Press.
- Johnson, M. (1993). *Moral Imagination: Implications of Cognitive Science for Ethics*. Chicago: University of Chicago Press.
- Johnson, M. (2008). Philosophy's Debt to Metaphor. In R. Gibbs (Ed.), *The Cambridge Handbook of Metaphor and Thought*. Cambridge: Cambridge University Press, pp. 39–52.
- Kovecses, Z. (2000). *Metaphor and Emotion: Language, Culture, and Body in Human Feeling*. Cambridge: Cambridge University Press.
- Kovecses, Z. (2010). *Metaphor: A Practical Introduction*. Oxford: Oxford University Press.
- Lakoff, G. (1987). *Women, Fire, and Dangerous Things: What Categories Reveal About the Mind*. Chicago: University of Chicago.
- Lakoff, G. and Johnson, M. (1999). *Philosophy in the Flesh: Embodied Mind and Its Challenge to Western Thought*. New York: Basic Books.
- Lakoff, G. and Nunez, R. (2000). *Where Mathematics Comes From: How the Embodied Mind Brings Mathematics Into Being*. New York: Basic Books.



# PROOF

116 Mark Johnson

- AQ8 McDermott, J. (Ed.) (1973). *The Philosophy of John Dewey*. Chicago: University of Chicago Press.
- McNeill, D. (1992). *Hand and Mind: What Gestures Reveal about Thought*. Chicago: University of Chicago Press.
- McNeill, D. (2005). *Gesture and Thought*. Chicago: University of Chicago Press.
- Magnani, L. and Nersessian, N. (Eds) (2002). *Model-based Reasoning: Science, Technology, Values*. New York: Kluwer Academic/Plenum Publishers.
- Narayanan, S. (1997). *KARMA: Knowledge-based Active Representations for Metaphor and Aspect*. PhD Dissertation, Computer Science Division, University of California, Berkeley.
- Noe, A. (2009). *Out of Our Heads: Why You Are Not Your Brain, and Other Lessons from the Biology of Consciousness*. New York: Hill and Wang.
- Regier, T. (1996). *The Human Semantic Potential: Spatial Language and Constrained Connectionism*. Cambridge, MA: MIT Press.
- Rizzolatti, G. and Fadiga, L. (1998). Grasping Objects and Grasping Action Meanings: the Dual Role of Monkey Rstroventral Premotor Cortex (area F5). *Sensory Guidance of Movement. Novartis Foundation Symposium 218*. Chichester: Wiley, pp. 81–103.
- Sperber, D. (2000). Metarepresentations in an Evolutionary Perspective. In D. Sperber (Ed.), *Metarepresentations: a Multidisciplinary Perspective. Vancouver Studies in Cognitive Science, Vol. 10*. Oxford University Press, pp. 117–38.
- Spitzer, M. (2004). *Metaphor and Musical Thought*. Chicago: University of Chicago Press.
- Varela, F., Thompson, E. and Rosch, E. (1991). *The Embodied Mind*. Cambridge, MA: MIT Press.
- Winter, S. (2001). *A Clearing in the Forest: Law, Life, and Mind*. Chicago: University of Chicago Press.
- Zbikowski, L. (2002). *Conceptualizing Music: Cognitive Structure, Theory, and Analysis*. Oxford: Oxford University Press.



## 5

## Neo-Pragmatism and Enactive Intentionality

*Shaun Gallagher*

### Introduction

Enactive approaches in cognitive science propose that perception, and more generally cognitive experience, are strongly mediated by embodied (sensory–motor) processes, and that our primary experience of the world is action-oriented or pragmatic (Noë, 2004; Thompson, 2007; Varela et al., 1991). Extended mind theorists propose that cognition supervenes on embodied and environmental processes such as gestures and the use of various technologies (Clark, 2008; Clark and Chalmers, 1998; Menary, 2010). Both enactive and extended conceptions of cognition suggest that the mind is not “in the head”—that cognitive processes are distributed over brain, body, and environment—but they also differ on a number of issues. Extended mind theorists defend a functionalist account of cognition and downplay the role of the body (e.g., Clark, 2008), and they argue that cognition and action can involve mental representations (e.g., Clark, 1997; Clark and Grush, 1999; Rowlands, 2006; Wheeler, 2005). In contrast, enactive theorists argue for radical embodiment (e.g., Thompson and Varela, 2001) and defend an antirepresentationalist view (e.g., Gallagher, 2008b; Hutto, in press; Thompson, 2007). There are also debates about how to define the boundaries, or lack of boundaries, involved in cognitive processes (e.g., Di Paolo, 2009; Wheeler, 2008).

In this chapter we do not propose to resolve all of these issues. Rather, we intend to address the following question: can enactive and extended conceptions of the mind agree on a model of intentionality? We explore several conceptions of intentionality in order to ask which conception of intentionality best supports the new concept of mind implied by both the enactive and the extended views. We argue (1) that, although

118 *Shaun Gallagher*

both enactive and extended views champion a non-Cartesian, noninternalist conception of mind, we only start to see what this conception of mind is when we adopt an enactivist conception of intentionality; (2) only by adopting this model of intentionality will the proponents of the extended mind hypothesis be able to fend off those critics who insist on defining the “mark of the mental” in terms of nonderivative (narrow or internal) content (e.g., Adams and Aizawa, 2001; 2009); and (3) working out this model of intentionality requires resolutions to a number of debates in the area of social cognition (sometimes called “theory of mind”(ToM)).

### Intentionality as a character of mental states

The concept of intentionality is a complicated one, with a long history. Both the phenomenological and the analytic tradition tend to point to Brentano’s definition of intentionality as a starting point. Brentano, in turn, took his orientation from medieval sources.

Every mental phenomenon is characterized by what the Scholastics of the Middle Ages called the intentional (or mental) inexistence of an object, and what we might call, though not wholly unambiguously, reference to a content, direction toward an object (which is not to be understood here as meaning a thing), or immanent objectivity. Every mental phenomenon includes something as object within itself...(Brentano, 1862/2008, p. 88)

In this regard, Brentano understood intentionality to be the mark of the mental, and understood the mental in terms of a psychological act, its content, and the relation between them.

In some ways this view shares an internalist bias with what Haugeland (1990) calls a neo-Cartesian concept of intentionality. Such a view is defended by Horgan and Kriegel, and they refer to it as a “traditional, strongly internalist, broadly Cartesian picture of the mind” (2008, p. 353). On this view the mind is in some way discontinuous with everything around it, but, at the same time, naturalistically continuous with the brain. Horgan and Kriegel (2008) summarize this internalist view in six propositions:

1. The mind exhibits intentionality in virtue of its nature as phenomenal consciousness (intrinsic phenomenal intentionality).



2. Intentional content is narrow – i.e., internal –and could be instantiated in a brain in a vat, reflecting the fact that prototypical mental states have strict neural correlates.
3. Intentionality is subjective – access is given only to the experiencing subject.
4. Intentionality applies to both sensory and cognitive states.
5. Intentionality is nonderivative.
6. Intentionality is the mark of the mental.

This is clearly a view that locates intentionality in the head and limits it to traditionally defined internal mental states. This neo-Cartesian internalist framework identifies nonderived intentionality as the mark of the mental (Horgan and Kriegel, 2008; also Horgan and Tienson, 2002), an idea posed against the extended mind hypothesis by those who insist that the mind cannot extend outside of brain-bound processes. Thus, Adams and Aizawa attempt to answer the question of where cognition stops and something noncognitive begins by appealing to nonderived content as the mark of the mental (Adams and Aizawa, 2001; 2009).

The concept of nonderived content, however, is not on settled ground. In this regard, there are disagreements that would apply to both the Brentanian and the neo-Cartesian concept concerning what is or is not intentional. Some, including Husserl [e.g., 1913/1982, §36] and Searle (1992), have argued that not all mental experience is intentional. For example, some theorists claim that the experience of pain is a mental experience, but is not necessarily intentional (see, e.g., Crane, 1998). Moreover, as Shapiro (in press) notes, “there is today no received theory of how original content comes to be in the first place.”

Furthermore, as an internalist argument against the extended mind hypothesis, the question about what constitutes the mark of the mental, or what constitutes nonderived intentionality, is not a question that can be answered without begging the question of whether some cognitive processes are extended. That is, it seems possible that there could be a theory of nonderived intentionality consistent with the very different conception of the mind suggested by enactive and extended accounts of cognition.

### Intentionality as a characteristic of agents

The discussion of intentionality goes beyond a narrow discussion of mental state intentionality in both phenomenology and analytic

AQ2





120 *Shaun Gallagher*

philosophy of mind. In phenomenology, Husserl introduces the concept of operative (*fungierende*) intentionality in contrast to act intentionality (the mental state concept). The concept of operative intentionality attempts to capture the fact that the experiencing agent is intentionally engaged with the world through actions and projects that are not reducible to simple mental states, but involve what Husserl refers to as bodily intentionality (1977, p. 39).<sup>1</sup> Merleau-Ponty (1945/1962) takes up the analysis of intentionality just at this point. Actions are intentional, not only in the sense that they are willed, but also in the sense that they are directed at some goal or project. Moreover, this intentionality of action is something that can be perceptually understood by others. We return to this concept of operative intentionality later in the paper.

Likewise, in analytic philosophy of mind we find discussions of agent intentionality that go beyond the question of mental state intentionality. Haugeland (1990), for example, contrasts the neo-Cartesian concept of intentionality with neo-behaviorist and neo-pragmatist conceptions. Both the neo-behaviorist and the neo-pragmatist conceptions of intentionality share a common feature: an externalist view that intentionality is something that we can discern in behavior and is not necessarily hidden away inside the head. In this respect they seem to be good candidates for the kind of intentionality needed to support the enactive and extended concepts of mind. We argue below (in the next two sections, respectively) that, for different reasons, both the neo-behaviorist version and, on its own, the neo-pragmatist version fall short.

### Neo-behaviorism and theory of mind

The neo-behaviorist view is exemplified in Dennett's intentional stance, which he explains in terms of observing an agent engaged in rational behavior, and on that basis ascribing intentionality, that is, treating the agent as someone "who harbors beliefs and desires and other mental states that exhibit intentionality or 'aboutness', and whose actions can be explained (or predicted) on the basis of the content of these states" (1991, p. 76).<sup>1</sup> In this account we can immediately see two things: first, that this conception of an intentional agent (or system) starts with considerations about the agent's behavior, but also refers us back to questions about mental state intentionality; second, that this conception relates intentionality to social cognition, a relation we will highlight as our discussion proceeds. Phillip Pettit provides a similar explanation: intentional agents "are agents that engage with their environment in such a way that we ascribe beliefs and desires to them" (1996, p. 10).



# PROOF



*Neo-Pragmatism and Enactive Intentionality* 121

We ascribe beliefs and desires to them on the basis of observed regularities in their behavior, most commonly identified as “rational regularities” (p. 11). Furthermore, defenders of the extended mind hypothesis, including Clark and Chalmers (1998) and Clark (2008), seem to adopt this conception of intentionality.<sup>2</sup> For them, it is reasonable to take an agent’s mind as extended only insofar as we gain explanatory and predictive advantage by doing so, that is, by ascribing mental states that require external physical vehicles for their realization, in order to explain action.

AQ4

Let’s note first that, in contrast to claims made by the neo-behaviorist model, in our actual practice of intentional ascription, we do not always treat another agent’s meaningful action as a rational behavior, or as an instrumental action directed at a particular desired goal. If, for example, we see someone gesturing or nodding their head as they listen to a lecture, we do not fail to attribute a certain intentionality to them in this respect, even though we do not always understand the intentionality expressed in gesturing or head-nodding as motivated by specific beliefs (e.g., about the meaning of the gesturing or head-nodding) or desires (e.g., to impress the lecturer) (Miyahara, in press).

Another example involves the intentionality associated with sexuality. As Merleau-Ponty puts it:

◎

Erotic perception is not a *cogitatio* aiming at a *cogitatum*; through one body it aims at another body, and it is carried out in the world, not in a consciousness. A spectacle has sexual signification for me, not when I represent, even confusedly, its possible relation to the sexual organs or to states of pleasure, but when it exists for my body, for this power which is always ready to form the given stimuli into an erotic situation and to behave therein in a sexual way. (Merleau-Ponty, 1945/1962, p. 139).

◎

Erotic intentionality is not a matter of instrumental rationality, and not reducible to a set of mental states, propositional attitudes like beliefs or desires, or even to a set of observable behaviors, or to some attributional/inferential link between the two. It’s a form of intentionality that seemingly goes beyond the terms of folk psychology.

As the descriptions of the intentional stance suggest, neo-behaviorists rely on a common-sense or folk psychology involving mental states (themselves understood in terms of mental state intentionality) as the basis for ascribing intentionality to the agent. In addition, however, neo-behaviorism asserts that we take an agent as having intentionality



# PROOF

122 *Shaun Gallagher*

"only in relation to the strategies of someone who is trying to explain and predict its behavior" (Dennett, 1971, p. 87). Accordingly, an agent actually has intentionality only relative to its being ascribed by an external observer, and this fails to explain how an agent might have intentionality on its own (or ascribe intentionality to itself) without the presence of an observer (or without trying to explain and predict its own behavior). This, however, apparently goes against our common-sense psychology. For we do usually think we can have intentional states on our own. Moreover, this inability to explain how an agent might have intentionality on its own seemingly forces neo-behaviorists to reintroduce the concept of nonderived intentionality. This problem becomes salient if we ask how the external observer, in virtue of whom an agent has intentionality, acquires the intention to explain and predict the agent's behavior. On the neo-behaviorist logic, to answer this question, we must appeal to another observer, who intends to explain and predict the first observer's behaviors. The intentionality of the second observer, however, would require a third observer, and so on. As Adams and Aizawa (2001) point out, for the neo-behaviorist account to get off the ground, it would be more reasonable to suppose an internalist story of nonderived intentionality for the first observer, or the agent whose intentionality was originally in question.

More generally, with respect to ToM and social cognition, the neo-behaviorist model of intentionality goes hand in hand with a theory (TT) approach, that is, an approach that conceives of intersubjective understanding as a form of inference based on folk psychology. There are, however, a number of objections that can be raised against the TT view (see Gallagher, 2005; Goldman, 2006). If, for example, as on some views of TT, mind-reading is both a conscious process and pervasive in our everyday understanding of others, then we should be able to find phenomenological evidence for the inferential process that relies on folk psychology. As a brief reflection should tell us, however, when we are engaged with others in our everyday situations, we rarely find ourselves making such mind-reading inferences to mental states. Even if, alternatively, we think of mind-reading inferences as nonconscious, TT still has difficulty explaining how young infants are able to interact with others in ways that suggest that they have a practical understanding of emotions and intentions, since it is unlikely that they have the cognitive capacities for theoretical inference.<sup>3</sup>

Finally, it is noteworthy that this approach to understanding intentionality, and the TT approach to social cognition, still depend on a rather standard model of the mind as a set of mental states "in the head,"

# PROOF



*Neo-Pragmatism and Enactive Intentionality* 123

and hidden behind behavior. Such hidden mental processes are said to constitute the canonical type of explanation of the intentional behaviors that we observe. The trace of this idea is obvious in the extended mind theorists' appeal to the "Parity Principle" (Clark, 2008; Clark and Chalmers, 1998) or "functional isomorphism" (Clark, 1997) in support of the extended mind hypothesis. The principle is that we should take a nonneural item (a body part or something in the environment) as a part of the vehicle of cognition only to the extent that the functional contribution it makes to a cognitive activity is isomorphic or similar enough to a contribution that could be made by some processes in the brain. A notebook containing information which guides an agent's behavior, for example, constitutes a part of the agent's mind because the function it plays is on a par with the function biological memory might play (Clark and Chalmers, 1998). Contrary to their intention, the Parity Principle invited troubling arguments against the extended mind hypothesis. Adams and Aizawa (2001), for example, reject the hypothesis, while endorsing the Parity Principle, by arguing that, as a matter of empirical fact, external processes are functionally dissimilar to internal, neural processes (see also Rupert, 2004 for a similar argument). The extended mind theorists, however, are concerned to establish the similarity between the external (nonneural) and internal (neural) items only because they still hang on to the canonical conception of the mind as "in the head." Thus, whether or not the extended mind theorists committed to the neo-behaviorist conception of intentionality can provide viable responses to such criticisms, this is not the concept of intentionality that will work for enactive and extended mind theorists who want to push toward an alternative conception of mind.<sup>4</sup>

## Neo-pragmatism<sup>5</sup>

In contrast to neo-behaviorists, some neo-pragmatists, like Brandom (1994;2000), appeal to an account of intentionality that depends on social/normative concepts. Brandom explains the concept of intentionality in terms of what he calls the practice of *deontic scorekeeping*, that is, our mutual implicit practice of keeping track of each other's and our own actions in terms of *normative status* (1994, Ch. 3). On this view we understand the intentionality of the other implicitly in terms of certain commitments or entitlements specified by social norms, although we do not always *acknowledge* such normative statuses *explicitly*. If, for example, I promise to give you a wake-up call tomorrow morning at 7, neither you nor I necessarily think explicitly of this in terms of my loss



# PROOF

124 *Shaun Gallagher*

of entitlement to sleep until 9. However, we *keep score* or *keep track* of such things *implicitly*, and this is revealed in the way we act. Thus, if I oversleep until 9 and do not fulfill the promise, I will apologize and you will blame me because I *should not* have slept until 9. This interaction reveals that we were both tracking my normative status implicitly.

Accordingly, we ascribe intentionality to an entity who is capable of having a particular set of commitments and entitlements, namely, *inferentially articulated* or *discursive* commitments and entitlements instituted by social linguistic norms – that is, the implicit norms that determine the social appropriateness of our linguistic practices, including inferential reasonings. In this regard, Brandom thinks it's "norms all the way down" and that "only communities, not individuals, can be interpreted as having original intentionality. [... T]he practices that institute the sort of normative status characteristic of intentional states must be *social practices*" (1994, p. 61). That is, we track, and occasionally acknowledge, other people's intentionality in virtue of what they are doing and saying, what they are expected to do or say, what roles they play, what kind of place and time it is and what such factors mean to us in the shared social situation, rather than by somehow looking for mental states hidden behind their behaviors.<sup>6</sup> We ascribe intentionality to actions to the extent that we have a practical grasp on their socially instituted significance.

On this basis, we have no problem in ascribing intentionality to gestures and head nods. According to neo-pragmatism, gestures are just another kind of the doing by which, by virtue of their socially instituted significance, people keep track of one another's normative statuses. In understanding another's head-nodding, for example, we attribute to that person a discursive commitment to the claim that the lecturer made, which he or she may or may not explicitly acknowledge, instead of positing a belief or internal mental state about the content of the claim. Furthermore, since we can keep track of our *own* normative status by relying on the social significance of our own circumstance, typically including our own previous sayings and doings, there is no problem with the possibility of having intentionality on one's own, as a social agent.

Thus, the neo-pragmatist account of intentionality avoids some of the problems found in the neo-behaviorist account. Neo-pragmatists, however, run into a different problem, namely, in their attempt to account for our common-sense ability to recognize intentionality in the behavior of a variety of non- or presocial entities, such as geometrical figures moving in certain patterns on a computer screen, nonhuman



animals, and human infants.<sup>7</sup> According to neo-pragmatism, something is an intentional agent only if it acts according to norms that are socially based. Certain insulting gestures, for example, are culturally relative, and we should not understand someone from a different culture who accidentally made an insulting gesture to be acting as an insulting intentional agent. More generally, if a creature (e.g., a non-human animal) *completely* lacks understanding of social norms, and is not expected to act in accordance with such norms, it seems that the ascribing of intentionality itself would be inappropriate. And yet we do ascribe intentionality to animals, and others who lack understanding of social norms (e.g., presocial infants). Empirical studies show that we also tend to see intentionality even in geometric figures if they make particular kinds of movements (Heider and Simmel, 1944; Michotte, 1963). Neo-pragmatists, then, seemingly fail to explain our everyday practices of ascribing intentionality to such creatures.

One suggestion for resolving this problem points to a more basic issue. Cash (2008;2009) suggests that, on a neo-pragmatist account, we can ascribe intentionality to animals and infants “*based on the similarity of their movement* to the kind of actions, which if performed by a person would entitle us to ascribe such intentional states as reasons” (2008, p. 101; emphasis added). That is, neo-pragmatists can ascribe intentionality to a nonsocial entity, but only by recognizing some kind of similarity between that entity’s behavior and the behavior of a socialized human. What this proposal entails, however, is not clear. Moreover, this lack of clarity extends back to the processes that are involved in ascribing intentionality to humans as well. That is, neo-pragmatism doesn’t make it clear how things work even in the human case – how, precisely, we recognize agents to be acting in accordance with social norms.

There seem to be two possibilities in the case of ascribing intentionality to nonhuman entities, if, as Cash indicates, similarity is the operative concept. The first would be a form of pattern recognition plus inference from analogy. That is, we might take certain nonhumans in motion as having intentionality by detecting a common dynamic pattern between their movements and behaviors displayed by other people. This solution fails, however, based on the simple fact that, at least in some instances when we ascribe intentionality to animals or moving geometrical figures on a computer screen, there is no behavioral similarity to humans involved (Miyahara, in press). Alternatively, if we take ourselves as the person or model on which to base the comparison, we can understand the emphasis on similarity as a move in the direction of the simulation theory of social cognition (ST). Goldman, for



# PROOF

126 *Shaun Gallagher*

example, calls the following view, as described by Dennett, a version of ST: “the view that when we attribute beliefs and other intentional states to others, we do this by comparing them to ourselves, by projecting ourselves into their states of mind” (Dennett, 1987, pp. 98–9; see Goldman, 2006, p. 57).

ST contends that we rely, not on folk psychological inferences, but on modeling the other person’s mental states using our own minds to simulate what we would do if we were in their shoes. Mind-reading, on this view, involves the projection of our own first-person pretend beliefs and desires onto the minds of others based on a similarity that we see between their actions and our own.

Some of the same objections raised against explicit versions of TT can be raised against explicit versions of ST. That is, there is no phenomenological evidence that we follow simulation routines in our everyday social interactions; we rarely find ourselves mind-reading, either by inference or by simulation. Furthermore, it is unlikely that infants have the cognitive capacities for what Goldman (2006) describes as a high-level or introspective simulation, and yet they have some practical understanding of others.

One way to avoid these objections is to appeal to neural ST. The recent neuroscience of mirror neurons (MNs) has motivated a reconception of simulation (see Gallese and Goldman, 1998; Rizzolatti et al., 2001). According to this view, MNs are characterized as simulating the actions of others since they are specific neurons that are activated in two circumstances: either when I perform an action or when I see the other person perform an action. In this regard, the claim is that the observer’s motor system goes into a matching state with the observed action of others. Simulation, on this view, would be equivalent to my system *matching* or establishing a *similarity* with the system that I observe – sometimes referred to as the matching hypothesis (Goldman, 2006; Rizzolatti et al., 2001). Mind-reading, or at least action comprehension, would be the result of an automatic mechanism that works entirely on a subpersonal level. The fact that these processes are nonconscious (not explicit) suggests that phenomenology is not in a position to offer evidence for or against simulation. Furthermore, since these processes are automatic, and do not involve higher cognitive function, they can explain how infants can simulate the actions of others.

This move to neural ST looks promising for neo-pragmatism, since it focuses on action understanding and treats intention as something implicit in the action itself. Theoretically, one can certainly stop short of stronger claims about MNs being a basis for mind-reading, rather than just action understanding. If neural ST avoids some objections,



however, it fails to avoid others. With respect to the issue of intention attribution to nonhuman entities, MN activation is limited to very few cases of cross-species action perception – that is, MNs cannot explain why we would attribute intentions to dogs, cats, roaches, geometrical figures, and so on, since MNs are not activated in these cases, although they may explain intention attribution to monkeys and apes (see Buccino et al., 2004). Supposedly, the bodily action similarity is strong enough between humans and monkeys. Outside these limited cases, we run into an objection similar to the one about simple pattern recognition – insufficient behavioral (bodily action) similarity.

Other objections arise if we take neural ST as a model of human, intersubjective understanding (see Gallagher, 2007). For example, on the standard definition simulation involves pretense (one system operates “as if” it were in the situation of the other system). MNs, however, are said to be neutral with respect to who the agent is (Gallese, 2005). That is, they fire indifferently whether I am acting or I am observing you acting. As a result, in the activation of MNs there is no distinction between self and other, a distinction which is necessary, but not sufficient, for pretense. Even if there were some basis in the MN activation for the self–other distinction (e.g., difference in firing rates for self-action vs. observation), more than the self–other distinction, something corresponding to the “as if,” is needed for pretense. This issue has motivated the matching hypothesis, which involves a more minimal definition of simulation, that is, simulation as simple matching (Goldman, 2006; Rizzolatti et al., 2001). Neuroscientific research and several recent experiments, however, show no matching activation between action-execution and action-observation for specific MNs (or MN areas).

Dinstein et al. (2008), for example, using the paper-scissors-rock game, show that, in fact, in certain areas of the brain where MNs are thought to exist – specifically the *anterior intraparietal sulcus* (aIPS) – areas activated for producing a particular hand action are not activated for observing that hand action in another. For matching gestures, “distinctly different fMRI response patterns were generated by executed and observed movements in aIPS... aIPS exhibits movement-selective responses during both observation and execution of movement, but... the representations of observed and executed movements are fundamentally different from one another” (Dinstein et al., 2008). Another study (Catmur et al., 2007) shows that learning can work against matching. The experimenters trained subjects to move their fingers in a manner incongruent with an observed hand, for example, moving the little finger when they observed movement of the index finger. After training, magnetic evoked potentials (MEPs) were greater in the little finger when index finger



128 *Shaun Gallagher*

movement was observed. "The important implication of this result is that study participants who exhibited incongruent MEP responses presumably did not mistake the perception of index finger movement for little finger movement..." (Hickok, 2009, p. 1,236). That is, the lack of matching in the motor system does not preempt some kind of recognition of what the other person is doing. More generally, Csibra (2005) points out that, conservatively, between 21 and 45 per cent of neurons identified as mirror neurons are sensitive to multiple types of action; of those activated by a single type of observed action, that action is not necessarily the same action as defined by the motor properties of the neuron; approximately 60 per cent of mirror neurons are "broadly congruent," which means there may be some relation between the observed action(s) and their associated executed action, but not an exact match. Only about one-third of mirror neurons show a one-to-one congruence.<sup>8</sup> Newman-Norlund et al. (2007, p. 55) suggest that activation of the broadly congruent mirror neurons may represent a complementary action rather than a similar action. In that case they could not be simulations defined on the matching hypothesis (Gallagher, 2008a).

From the neo-pragmatist perspective, even if we could set these problems aside, the simulation theory of social cognition, like TT, remains too closely tied to internalist models of intentionality. For example, ST shares two important assumptions with TT. First, both approaches assume that the problem of social cognition is best posed as one that involves mind-reading – a capacity that allows us to address our lack of access to other minds, characterized as hidden mental states which explain their manifest behavior. If some proponents of neural ST sometimes move away from this mind-reading model to focus on action understanding (e.g., Gallese, 2009), many others take neural ST as a form of mind-reading (e.g., Oberman and Ramachandran, 2009) or a support for mind-reading (e.g., Keysers and Gazzola, 2009). This assumption goes against the neo-pragmatic claim that intentionality has a normative status specified by social norms, publicly accessible to other people as well as to the one to whom intentionality belongs. Furthermore, we have already seen that extended mind theorists are confronted by difficult criticisms because of their failure to remove the trace of this internalist model from their argument for the extended mind hypothesis.

Second, and closely aligned with the first assumption: the assumption of methodological individualism.

*Methodological individualism:* the working assumption that access to knowledge about the minds of others depends on cognitive



capabilities or mechanisms of an isolated individual, or on processes that take place inside an individual brain. (Froese and Gallagher, submitted)

In other words, for TT and ST, the solution has to come by some cognitive ("in the head") operation (theoretical inference, introspective modeling, or neuronal process), or theory-of-mind mechanism (ToMM or MNs), that would allow us to grasp the other person's mental states. Thus, both the attribution process and the intentionality attributed are narrowed down to the workings of internal mental states or brain states. According to neo-pragmatism, however, the cognitive understanding of others, or the explicit acknowledgement of intentionality, is secondary, possible only on the background of implicit shared practices; the basic mode of understanding others is not supported by the individual cognitive capacities for acknowledging others' intention as such, but rather by actual or potential interaction with others in a socially appropriate way. Thus, according to neo-pragmatism, some agents (or pre-agents) may be capable of tracking another's intentionality in practice, while being unable to understand it in isolation from the other in question, without engaging in an actual social interaction, which is a possibility that methodological individualism shuts out right away.

These models of intentionality attribution and social cognition thus do not provide a good fit with either the neo-pragmatist view or the enactive or extended theories of cognition. What we propose in the following sections is a neo-pragmatist account of intentionality without simulation; a neo-pragmatism consistent with enactive, extended, externalist conceptions of mind.

### Interaction theory and operative intentionality

It may be helpful to summarize what we have done so far. We're still searching for an account of intentionality consistent with the alternative concept of mind suggested by enactive and extended approaches. So far we have ruled out the Brentanian, the neo-Cartesian, and the neo-behaviorist views of intentionality for a variety of reasons. We have also ruled out analogical and simulationist versions of neo-pragmatism. We turn now to our main proposal, namely that the phenomenological conception of operative intentionality is consistent with neo-pragmatism, and, as we argue in this section, is supported by an alternative conception of social cognition known as interaction theory. Moreover, this conception of intentionality is enactive, and is precisely the view of





130 *Shaun Gallagher*

intentionality needed by extended mind theorists to counter the argument concerning nonderived content and the mark of the mental.

The interaction theory of social cognition (IT) cites evidence from developmental psychology, phenomenology, and dynamic systems modeling to show that interaction and social contexts are important constitutional factors in social cognition: that is, that processes and events external to the individual – specifically, interactions themselves – have a transformative effect on individuals who engage in them (Gallagher, 2001; 2005; Gallagher and Hutto, 2007; Gallagher and Zahavi, 2008; Ratcliffe, 2007). This involves the concept of ‘strong interaction’ (Gallagher, *in press*; Froese and Gallagher, *submitted*), which we can define, following De Jaegher et al. (2010):

*Strong Interaction:* a mutually engaged coregulated coupling between at least two autonomous agents where the coregulation and the coupling mutually affect each other, constituting a self-sustaining organization in the domain of relational dynamics.

It is beyond the scope of this chapter to present the full range of evidence for the notion of strong interaction. We can provide a brief summary of some of the evidence found in developmental studies, phenomenological and behavioral studies, and dynamic systems modeling.



### (a) Developmental studies

IT appeals to evidence from developmental studies, starting with primary and secondary intersubjectivity (Trevarthen, 1979; Trevarthen and Hubley, 1978). Primary intersubjectivity consists of the innate or early-developing sensory-motor capacities that bring us into relation with others and allow us to interact with them. These capacities are manifested at the level of perceptual experience –we *see* or more generally *perceive* in the other person’s bodily movements, gestures, facial expressions, eye direction, and so on what they intend and what they feel, and we respond with our own bodily movements, gestures, facial expressions, gaze, and so on. From birth the infant is pulled into these interactive processes. This can be seen in the very early behavior of the newborn. Infants from birth are capable of perceiving and imitating facial gestures presented by another (Meltzoff and Moore, 1977; 1994). Importantly, this kind of imitation is not an automatic or mechanical procedure; Csibra and Gergely (2009) have shown, for example, that the infant is more likely to imitate only if the other person is attending to it.





Primary intersubjectivity can be specified in more detail as the infant develops. At two months old, for example, infants are able to follow the gaze of the other person, to see that the other person is looking in a certain direction, and to sense what the other person sees (which is sometimes the infant herself), in a way that throws the intention of the other person into relief (Baron-Cohen, 1995; Maurer and Barrera, 1981). In addition, second-person *interaction* is evidenced by the timing and emotional response of infants' behavior. Infants "vocalize and gesture in a way that seems [affectively and temporally] 'tuned' to the vocalizations and gestures of the other person" (Gopnik and Meltzoff, 1997, p. 131). Murray and Trevarthen (1985) have shown the importance of the mother's live interaction with two-month-old infants in their double TV monitor experiment in which mother and infant interact by means of a live television link. The infants engage in lively interaction in this situation. When presented with a recorded replay of their mother's previous actions, however, they quickly disengage and become distracted and upset.

AQ7

At five to seven months, infants are able to detect correspondences between visual and auditory information that specify the expression of emotions (Hobson, 1993; 2002; Walker, 1982). At six months infants start to perceive grasping as goal-directed, and at 10–11 months infants are able to parse some kinds of continuous action according to intentional boundaries (Baird and Baldwin, 2001; Baldwin and Baird, 2001; Woodward and Sommerville, 2000). They start to perceive various movements of the head, the mouth, the hands, and more general body movements as meaningful, goal-directed movements (Senju et al., 2006).

By the end of the first year of life, infants have a nonmentalizing, perceptually based, embodied, and pragmatic understanding of the intentions and dispositions of other persons. With the advent of joint attention (at around nine months) and secondary intersubjectivity (at around one year), infants start to use context and enter into situations of participatory sense-making (De Jaegher and Di Paolo, 2007). That is, infants begin to co-constitute the meaning of the world in their interactions with others. They start to understand the world through their interactions with others, and they gain a more nuanced understanding of others by situating their actions in contexts that are defined by both pragmatic tasks and cultural practices.

Insights about secondary intersubjectivity have recently been complemented by a new line of evidence, namely, spontaneous response tasks which test the ability of young infants to understand false beliefs



132 *Shaun Gallagher*

(see Baillargeon et al., 2010 for a review). The experimenters suggest that series of such tasks show that infants can attribute to an agent a false belief about an object's location (e.g., Onishi and Baillargeon, 2005; Southgate et al., 2007; Song et al., 2008; Träuble et al., 2010). In addition, the experiments have been extended to test and demonstrate infants' understanding of another's false perception of an object, and another's false belief about an object's identity (e.g., Baillargeon and Song, 2008). While the design and interpretation of these studies is still largely determined by ToM approaches, which treat the infant as a detached and isolated observer, more interactive experimental designs are also starting to appear. For instance, in a study by Buttelmann et al. (2009) it was found that 18-month-olds would try to practically help an experimenter retrieve a toy in a way that took into account that the experimenter's previous actions were informed by a false belief about the hidden toy's location (see Southgate et al., Csibra, 2010 for similar results). These studies indicate that the capacities for understanding context and social situations, and for understanding others in such situations, are closely intertwined with the ability to deploy social competences to engage with those situations and to interact with other agents.

### **(b) Behavioral and phenomenological evidence**

Neither primary nor secondary intersubjectivity disappears after the first year of life. These are not stages that we leave behind, and they are not, as Greg Currie suggests, a set of precursor states "that underpin early intersubjective understanding, and *make way* for the development of later theorizing or simulation" (2008, p. 212; emphasis added; cf. Baron-Cohen, 1991; 1995). Rather, citing both behavioral and phenomenological evidence, IT argues that primary and secondary intersubjectivity are both operative in adult social engagements; the processes involved here don't "*make way*" for the purportedly more sophisticated mind-reading processes – these embodied interactive processes continue to characterize our everyday encounters even as adults. That is, we continue to understand others in strong interactional terms, facilitated by our recognition of facial expressions, gestures, postures, and actions as meaningful.

Scientific experiments bear this out. Point-light experiments (actors in the dark wearing point lights on their joints, presenting abstract physical patterns of emotional and action postures), for example, show that not only children (although not autistic children) but also adults perceive emotion even in movement that offers minimal



information (Dittrich et al., 1996; Hobson and Lee, 1999). Close analysis of facial expression, gesture, and action in everyday contexts shows that as adults we continue to rely on embodied interactive abilities to understand the intentions and actions of others and to accomplish interactive tasks (Lindblom, 2007; Lindblom and Ziemke, 2007).

AQ8

Accordingly, meaning and emotional significance are co-constituted in the interaction—not in the private confines of one or the other's head. The analyses of social interactions in shared activities, in working together, in communicative practices, and so on, show that agents unconsciously coordinate their movements, gestures, and speech acts (Issartel et al., 2007; Kendon, 1990; Lindblom, 2007). In the contextualized practices of secondary intersubjectivity, timing and emotional attunement continue to be important as we coordinate our perception-action sequences; our movements are coupled with changes in velocity, direction, and intonation of the movements and utterances of the speaker.

Phenomenology also bears this out. A frequently heard objection is that phenomenology cannot tell us anything about subpersonal processes and so is irrelevant to explanations of social cognition (see, e.g., Spaulding, 2010). If, however, social cognition is not something reducible to individual subpersonal processes, that is, if social cognition depends on interaction processes that are best understood on personal and *superpersonal* levels of explanation, then phenomenology has some relevance. Furthermore, even if a phenomenological assessment of lived experience is insufficient by itself to make the case for interaction, the fact that it supports both the behavioral and developmental evidence suggests that phenomenological and enactive accounts may provide a better explanatory model than ToM, which lacks the coherence offered by an approach to the social which encompasses the first- and third-person perspectives within a framework of second-person engagement (e.g., Reddy and Morris, 2004).

In this spirit, consider Merleau-Ponty's concept of intercorporeity. The evidence suggests that from birth the action of the infant and the perceived action of the other person are coded in the same "language," a cross-modal sensory-motor system that is directly attuned to the actions and gestures of other humans (Gallagher and Meltzoff, 1996; Meltzoff and Moore, 1994). Phenomenology suggests that in this kind of interaction there is a bodily intentionality distributed across the interacting agents, an intentionality that couldn't be realized without there being an actual interaction. Merleau-Ponty calls this "intercorporeity," and characterizes it in this way: "between this phenomenal body of



134 *Shaun Gallagher*

mine, and that of another as I see it from the outside, there exists an internal relation which causes the other to appear as the completion of the system" (1945/1962, p. 352; see 1968, pp. 141, 143). Intercorporeity involves a mutual influence of body schemas, but not, as Gallese (2009; 2010) suggests, in an isomorphic format, whereby one mirrors, or maps the other's actions onto one's own motor representations. Rather, intercorporeity involves a reciprocal, dynamic, and enactive response to the other's action, taking that action as an affordance for further action rather than an opportunity for replication (simulation). This enactive approach offers an alternative (nonsimulationist) interpretation of MN activation in the social cognition process (Gallagher, 2008a). Consistent with the suggestion made by Newman-Norlund et al. (2007), activation of the broadly congruent mirror neurons may be preparatory for an enactive response rather than a matching action.

### (c) Dynamic systems modeling

Dynamic systems modeling relies on the use of computer simulations, as found, for example, in the field of evolutionary robotics (e.g., Beer, 2000; 2003; Harvey et al., 2005). These models can serve as a useful technological supplementation of phenomenological methodology (Froese and Gallagher, 2010). Theorists have used this approach to investigate minimally social behavior (Froese and Di Paolo, 2010; in press). Basing their model on Murray and Trevarthen's (1985) "double TV monitor" contingency study (discussed above), Iizuka and Di Paolo (2007), for example, used an evolutionary robotics approach to show that the detection of social contingency emerges from the dynamics of the interaction process itself. In their simulation model the evolved agents successfully acquired the capacity to discriminate between 'live' (interactive) and 'recorded' (one-way, noninteractive) relations. Dynamic systems analysis demonstrates that this capacity cannot be reduced to the isolated individual agent, but that the dynamics of the interaction process itself play an essential role in enabling this behavior. When the agent attempts to interact with a nonresponsive "partner" whose movements are merely played back from a recording of a previously highly successful encounter, the interaction fails to materialize. Individual interactors do *not* achieve their performance by utilizing internal computational mechanisms, such as "social contingency detection modules." Rather, their successful performance constitutively depends on dynamical properties of their mutual coupling.

The evidence for interaction theory shows that social cognition is both dynamic and enactive in nature. On the enactive view, we engage

# PROOF



with others in ways that depend on embodied sensorimotor processes. We do not first perceive nonintentional movements, and then make inferences to what they mean. We enactively perceive the actions and emotional expressions of others as a form of intentionality –that is, as meaningful and directed. Enactive perception of others means that we see their emotional expressions and contextualized actions as meaningful in terms of how we might respond to or interact with them. Others present us with social affordances. Accordingly, our understanding of others is pragmatic and it references their actions in context: it is not indexed to Cartesian mental states that would explain their actions.

Another way to say this is that we ordinarily perceive another's intentionality in the form of "operative intentionality" rather than mental "act intentionality." As we indicated above, the concept of operative intentionality attempts to capture the fact that the experiencing agent is intentionally engaged with the world through actions and projects that are not reducible to simple mental states, but involve an intentionality that is motoric and bodily. Actions have intentionality because they are directed at some goal or project, and this is something that we can see in the actions of others. Operative intentionality is quite different from mental state (or act) intentionality, which is garnered in reflective inference or judgment (Merleau-Ponty, 1962, p. xviii). The latter seems to be what we appreciate when we try to explain or predict others' behaviors from a detached, observatory standpoint, or reflect upon others' behaviors rather than when we enactively engage with their intentional behavior. In contrast, we usually experience both others and ourselves in terms of operative intentionality, an intentionality "which brings about the natural and prepredicative unity of the world and of our lives, which appears more clearly...in our visual field than in objective knowledge" (Merleau-Ponty, 1962, p. xviii). With respect to social cognition, we normally perceive another's intentionality in terms of its appropriateness, its pragmatic and/or emotional value for our particular way of being, constituted by the particular goals or projects we have at the time, our implicit grasp on cultural norms, our social status, and so on, rather than as reflecting inner mental states, or as constituting explanatory reasons for her further thoughts and actions.

Consider the following example (from Miyahara, in press). Suppose you are driving a car along a busy street and see a person restlessly looking left and right at the edge of the street where there are no crosswalks. You slow down a little in case he runs onto the street, or at least you ready yourself to press the brake pedal. If the passenger in the car with you asked you why you slowed down, you might answer that the person



# PROOF

136 *Shaun Gallagher*

looked like he *wanted* to cross the road. In this reflective explanation it seems as if the person had been experienced in terms of his mental states, that is, his *desire* to cross the road, which constitutes a reason for a further action of crossing the road. This, however, is a way of putting it that is forced by reflection. In fact, in the original action, placing your foot on the brake pedal just is part of what it means to experience the intentionality of the person at the edge of the road. As Merleau-Ponty puts it:

Our bodily experience of movement is not a particular case of knowledge; it provides us with a way of access to the world and the object, with a “praktognosia”, which has to be recognized as original and perhaps as primary. My body has its world, without having to make use of “symbolic” or “objectifying function” (1962, pp. 140–1).

Making such bodily responses to the world or to an object, or, in social contexts, to others, is a way of encountering such entities that not only cannot be reduced to actions guided by the mediation of reasonings but is also more primitive than the kind of recognition of the world that guides action only indirectly.

Enactive phenomenologists (and the interaction theory of social cognition) claim that this intersubjective and pragmatic understanding is the basic kind of understanding we have of others' and our own intentionality, and that this intentionality is primary and nonderived. On this notion of intentionality “the unity of the world, before being posited by knowledge in an explicit act of identification, is lived as already made or already there” (Merleau-Ponty, 1962, p. xvii). Intentionality is determined by what the agent is doing and what the agent is ready to do – that is, in the agent's sensorimotor skills to cope with the situation at hand – and that holds for both stepping off a curb and stepping on the brake, and for any interaction that might follow.

## Enactive and extended minds

On the enactive view, one doesn't need to go to the level of mental states (propositional attitudes, beliefs, desires, inside the head) to encounter intentionality – operative intentionality is in the movement, in the action, in the environmentally attuned responses. This operative intentionality is the real (nonderived, primary) intentionality. Anything like attributed intentionality in terms of mental states is derived from this,

# PROOF



*Neo-Pragmatism and Enactive Intentionality* 137

and, in most cases of everyday interaction, is unnecessary, redundant, and not necessarily real. Moreover, operative intentionality is clearly distributed over brain–body–environment – and so it is precisely the concept/theory of intentionality that enactive and extended accounts need.

This account is also consistent with the neo-pragmatist view. Indeed, IT shows us how to connect the very basic operative intentionality with the neo-pragmatist emphasis on social/normative aspects of behavior. Brandom makes the connection between operative intentionality and neo-pragmatism very clear.

A founding idea of pragmatism is that the most fundamental kind of intentionality (in the sense of directedness towards objects) is the *practical* involvement with objects exhibited by a sentient creature dealing skillfully with its world. (2008, p. 178)

Brandom pictures this intentionality as more basic than language-based “*semantic* intentionality,” and as involving feedback-governed processes that extend into the world, and which exhibit “a complexity [that] cannot in principle be specified without reference to the changes in the world that are both produced by the system’s responses and responded to.... [Such practices] are ‘thick’, in the sense of essentially involving objects, events, and worldly states of affairs. Bits of the world are *incorporated* in such practices.” (p. 178)

It’s clear, however, that (according to IT and developmental studies) this kind of intentionality (and the possibility of recognizing and attributing such intentionality to others) comes in earlier than any discussion of social norms might indicate, namely, in the intercorporeity of primary intersubjectivity. This is clearly prior to what develops as social/normative aspects of behavior later in childhood. IT understands this later development as primarily depending on communicative and narrative practices (Gallagher and Hutto, 2008).

Again, my understanding of the other’s intentionality is not based on the idea that I take myself as simulating the other, or as a perception of a normative similarity between my own actions and those of the other; it is, rather, that I see the other’s actions as an affordance for my own possible action (which may be very different from hers); I see the other’s action as interactional or as calling forth a response on my part.

We note that this notion of intentionality also provides a better account of both *erotic* intentionality and our attribution of intentionality to



# PROOF

138 *Shaun Gallagher*

nonhumans. In erotic perception, which is not a *cogitatio* but a sexual significance for me “when it exists for my body,”

We discover both that sexual life is one more form of original intentionality, and also brings to view the vital origins of perception, motility and [symbolic] representation by basing all these “processes” on an intentional arc.... (Merleau-Ponty, 1962, p. 157)

Erotic intentionality, like every instance of operative intentionality, is not an “I think that...” but an “I can....” – it is whatever I recognize as something to which I could respond or interact erotically.

Consider again the attribution of intentionality to geometrical figures on a computer screen. I do not have to take them as *similar* to human, socially normative actions to understand the intentionality which, in fact, is in their very movement. In fact, as long as it was programmed into them by human experimenters, their movement *is*, rather than being merely similar to, a display of human action, just as a car cruising on the street is a display of the driver’s intention, project, or commitment to drive the car. In both cases, I perceive the movement as something with which I could interact to some end. In the experimental case, one could easily picture a virtual reality where I, as a human subject, am in the scene with the geometrical objects, and where I could intervene, for example, to prevent one object from “chasing” another. This possibility for intervention on my part is what I see in their movement as meaningful, and what constitutes the basis for my attribution of intentionality. At the same time, this notion of intentionality can explain why I don’t take those figures as true intentional agents: this feeling is due to the fact that they are presented in an environment where the possibility for interactions is largely restricted. On the one hand, I see the possibility for intervening, but, on the other hand, I know that I actually cannot intervene. It is this ambiguity that is reflected in my ambivalent appraisal of the figure’s intentionality.

This enactive, neo-pragmatic, operative concept of intentionality is precisely the relevant concept needed to support the extended mind hypothesis. As Dewey made clear, long before the proponents of the extended mind formed their hypothesis, this is a pragmatic concept of mind: the mind “is formed out of commerce with the world and is set toward that world;” it should never be regarded as “something self-contained and self-enclosed” (1934, p. 269). That this concept of mind is intersubjective from the very start means that there is no mystery about where this nonderived intentionality comes from. It comes from



the others with whom we interact, or, more precisely, it is generated in our interaction. To the extent that we are all born into a community, our environment is full of intentional practices from the very beginning of our life. We develop and shape our intentionality by being initiated into this communal practice in virtue of actual interactions with other people, primarily with our caregivers, and in virtue of our innate or early-learned sensitivity to them or to opportunities for such interactions. This means that nonderived intentionality is not something that is first generated in my own isolated mind, or in brain processes that are not already directed to and by others. What Adams and Aizawa call nonderived content is surely derived from these originary interactive practices. In this regard, the mind is constituted by our enactive engagements with the environment, which is both social and physical; and intentionality means that we are “in-the-world,” distributed over brain–body–environment, and extended in pragmatic and communicative practices that may further supervene on the tools, technologies, and institutional practices through which we engage with the world.

### Notes



Miyahara's work on this paper was supported by a Japanese Society for the Promotion of Science Research Fellowship for Young Scientists.



1. Here is his original formulation. “Here is how it works: first you decide to treat the object whose behavior is to be predicted as a rational agent; then you figure out what beliefs that agent ought to have, given its place in the world and its purpose. Then you figure out what desires it ought to have, on the same considerations, and finally you predict that this rational agent will act to further its goals in the light of its beliefs. A little practical reasoning from the chosen set of beliefs and desires will in most instances yield a decision about what the agent ought to do; that is what you predict the agent will do.”(Dennett, 1987, p. 17)
2. The neo-behaviorist view is quite consistent with the functionalist position defended by Clark. Indeed, Horgan and Kriegel (2008) contend that the neo-behaviorist position would be the only viable option for the extended mind hypothesis, although more generally they think the extended mind hypothesis is not viable.
3. See, for example, recent false-belief experiments in infants 15 months of age and younger (e.g., Baillargeon et al., 2010), discussed below.
4. Adams and Aizawa (2009) state that we can understand “why even transcranialists [i.e., extended and enactive mind theorists] maintain that cognition extends from brains into the extraorganismal world rather than from the extraorganismal world into brains” (p. 92) based on the fact that nonneural external processes are actually noncognitive. We agree with them that it





is misleading for the extended and enactive mind theorists to describe the mind as *extending from brains into the world*. We disagree with them that such misleading descriptions constitute evidence in support of the standard model of the mind as “in the head;” rather, such descriptions reflect remnants of the old model in the extended mind literature, which need to be removed in order to fully appreciate the potential of the claim.

5. In this section we consider one version of neo-pragmatism, as found in Brandom and a few other theorists. There are a number of different contemporary versions of neo-pragmatism being developed, as in the work of Johnson (2008), Schulkin (2008), and others. Our focus on Brandom is purely pragmatic; he directly addresses the issue of intentionality in a way that facilitates our analysis here.
6. Cash (2010) describes it as follows: “On this normative view...the paradigmatic cases of such ascriptions are made by another member of the agent’s linguistic and normative community; the ascriptions abide by, and are justified by, the norms of that community’s practice of giving intentional states as reasons for actions. This practice is firmly situated in and supported by that community’s shared, public language, with its norms regulating the appropriate uses of words to give content to intentional states....This practice constrains what ascriptions an observer is licensed to ascribe according to the agent’s behavior. But they also normatively constrain the further actions of the agent. Agents who recognize that observers are licensed to ascribe particular intentional states to them ought to take themselves to be committed to further actions consistent with those intentional states. If I say to you that I intend to go for a walk, I should recognize that this utterance licenses you to ascribe to me the intention to go for a walk; I have licensed you to expect me to go for a walk, and thus I have placed myself under a commitment (*ceteris paribus*) to go for a walk.”
7. As Miyahara (in press) points out, this goes against Clark’s commitment to the possibility of cognition in nonhuman animals or the “biological mind” (1997, p. 1). In this regard, he has good reason to adopt the neo-behaviorist concept of intentionality rather than the neo-pragmatic one.
8. Csibra concludes: “With strongly unequal distribution of types of action or types of grip, one could find a relatively high proportion of good match between the [observed action vs. executed action] domains even if there were no causal relation between them. Without such a statistical analysis, it remains uncertain whether the cells that satisfy the definition of ‘mirror neurons’ (i.e., the ones that discharge both with execution and observation of actions) do indeed have ‘mirror properties’ in the everyday use of this term (i.e., are generally activated by the same action in both domains)” (2005, p. 3).

## References

- Adams, F. and Aizawa, K. (2001). The bounds of cognition. *Philosophical Psychology*, 14, no. 1, 43–64.
- AQ9 Adams, F. and Aizawa, K. (2008). *The Bounds of Cognition*. Malden, MA: Blackwell.
- Adams, F. and Aizawa, K. (2009). Why the mind is still in the head. In P. Robbins and M. Aydede (Eds), *The Cambridge Handbook for Situated Cognition*. New York: Cambridge University Press, pp. 78–95.



# PROOF



- Baillargeon, R. and Song, H.-J. (2008). Infants' reasoning about others' false perceptions. *Developmental Psychology, 44*, no. 6, 1,789–95.
- Baillargeon, R., Scott, R. M., and He, Z. (2010). False-belief understanding in infants. *Trends in Cognitive Sciences, 14*, no. 3, 110–18.
- Baird, J. A., and Baldwin, D. A. (2001). Making sense of human behavior: Action parsing and intentional inference. In B. F. Malle, L. J. Moses, and D. A. Baldwin (Eds), *Intentions and intentionality: Foundations of social cognition*. Cambridge, MA: MIT Press, pp. 193–206.
- Baldwin, D. A., and Baird, J. A. (2001). Discerning intentions in dynamic human action. *Trends in Cognitive Science, 5*, no. 4, 171–8.
- Baron-Cohen, S. (1991). Precursors to a theory of mind: Understanding attention in others. In A. Whiten (Ed.), *Natural theories of mind: Evolution, development and simulation of everyday mindreading*. Oxford, UK: Basil Blackwell, pp. 233–51.
- Baron-Cohen, S. (1995). *Mindblindness: An Essay on Autism and Theory of Mind*. Cambridge, MA: MIT Press.
- Beer, R. D. (2000). Dynamical approaches to cognitive science. *Trends in Cognitive Sciences, 4*, no. 3, 91–9.
- Beer, R. D. (2003). The dynamics of active categorical perception in an evolved model agent. *Adaptive Behavior, 11*, no. 4, 209–43.
- Brandom, R. B. (1994). *Making It Explicit: Reasoning, Representing, and Discursive Commitment*. Cambridge, MA: Harvard University Press.
- Brandom, R. B. (2000). *Articulating Reasons: An Introduction to Inferentialism*. Cambridge, MA: Harvard University Press.
- Brandom, R. B. (2008). *Between Saying and Doing: Towards an Analytic Pragmatism*. Oxford: Oxford University Press.
- Brentano, F. (2008/1862). *Psychologie vom empirischen Standpunkte. Von der Klassifikation der psychischen Phänomene*. (Mauro Antonelli, Ed.) Heusenstamm: Ontos.
- Buccino, G., Lui, F., Canessa, N., Patteri, I., Lagravinese, G., Benuzzi, F., Porro, C. A., and Rizzolatti, G. (2004). Neural circuits involved in the recognition of actions performed by non-conspecifics: an fMRI study. *Journal of Cognitive Neuroscience, 16*, no. 1, 114–126.
- Buttelmann, D., Carpenter, M., and Tomasello, M. (2009). Eighteen-month-old infants show false belief understanding in an active helping paradigm. *Cognition, 112*, 337–42.
- Cash, M. (2008). Thought and Oughts. *Philosophical Explorations, 11*, no. 2, 93–119.
- Cash, M. (2009). Normativity is the mother of intentionality: Wittgenstein, normative practices and neurological representations. *New Ideas in Psychology, 27*, 133–47.
- Cash, M. (2010). Extended cognition, personal responsibility, and relational autonomy. *Phenomenology and the Cognitive Sciences, 9*, no. 4, 645–71.
- Catmur, C., Walsh, V., and Heyes, C. (2007). Sensorimotor learning configures the human mirror system. *Current Biology, 17*, no. 1, 527–31.
- Clark, A. (1997). *Being There: Putting Brain, Body and World Together Again*. Cambridge, MA: MIT Press.
- Clark, A. (2008). *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*. Oxford: Oxford University Press.



# PROOF

142 *Shaun Gallagher*

- Clark, A. and Chalmers, D. (1998). The extended mind. *Analysis*, 58, no. 1, 7–19.
- Clark, A. and Grush, R. (1999). Towards a cognitive robotics. *Adaptive Behavior*, 7, no. 1, 5–16.
- Crane, T. (1998). Intentionality as the mark of the mental. In A. O'Hear (Ed.), *Contemporary Issues in the Philosophy of Mind*. Royal Institute of Philosophy Supplement (No. 43). Cambridge: Cambridge University Press.
- AQ10 Csibra, G. (2005). Mirror neurons and action observation. Is simulation involved? *ESF Interdisciplines*. <http://www.interdisciplines.org/mirror/papers/>
- Csibra, G. and Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences*, 13, 148–53.
- Currie, G. (2008). Some ways to understand people. *Philosophical Explorations*, 11, no. 3, 211–18.
- De Jaegher, H. and Di Paolo, E. A. (2007). Participatory sense-making: An enactive approach to social cognition. *Phenomenology and the Cognitive Sciences*, 6, no. 4, 485–507.
- De Jaegher, H., Di Paolo, E. A., and Gallagher, S. (2010). Can social interaction constitute social cognition? *Trends in Cognitive Sciences*, 14, no. 10, 441–7.
- Dennett, D. (1971). Intentional systems. *The Journal of Philosophy*, 68, no. 4, 87–106.
- Dennett, D. (1987). *The Intentional Stance*. Cambridge, MA: MIT Press.
- AQ11 Dennett, D. (1991). *Consciousness Explained*. Boston: Little, Brown, and Co.
- Dewey, J. (1934). *Art as Experience*. New York: Perigee/Berkley.
- Dinstein, I., Thomas, C., Behrmann, M., and Heeger, D. J. (2008). A mirror up to nature. *Current Biology*, 18, no. 1, R13–R18.
- Di Paolo, E. (2009). Extended life. *Topoi*, 28, no. 1, 9–21.
- Dittrich, W.H., Troscianko, T., Lea, S.E.G., and Morgan, D. (1996). Perception of emotion from dynamic point-light displays represented in dance. *Perception*, 25, 727–38.
- Froese, T. and Di Paolo, E. A. (2010). Modeling social interaction as perceptual crossing: An investigation into the dynamics of the interaction process. *Connection Science*, 22, no. 1, 43–68.
- Froese, T. and Di Paolo, E. A. (in press). Toward Minimally Social Behavior: Social Psychology Meets Evolutionary Robotics. *Advances in Artificial Life: Proc. of the 10th Euro. Conf. on Artificial Life*. Berlin, Germany: Springer.
- Froese, T. and Gallagher, S. (2010). Phenomenology and Artificial Life: Toward a Technological Supplementation of Phenomenological Methodology. *Husserl Studies*, 26, no. 2, 83–106.
- Froese, T. and Gallagher, S. (submitted). Being together: Strong interaction and dynamical systems theory.
- Gallagher, S. (2001). The Practice of Mind: Theory, Simulation or Primary Interaction? *Journal of Consciousness Studies*, 8, no. 5–7, 83–108.
- Gallagher, S. (2005). *How the Body Shapes the Mind*. New York, NY: Oxford University Press.
- Gallagher, S. (2007). Simulation trouble. *Social Neuroscience*, 2, no. 3–4, 353–65.
- Gallagher, S. (2008a). Are minimal representations still representations? *International Journal of Philosophical Studies*, 16, no. 3, 351–69.
- Gallagher, S. (2008b). Neural simulation and social cognition. In J. A. Pineda (Ed.), *Mirror Neuron Systems: The Role of Mirroring Processes in Social Cognition*. Totowa, NJ: Humana Press, pp. 355–71.

# PROOF



- Gallagher, S. (in press). Narrative competency and the massive hermeneutical background. In P. Fairfield. *Education, Dialogue, and Hermeneutics*. New York: Continuum.
- Gallagher, S. and Hutto, D. (2008). Understanding others through primary interaction and narrative practice. In J. Zlatev, T. P. Racine, C. Sinha, and E. Itkonen (Eds), *The Shared Mind: Perspectives on Intersubjectivity*. Amsterdam: John Benjamins, pp. 17–38.
- Gallagher, S. and Meltzoff, A. (1996). The earliest sense of self and others: Merleau-Ponty and recent developmental studies. *Philosophical Psychology*, 9, 213–36.
- Gallagher, S. and Zahavi, D. (2008). *The Phenomenological Mind: An Introduction to Philosophy of Mind and Cognitive Science*. London, UK: Routledge.
- Gallese, V. (2005). ‘Being like me’: Self-other identity, mirror neurons and empathy. In S. Hurley and N. Chater (Eds), *Perspectives on Imitation I*. Cambridge, MA: MIT Press, pp. 101–18.
- Gallese, V. (2009). The Two Sides of Mimesis: Girard’s Mimetic Theory, Embodied Simulation and Social Identification. *Journal of Consciousness Studies*, 16, no. 4, 21–44.
- Gallese, V. (2010). Embodied Simulation and its Role in Intersubjectivity. In T. Fuchs, H. C. Sattel, and P. Henningsen (Eds), *The Embodied Self: Dimensions, Coherence and Disorders*. Stuttgart, Germany: Schattauer, pp. 78–92.
- Gallese, V., and Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, 2, 493–501.
- Goldman, A. I. (2006). *Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading*. New York: Oxford University Press.
- Gopnik, A. and Meltzoff, A. N. (1996). *Words, Thoughts, and Theories*. Cambridge, MA: MIT Press.
- Harvey, I., Di Paolo, E.A., Wood, R., Quinn, M., and Tuci, E. A. (2005). Evolutionary Robotics: A new scientific tool for studying cognition. *Artificial Life*, 11, no. 1–2, 79–98.
- Haugeland, J. (1990). Intentionality all-stars. *Philosophical Perspectives*, 4, 383–427. Reprinted in J. Haugeland, *Having Thought: Essays in the Metaphysics of Mind*. Cambridge, MA: Harvard University Press, pp. 127–70.
- Heider, F., and Simmel, M. (1944). An experimental study of apparent behavior. *The American Journal of Psychology*, 57, no. 2, 243–59.
- Hickok, G. (2009). Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *J CognNeurosci*, 21, no. 7, 1,229–43.
- Hobson, P. (1993). The emotional origins of social understanding. *Philosophical Psychology*, 6, 227–49.
- Hobson, P. (2002). *The Cradle of Thought*. London: Macmillan.
- Hobson, P. and Lee, A. (1999). Imitation and identification in autism. *Journal of Child Psychology and Psychiatry*, 40, 649–59.
- Horgan, T. E. and Kriegel, U. (2008). Phenomenal intentionality meets the extended mind. *The Monist*, 91, 353–80.
- Horgan, T. E., and Tienson, J. L. (2002). The intentionality of phenomenology and the phenomenology of intentionality. In D. J. Chalmers (Ed.), *Philosophy of Mind: Classical and Contemporary Readings*. Oxford: Oxford University Press.
- Husserl, E. (1977). *Phenomenological Psychology*. (J. Scanlon, Trans.). The Hague: Martinus Nijhoff.



# PROOF

144 *Shaun Gallagher*

- Husserl, E. (1982/1913). *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy – First Book.* (F. Kersten, Trans.). The Hague: Martinus Nijhoff.
- Hutto, D. (in press). Philosophy of mind's new lease on life: Autopoietic enactivism meets teleosemiotics. *Journal of Consciousness Studies* (Spring 2011).
- Iizuka, H. and Di Paolo, E. A. (2007). Minimal agency detection of embodied agents. In F. Almeida e Costa, L. M. Rocha, E. Costa, I. Harvey, and A. Coutinho (Eds), *Advances in Artificial Life: Proc. of the 9th Euro. Conf. on Artificial Life*. Berlin, Germany: Springer, pp. 485–94.
- Issartel, J., Marin, L., and Cadopi, M. (2007). Unintended interpersonal co-ordination: 'can we march to the beat of our own drum?' *Neuroscience Letters*, 411, no. 3, 174–9.
- Johnson, M. (2008). *The Meaning of the Body: Aesthetics of Human Understanding*. Chicago: University of Chicago Press.
- Kendon, A. (1990). *Conducting Interaction: Patterns of behavior in focused encounters*. Cambridge, UK: Cambridge University Press.
- Keyser, C. and Gazzola, V. (2006). Towards a unifying neural theory of social cognition. In S. Anders, G. Ende, M. Junghofer, and J. Kissler (Eds), *Understanding Emotions*. Amsterdam: Elsevier, 379–402.
- Lindblom, J. (2007). *Minding the Body – Interacting socially through embodied action*. Dissertation, University of Linköping and University of Skövde.
- Lindblom, J. and Ziemke, T. (2008). Interacting socially through embodied action. In F. Morganti, A. Carassa, and G. Riva (Eds), *Enacting Intersubjectivity: A Cognitive and Social Perspective on the Study of Interactions*. Amsterdam, Netherlands: IOS Press, pp. 49–63.
- Maurer, D. and Barrera, M. E. (1981). Infants' perception of natural and distorted arrangements of a schematic face. *Child Development*, 52, no. 1, 196–202.
- Meltzoff, A. N. and Moore, M. K. (1977). Imitation of facial and manual gestures by human neonates. *Science*, 198, 75–8.
- Meltzoff, A. N. and Moore, M. K. (1994). Imitation, memory, and the representation of persons. *Infant Behavior and Development*, 17, 83–99.
- Menary, R. (Ed.) 2010. *The Extended Mind*. Cambridge, MA: MIT Press.
- Merleau-Ponty, M. (1962/1945). *Phenomenology of Perception*. (C. Smith, Trans.). London: Routledge and Kegan Paul.
- Merleau-Ponty, M. (1968). *The Visible and the Invisible*. (A. Lingis, Trans.). Evanston: Northwestern University Press.
- Michotte, A. (1963). *The Perception of Causality*. (T. Miles, and E. Miles, Trans.). New York: Basic Books.
- Miyahara, K. (in press). Neo-pragmatic intentionality and enactive perception: A compromise between extended and enactive minds. *Phenomenology and the Cognitive Sciences*.
- Murray, L. and Trevarthen, C. (1985). Emotional regulations of interactions between two-month-olds and their mothers. In T. M. Field and N. A. Fox (Eds), *Social perception in infants*. Norwood, NJ: Ablex Publishing, pp. 177–97.
- Newman-Norlund, R.D., Noordzij, M.L., Meulenbroek, R.G.J., and Bekkering, H. (2007). Exploring the brain basis of joint attention: Co-ordination of actions, goals and intentions. *Social Neuroscience*, 2, no. 1, 48–65.
- Noë, A. (2004). *Action in Perception*. Cambridge, MA: MIT Press.

# PROOF



- Oberman, L. M. and Ramachandran, V. S. (2007). The simulating social mind: The role of the mirror neuron system and simulation in the social and communicative deficits of autism spectrum disorders. *Psychological Bulletin*, 133, no. 2, 310–27.
- Onishi, K.H., and Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, 308, 255–8.
- Pettit, P. (1996). *The Common Mind*. Oxford: Oxford University Press.
- Ratcliffe, M. (2007). *Rethinking Commonsense Psychology: A Critique of Folk Psychology, Theory of Mind and Simulation*. Basingstoke: Palgrave Macmillan.
- Reddy, V. and Morris, P. (2004). Participants don't need theories: Knowing minds in engagement. *Theory and Psychology*, 14, no. 5, 647–65.
- Rizzolatti, G., Fogassi, L., and Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews Neuroscience*, 2, 661–70.
- Rowlands, M. (2006). *Body Language*. Cambridge, MA: MIT Press.
- Rupert, R. (2004). Challenges to the hypothesis of extended cognition. *The Journal of Philosophy*, 101, no. 8, 389–428.
- Schulkin, J. (2008). *Cognitive Adaptation: A Pragmatist Perspective*. Cambridge: Cambridge University Press.
- Searle, J. (1992). *The Rediscovery of the Mind*. Cambridge, MA: MIT Press.
- Senju, A., Johnson, M. H., and Csibra, G. (2006). The development and neural basis of referential gaze perception. *Social Neuroscience*, 1, no. 3–4, 220–34.
- Shapiro, L. (in press). Review of Fred Adams and Ken Aizawa, *The Bounds of Cognition* (Malden, MA: Blackwell, 2008). *Phenomenology and the Cognitive Sciences*.
- Song, H.-J., Onishi, K. H., Baillargeon, R., and Fisher, C. (2008). Can an agent's false belief be corrected by an appropriate communication? Psychological reasoning in 18-month-old infants. *Cognition*, 109, no. 3, 295–315.
- Southgate, V., Senju, A., and Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science*, 18, no. 7, 587–92.
- Southgate, V., Chevallier, C., and Csibra, G. (2010). Seventeen-month-olds appeal to false beliefs to interpret others' referential communication. *Developmental Science*, 13, no. 6, 907–12.
- Spaulding, S. (2010). Embodied cognition and mindreading. *Mind and Language*, 25, no. 1, 119–40.
- Thompson, E. (2007). *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*. Cambridge, MA: Harvard University Press.
- Thompson, E. and Varela, F. (2001). Radical embodiment: Neural dynamics and consciousness. *Trends in Cognitive Sciences*, 5, no. 10, 418–25.
- Träuble, B., Marinović, V., and Pauen, S. (2010). Early theory of mind competencies: Do infants understand others' beliefs? *Infancy*, 15, no. 4, 434–44.
- Trevarthen, C. B. (1979). Communication and cooperation in early infancy: A description of primary intersubjectivity. In M. Bullowa (Ed.), *Before Speech*. Cambridge, MA: Cambridge University Press, pp. 321–47.
- Trevarthen, C. and Hubley, P. (1978). Secondary intersubjectivity: Confidence, confiding and acts of meaning in the first year. In A. Lock (Ed.), *Action, Gesture and Symbol: The Emergence of Language*. London: Academic Press, pp. 183–229.
- Varela, F., Thompson, E., and Rosch, E. (1991). *The Embodied Mind*. Cambridge, MA: MIT Press.



# PROOF

146 *Shaun Gallagher*

- Walker, A. S. (1982). Intermodal perception of expressive behaviors by human infants. *Journal of Experimental Child Psychology*, 33, 514–35.
- Wheeler, M. (2005). *Reconstructing the Cognitive World: The Next Step*. Cambridge, MA: MIT Press.
- Wheeler, M. (2008). Mind, things and materiality. In Renfrew C. and Malafouris L. (Eds.), *The Cognitive Life of Things: Recasting the Boundaries of the Mind*. Cambridge: McDonald Institute for Archaeological Research Publications.
- Woodward, A. L., and Sommerville, J. A. (2000). Twelve-month-old infants interpret action in context. *Psychological Science*, 11, 73–7.



## 6

## Minds, Things, and Materiality

*Michael Wheeler***Taking material culture seriously**

In a rich and thought-provoking paper, Lambros Malafouris argues that taking material culture seriously means to be “systematically concerned with figuring out the causal efficacy of materiality in the enactment and constitution of a cognitive system or operation” (Malafouris, 2004, p. 55). As I understand this view, there are really two intertwined claims to be established. The first is that the things beyond the skin that make up material culture (in other words, the physical objects and artifacts in which cultural networks and systems of human social relations are realized) may be essential to the enactment of, and be partly constitutive of, certain cognitive systems or operations. The consequence of establishing this claim is supposed to be that we have a mandate to recast the boundaries of the mind so as to include, as proper parts of the mind, things located beyond the skin. Thus, in talking about the contribution of the world to cognition, Malafouris (2004, p. 58) concludes that “what we have traditionally construed as an active or passive but always clearly separated external stimulus for setting a cognitive mechanism into motion, may be after all a continuous part of the machinery itself; at least *ex hypothesi*.” This is the position that, in philosophical circles, is known increasingly as the *extended mind hypothesis* (Clark and Chalmers, 1998; Menary, forthcoming). Henceforth I shall refer to this hypothesis as *EM*. Further explication of *EM* will come later, but a stock example, due originally to Rumelhart et al. (1986), will help bring the idea into preliminary view. Most of us solve difficult multiplication problems by using “pen and paper” as an external resource. This environmental prop enables us to transform a difficult cognitive problem into a set of simpler ones,



# PROOF

148 Michael Wheeler

and to temporarily store the results of intermediate calculations. For the fan of EM, the coupled combination of pen-and-paper resource, appropriate bodily manipulations, and in-the-head processing constitutes a cognitive system in its own right.

The second claim that Malafouris needs to establish is that, when things-beyond-the-skin achieve the status of being essential to the enactment of, and partly constitutive of, certain cognitive systems or operations, they often do so in virtue of a kind of causal contribution that is, in some way to be determined, a product of those things' essential materiality, rather than in virtue of some other kind of causal contribution. (I shall later identify the former kind of causal contribution as one that involves *vital materiality* and the latter as one that involves only *implementational materiality*.) One implication of this second claim is that, for EM to be characterized correctly, it needs to pay more than lip service to what we might call the very materiality of material culture.

It is worth noting that, in the context of Malafouris's interests, the payoff from adopting EM is plausibly nothing less than a reconfiguration of the intellectual landscape inhabited by the discipline known as *cognitive archaeology*. Consider the question "What does a Palaeolithic stone tool do for the mind?" Conventional cognitive archaeology is committed to an "in-the-head" ontology of mind. This condemns material culture to a life outside of cognition proper, and so the things studied by cognitive archaeology are (roughly) things that minds have made and/or used. The cognitive states and processes concerned are not themselves on show in those things, although certain inferences about the nature of those states and processes, inferences of an unavoidably hazardous nature, might be ventured. However, if past ways of thought were not just *expressed in* material culture but were often partly *constituted by* material culture, as EM implies, then cognitive archaeology gets to study past minds in a rather more direct fashion. In fact, the things studied by cognitive archaeology are literally parts of (no longer functioning) minds. If this is right, then the interdisciplinary collective that is cognitive science is poised to welcome a new member of the team.

So where are we going in this paper? In what follows I shall spell out what I take to be the only plausible reading of what is the canonical statement of EM, and argue that, on this reading, the distinctive EM conclusion is purchased using a currency of what I shall call "implementational materiality." I shall also submit evidence that Malafouris would judge such implementational materiality to be an inadequate

# PROOF

*Minds, Things, and Materiality* 149

basis for capturing the full cognitive life of things. This goes some way towards breaking the link between Malafouris's vision of what it is to take material culture seriously and EM. To go the rest of the distance I shall argue that the enactive aspect of Malafouris's approach is plausibly in tension with EM. (Recall that we are concerned with the causal efficacy of materiality in the *enactment* and constitution of a cognitive system or operation.) If I am right about this tension, then taking material culture seriously in the way that Malafouris urges us to might actually require us to give up on EM.

Although I have just set things up in terms of a critical response to Malafouris's paper, and although I shall refer back to the argument of that paper as the present investigation unfolds, let me stress at the outset that my real target here is certainly not Malafouris, but, rather, an increasingly widespread tendency in the region of cognitive-scientific space occupied by EM to run together certain importantly different contemporary styles of thinking about thinking. If Malafouris is guilty of this crime, he is far from the only perpetrator. To explain: the wearers of intellectual labels such as situated cognition, embodied-embedded cognitive science, distributed cognition, enactive cognitive science, and EM are wont to march together against the common enemy of a residual Cartesianism in cognitive science. Now, I'm as anti-Cartesian as the next right-thinking cognitive theorist (see, e.g., Wheeler, 2005), but the fact remains that, even if there is a sense in which all these movements are on the same side, their unity against the shared foe serves to obscure some crucial differences between them. (For what I take to be a complementary recent attempt to distinguish EM from certain other embodiment-based approaches, see Clark, 2008). I realize that, from the outside, such differences might look like nothing more than local spats on a par with the tension between the People's Front of Judea and the Judean People's Front, whose joint opposition to Roman rule is overridden by their doctrinal differences. (If this reference is a mystery to you, go and watch Monty Python's *Life of Brian* – now!) However, whatever the pros and cons may be of pursuing interdoctrinal disputes in revolutionary politics, the ignoring of fundamental philosophical differences in cognitive science is a sure-fire recipe for confusion. There must come a time when it is right to recognize and to debate those differences. Indeed, sometimes it's the conflicts between thinkers who are, broadly speaking, on the same side that are the most illuminating. As Reg, the leader of the People's Front of Judea, exclaims, "The only people we hate more than the Romans are the ... Judean People's Front."

# PROOF



150 Michael Wheeler

## So what exactly is the extended mind hypothesis?

Here is what I take to be the canonical (and thus the default) statement of EM:

[Under certain conditions, the] organism is linked with an external entity in a two-way interaction, creating a *coupled system* that can be seen as a cognitive system in its own right. All the components in the system play an active causal role, and they jointly govern behavior in the same sort of way that cognition usually does. If we remove the external component the system's behavioral competence will drop, just as it would if we removed part of its brain. Our thesis is that this sort of coupled process counts equally well as a cognitive process, whether or not it is wholly in the head. (Clark and Chalmers, 1998, p. 7)

This introduces what has come to be known (by those of us affiliated to the EM club, anyway) as *the parity principle*. In broad terms, the parity principle states that if there is equality, with respect to governing behavior, between the causal contribution of certain internal elements and the causal contribution of certain external elements, then there is no good reason to count the internal elements concerned as proper parts of the cognitive system while denying that status to the external elements concerned. Parity of causal contribution mandates parity of cognitive status. Notice that, as stated, EM does *not* claim that there are conditions under which the very idea of any internal–external boundary becomes problematic or misleading. It claims, rather, that there are conditions under which something that counts as a single cognitive system or as a single cognitive process contains some elements which are internal and some which are external. The cognitive process, and thus the mind, is held to be *extended over* that still-in-place internal–external boundary. Of course, this way of talking presupposes that the internal–external boundary at issue is to be fixed by the limits not of the mind, but by those of the brain, or of the skull, or of the central nervous system, or of the skin – take your pick. Clark and Chalmers are not specific on this issue, and for present purposes we need not be either, since, by the lights of the parity principle, there are entities on the external side of any of the latter four interfaces whose causal contribution to behavior (so the EM theorist argues) will end up counting as cognitive.

With that clarification on board, we can ask the following question: what are the benchmarks by which parity of causal contribution is to be



# PROOF



*Minds, Things, and Materiality* 151

judged? Here is the *wrong* way to answer this question. First, we fix the benchmarks for what it is to count as a proper part of a cognitive system by identifying all the details of the causal contribution made by (say) the brain. Then we look to see if any external elements meet those benchmarks. Why is this the wrong way to go? Because it opens the door to the following style of anti-EM argument: we identify some features of, say, internal memory that are not shared by external memory, and we conclude that, since the parity principle is not satisfied, EM is false. Rupert (2004) uses this very strategy as the first part of his memory-oriented critique of EM. Here is his own summary of his own argument:

I argue that the external portions of extended “memory” states (processes) differ so greatly from internal memories (the process of remembering) that they should be treated as distinct kinds; this quells any temptation to argue for [EM] from brute analogy (*viz.* *extended cognitive states are like wholly internal ones; therefore, they are of the same explanatory cognitive kind; therefore there are extended cognitive states*). (Rupert, 2004, p. 407, my emphasis)

Rupert proceeds to discuss empirical psychological data which putatively indicate significant differences between (a) the profile of internal memory and (b) the profile of certain external resources, as such external resources figure in the process of remembering. According to Rupert, such differences tell against any attempt to see the latter phenomena as being of the same explanatory kind as the former. For example, there are psychological experiments which show that internal memory is sensitive to what is called the generation effect. Where this effect is in evidence, subjects gain a mnemonic advantage by generating their own meaningful connections between paired associate items

AQ1

Rupert argues that the generation effect will simply not occur in some extended “memory” systems (e.g., in a system according to which, during recall, the subject refers to a notebook in which the paired associates are accompanied by connection sentences produced by those subjects during learning, but which were entered into the notebook by the experimenter). He concedes that it might occur in others (e.g., in a system according to which, during recall, the subject refers to a notebook in which the paired associates to be learned are accompanied by connection sentences produced and entered by the subjects during learning). In the latter case, however, he concludes that the effect is an accidental feature, rather than a defining dimension, of the memory system.



# PROOF

152 Michael Wheeler

For the sake of argument, let's just agree that Rupert is right about the presence of such differences here. What the fan of EM needs to do is simply refuse to accept that one should allow the extant details of internal memory to set the benchmarks for what counts as memory in general. To be clear, the Rupert-style argument under consideration isn't suspect in virtue of being anti-EM. Rather, it's suspect because it begs the question against EM by assuming that what counts as cognitive should be fixed by the fine-grained profile of the inner. Such question-begging can be avoided, and Rupert's criticism resisted, if we adopt the following alternative strategy for saying what the benchmarks are by which parity of causal contribution is to be judged. First, we give an account of what it is to be a proper part of a cognitive system that is essentially independent of where a candidate element happens to be located with respect to the internal–external boundary (however that boundary is to be determined). Then we look to see where cognition falls – in the brain, in the nonneural body, in the environment, or, as the EM theorist predicts will sometimes be the case, in a system that extends across all of these aspects of the world.

Rupert sees this sort of response coming, and so develops his memory-oriented critique further by arguing that any attempt to fix a generic kind that would subsume internal and extended systems would need to be so devoid of detail (in order to subsume all the different profiles) that it would fail to earn its explanatory keep. But this seems wrong. Indeed, it's important to note that we would surely not intuitively withdraw the epithet "memory" from an internally located system which did not exhibit the generation effect, but which continued to achieve (something like) the selective storage and context-sensitive retrieval of information, so why should we withdraw that epithet from an extended system with a similar profile? But, if that's right, why think that exhibiting the generation effect is a defining dimension of memory, rather than an accidental feature? This gives us some reason to think that there must be a generic account of what memory is that covers both cases, and that has explanatory bite.

Here it is worth pausing momentarily to consider a different response to the clearly prejudiced fix-according-to-the-inner strategy. This response says that we should first work out the details of what the brain does, and then remove from our list of features any details that are inessential to that contribution *as cognitive*. That way we will be able to rule out the arbitrary exclusion of external elements and arrive at a viable set of criteria for what it is for something to count as a proper part of a cognitive system. However, any *decent* version of this response must,



it seems, collapse into a version of the strategy that I am recommending, since, in order for some detail of a causal contribution to be judged inessential to that contribution as cognitive, one must have access to an independent theory of the cognitive. The alternative, which would involve ruling out a detail purely on the grounds that it is not shared by some external element under consideration, would, of course, beg the question against the opponent of EM.

So what sort of overarching theory of the cognitive is favored by EM theorists? As Clark (2008, p. 44) notes, the fact is that “[a]rguments in favour of [EM] appeal mainly, if not exclusively, to the *computational role* played by certain kinds of non-neural events and processes in online problem-solving.” In other words, EM theorists overwhelmingly conceive of cognition as a matter of *information processing*. Their distinctive observation is that, given this view of what cognition is, extraneuronal factors – including the stuff of material culture – may, in some cases anyway, realize the target phenomenon just as readily as neural tissue. For example, taking it that memory is, at least in part, a matter of the selective storage and context-sensitive retrieval of information, the EM theorist with a cognitive–archaeological bent might contend that information that is poised appropriately for context-sensitive retrieval may be stored in a Mycenaean Linear B tablet just as readily as in a Mycenaean brain.



### Extended functionalism and implementational materiality

Of course, not any old kind of information processing profile will do here. To say [otherwise] would be to fall prey to Rupert’s worry about explanatory inefficacy. No, genuine cognition will be found only in a (perhaps rather small) subset of information-processing systems. But, whatever the detailed story one eventually might arrive at via painstaking philosophical and empirical research, the basic commitment to cognition as a kind of information-processing is precisely what Malafouris thinks of as the kind of residual cognitivism that stymies progress towards understanding the rich way in which things may have cognitive lives. As Malafouris (2004, p. 55) puts it, “material culture has a place in the mind only as a disembodied digit of information written somehow on the neural tissue.” But, he argues, material culture has a far richer cognitive life than that image allows. Using striking examples from the contributions of artifacts within Mycenaean funeral rituals, he suggests that the role of “intentional or

AQ2





154 Michael Wheeler

non-intentional mnemo-technical artifacts and practices is far more dynamic and dialogical" and so artifact-involving memory reveals itself to be a "process of active discovery spanning the monumental and the minute, the conspicuous and the commonplace, iconicity and iconoclasm" (Malafouris, 2004, p. 57).

What is going on here? Although Malafouris doesn't put things in quite the way that I'm about to, his view, I think, is that those objects whose contribution to cognitive life is merely a matter of enabling enhanced performance in the storage and manipulation of information display only what we might call *implementational materiality*. In other words, given that what is important in understanding the mental status of such objects is that they perform certain computational functions that are allegedly distinctive of cognitive processing, the materiality of those objects is relevant only as the explanation of how the computational functions in question are *implemented in the physical world*. The real explanatory action is reserved for the more abstract, and in some sense nonmaterial (that is, disembodied), information-processing story.

But is EM's standard reliance on the concept of information the real issue? I don't think so. Let's allow the EM theorist to shed the grand theory of cognition as information-processing. Where might she go next? Well, she might preserve much of what is important about the computational approach without actually thinking of cognition as computation. In other words, she might declare herself to be a *noncomputational functionalist*. According to the traditional formulation of functionalism in the philosophy of mind, the canonical statement of which is arguably due to Putnam (1967), a mental state is constituted by the causal relations that it bears to sensory inputs, behavioral outputs, and other mental states. The information-processing theory of cognition, which glosses functional role as computational role, is just one possible version of functionalism, so conceived. As an alternative to computational functionalism one might, for example, gloss one kind of noncomputational functional role in terms of maintaining systemic stability in a homeostatic dynamic system by keeping the values of certain critical variables within certain limits. What this tells us is that the standard way of unpacking EM in information-processing terms is just one possible version of a position that Clark (2008) has dubbed *extended functionalism*. According to extended functionalism, a mental state is constituted by the causal relations that it bears to systemic inputs, systemic outputs, and other systemic states. (This is merely a more general formulation of the functionalist line, one that, in principle, allows the borders of



# PROOF

*Minds, Things, and Materiality* 155

the cognitive system to fall somewhere other than the sensory-motor interface of the organic body.) The very possibility of a noncomputational extended functionalism shows that information-processing (at least as construed in standard cognitive-scientific terms) cannot be at the center of the present issue. For all forms of extended functionalism will be just as committed to implementational materiality as the more narrowly defined information-processing variety.

To illustrate this point, consider the following example. Air traffic controllers typically coordinate their activity using flight strips – bands of paper printed with flight information (e.g., airline name, flight number, and type of aircraft, plus the speed, level, and route of the flight plan, both as requested and as authorized). When thinking about how air traffic controllers succeed in their complex, high-pressure job, and about how contemporary technology may enhance that success, one is inclined to focus, naturally enough, on the information carried by these strips. But this is not the only contribution of the strips. Mackay et al. (1998) argue convincingly that the physical embodiment of the strips supports a number of workplace strategies employed successfully by the controllers. For example, individuals often hold the strips as reminders to perform some action, or slide them to the left or right to indicate certain conditions (e.g., two planes in a potential conflict situation). Moreover, two controllers may work simultaneously on the same strip-holding board, using body language to signal the importance of particular movements or rearrangements of the strips.

From a practical perspective, this recognition of the noninformational contribution of the flight strips is far from idle. The testimonial evidence suggests that a number of previous attempts to introduce new computer technology into air traffic control may ultimately have been rejected as unworkable by the controllers precisely because the proposed replacement systems attempted to reproduce the straightforwardly informational aspects of the flight strips while ignoring the extra factors. Thus, and recognizing that the noninformational interactions supported by the flight strips would be difficult to reconstruct in any keyboard/monitor interface, Mackay et al. advocate the use of augmented electronic strips. From a theoretical perspective, and pending further analysis, one might at least consider the first two of the identified contributions of the flight strips (memory, the modeling of possible states of affairs) to put those artifacts in the ballpark to be considered elements in the cognitive architecture of the flight controller. And now note that nothing about this story undermines the extended functionalist line. The flight strips reveal their (provisionally) cognitive

156 Michael Wheeler

status precisely by carrying out certain functionally defined systemic roles that are distinctive of cognition (memory, the modeling of possible states of affairs). The materiality of the flight strips is thus implementational in character.

It is easy enough to see through to the theoretical heart of this issue. In order to fly, EM needs to embrace a key feature supported by functionalist theorizing – namely *multiple realizability*. A little philosophical history will help here. Functionalism (in its nonextended form) freed physicalist philosophy of mind from a kind of neural chauvinism. If our mental states were constituted by their functional roles, and the material contribution of our brains were merely implementational in character, then robots, Martians, Klingons, and gaseous creatures from the outer limits of the universe could all join us in having mental states, just so long as the physical stuff out of which they were made could implement the right functional profiles. Stretching the word “skin” to include boundaries made of tin and gas, traditional functionalism bequeathed to the mind what we might call within-the-skin multiple realizability. And within-the-skin multiple realizability requires within-the-skin implementational materiality. But now extended functionalism merely plays out the same logic beyond the skin. If the specific materiality of the substrate doesn’t matter to cognition, outside of the fact that it must be able to support the required functional profile, then what, in principle, is there to stop things-beyond-the-skin counting as proper parts of a cognitive architecture? Nothing, that’s what. And this beyond-the-skin species of multiple realizability, which is just another way of characterizing the core philosophical commitment of EM, requires beyond-the-skin implementational materiality. If we look at things this way, the really radical and revolutionary movement was functionalism, not EM. EM simply makes manifest one of the implications of functionalism. In other words, EM is just a footnote to Putnam.

### Enactivism and vital materiality

One could, I suppose, develop a version of EM (as standardly conceived) without buying into functionalism, if one could have multiple realizability without functionalism. Churchland (2005) has argued recently that the latter is possible. But, even if one could have EM without functionalism, one couldn’t have EM (as standardly conceived) without multiple realizability, and so one couldn’t have EM (as standardly conceived) without implementational materiality. Given that Malafouris finds implementational materiality wanting as the basis of a full account of

# PROOF

*Minds, Things, and Materiality* 157

the cognitive life of material culture, this goes some way towards breaking the link between EM and what he calls “taking material culture seriously.” Now recall that taking material culture seriously means to be “systematically concerned with figuring out the causal efficacy of materiality in the enactment and constitution of a cognitive system or operation” (Malafouris, 2004, p. 55). So what is this other species of material causal efficacy, the one that, according to Malafouris, goes beyond mere implementational materiality, and which matters to the cognitive life of things? Let’s label this form of material causal efficacy “vital materiality.” What, then, is vital materiality? To answer this question we need to turn to Malafouris’s compelling image of the potter at her wheel.

[The] cognitive map of knowledge and memory may well be extended and distributed in the neurons of the potter’s brain, the muscles of the potter’s body, the “affordances”... of the potter’s wheel, the material properties of the clay, the morphological and typological prototypes of existing vessels as well as the general social context in which the activity occurs. (Malafouris, 2004, p. 59)

So far, nothing new seems to be on the table – or rather the wheel. Indeed, it looks as if what we have here is EM. It might seem, therefore, that the lesson of the potter’s wheel is that we should replace our view of cognition as residing inside the potter’s head with that of cognition as spatially distributed over brain, body, and world. But Malafouris develops his case in a different direction:

It is at the potter’s fingers that the form and shape of the vessel is perceived as it gradually emerges in the interactive tension between the centrifugal force and the texture of the wet clay. Materiality enters the cognitive equation at a much more basic level, shaping the phenomenology of what Searle has defined as the “Background” i.e., the set of non-representational mental capacities that enable all representing to take place...In other words...we should replace our view of cognition as residing inside the potter’s head, with that of cognition enacted *at the potter’s wheel*. (Malafouris, 2004, p. 59)

Clearly, we need to understand the term “enacted.” As the references given by Malafouris (2004, p. 57) indicate, the relevant interpretation of this idea is provided by Varela et al.’s enactive approach to cognition (Varela et al., 1991; see also Thompson, 2007), in which that approach is

# PROOF



158 Michael Wheeler

seen as a development of the autopoietic theory of Maturana and Varela (1980). And there's the rub. For, as far as I can see, on that interpretation of the pivotal idea, cognition enacted *cannot* be cognition extended. Here's why.

The smart money says that the theory of autopoiesis (an all-too-brief introduction to which will be given in a moment) is a nonnegotiable component of enactivism. This might come as a surprise to some fans of the enactivist "bible" *The Embodied Mind* (Varela et al., 1991) since that text doesn't foreground autopoiesis as a term, but the fact is that many of the component concepts from autopoietic theory, such as "autonomy" and "structural coupling," all appear there in pivotal roles, and are used in ways identifiable from the theory of autopoiesis. More recently, Thompson, in a book that I believe is poised to become the leading post-*Embodied-Mind* development of enactivism, explicitly treats the theory of autopoiesis as one of the conceptual keystones of the position (Thompson, 2007, Ch. 5). This suggests that getting straight about autopoiesis will help us to understand the enactivist's core commitments. The bad news for the uninitiated is that the theory of autopoiesis is so devilishly complex that I can't hope to do it justice in the space available here. The good news is that I don't need to worry about this exegetical shortfall since I'll be concerned with the overall shape of things, not the fine-grained details. Under such circumstances, the roughest of rough guides will do.

Autopoietic systems form a subset of *self-organizing* systems (Di Paolo, 2005; Weber and Varela, 2002), where a self-organizing system is one in which the intrasystemic components, on the basis of purely local rules, interact with each other in nonlinear ways to produce the emergence and maintenance of structured global order. More specifically, autopoietic systems form a subset of *autonomous* self-organizing systems, where an autonomous system is one in which the constituent processes "(i) recursively depend on each other for their generation and their realization as a network, (ii) constitute the system as a unity in whatever domain they exist, and (iii) determine a domain of possible interactions with the environment" (Thompson, 2007, p. 44; citing Varela, 1979, p. 55). Examples of autonomous self-organizing systems include cells, nervous systems, and insect colonies. However, while self-organization and autonomy are necessary for autopoiesis, they are not sufficient. To be autopoietic, a self-organizing autonomous system must also, through its own endogenous activity, produce and maintain a *physical* boundary that distinguishes that system as a material unity (a minimal self) in the space in which it exists. The paradigm example of



# PROOF

Minds, Things, and Materiality 159

an autopoietic system is the single cell (a network of chemical reactions that produces its own membrane), and one might reasonably think of the concept of autopoiesis as capturing the distinctive mode of systemic organization that is realized in material living systems as metabolism (Di Paolo, 2005). Crucially, the self-distinguishing process characteristic of autopoiesis takes place in the face of physical perturbations from the system's environment. This is what autopoietic theorists call *structural coupling*, a process in which an autopoietic system encounters its environment while maintaining its organization.

Now for the point of all this. According to Maturana and Varela, autopoiesis is necessary and sufficient for life (see, e.g., Maturana and Varela, 1980, p. 82). In other words, for Maturana and Varela, any living system is an autopoietic system and any autopoietic system is a living system. But what about cognition? One striking claim in the core literature is that living "simply" is cognition: "[l]iving systems are cognitive systems, and living as a process is a process of cognition" (Maturana, 1970, p. 13). How could this be? For Maturana, cognition is effective (i.e., viability-maintaining) activity in a domain of interactions defined by the autopoietic system's organization. In this context, then, *enaction* is the process by which significance or relevance is brought forth through the viable structural coupling of the autopoietic system with its environment. The idea here is that the autopoietic organization, in establishing the distinction between the self-maintenance and the collapse of the system as a material unity, institutes a norm of survival, and thereby the significance or relevance of certain environmental perturbations as either leaving the system organizationally intact or resulting in its disintegration. (Notice that the tight link between cognition and autopoiesis plausibly introduces something that we might be moved to call "vital materiality" in the case of the former, given the stress in the latter on the endogenous construction and maintenance of a self-distinguishing *material* boundary. However, as I shall explain later, I don't think it is this idea of vital materiality that Malafouris has in mind.)

Of course, as Thompson (2004) notes, it is possible to hear Maturana's striking claim in more than one way – as advancing either (a) the view that life and cognition are identical, or (b) the view that living systems are a subset of cognitive systems. For the moment let's work with (a). I'll come back to (b) later. If the living system is identical with the cognitive system, then the boundary of the living system will coincide with the boundary of the cognitive system. And it's this that (finally) generates the inconsistency with EM, given, that is, that we accept what I take to be a *highly* plausible claim, namely that, where we have an extended

160 Michael Wheeler

cognitive system, the living system (the organism, although not necessarily the phenotype) remains bounded by its skin. The external elements that, according to EM, count as proper parts of the cognitive system do not thereby become proper parts of the living organism. They remain abiotic. In other words, and *pace* Scott Turner (2000), organisms don't extend even if minds and, as Dawkins (1982) has pointed out, phenotypes do. But if, in cases of extended cognition, the boundary of the cognitive system and the boundary of the living system come apart, as I have suggested they do, then one cannot simultaneously sanction EM and identify the cognitive system with the living system. Since autopoietic theory and (therefore) enactivism are committed to identifying the cognitive system with the living system, the enactivist simply cannot endorse EM. In other words, enacted minds are not extended minds.

So far so bad for the attempt to bring EM and enactivism together. But the game is not yet over, because a more complex understanding of the relationship between life and cognition has recently entered the autopoiesis-enactivist literature (Di Paolo, 2005; see also Thompson, 2007). In his later work, Varela (e.g., 1991) argued that cognition is a process of (what he dubbed) *sense-making* and that living is just such a process. Sense-making enriches the autopoietic picture of an enacted domain of significance by introducing the idea that that domain is "a place of *valence*, of attraction and repulsion, approach or escape" (Thompson, 2004, p. 386). Varela (1991) and Thompson (2004) both suggest that autopoiesis is sufficient to instantiate sense-making (hence living is a process of sense-making). However, Di Paolo (2005) has argued – convincingly, as Thompson (2007) notes – that sense-making requires something over and above what we might now think of as *raw* autopoiesis. This is because raw autopoiesis bestows only a kind of robustness or conservation of systemic organization in the face of environmental perturbation. The system doesn't alter its behavior in response to changes in its environment. It either survives any perturbations it experiences, or it doesn't. As Di Paolo usefully puts it, the norm of survival as established by raw autopoiesis is an all-or-nothing affair. But sense-making (as Thompson's talk of attraction, repulsion, approach, and escape indicates) requires a system to be sensitive to graded differences between states. The organism needs to monitor *how it is doing* with respect to the norm of survival and to regulate its behavior accordingly in order to improve its situation. In other words, it needs to be an *adaptive* system.

Crucially, while adaptivity is a phenomenon over and above raw autopoiesis, it needs to be established on the basis of an autopoietic organization. Without the connection to the self-distinguishing process



of autopoiesis, the meaning generated by adaptivity and sense-making would not be established as original to the activity of the system (that is, it would not be meaning *for* the system), but would merely be attributable to the system by some external observer (Di Paolo, 2005; Thompson, 2007). So, given that raw autopoiesis is necessary but not sufficient for sense-making, and given that sense-making is to be identified with cognition, raw autopoiesis is necessary but not sufficient for cognition, which means, of course, that cognition is sufficient for raw autopoiesis. But now, since on our revised picture raw autopoiesis remains necessary and sufficient for life – Di Paolo never suggests that a raw autopoietic system wouldn't be alive, only that it wouldn't be a cognitive system – being a cognitive system is sufficient for being a living system. And that means that the tension between enactivism and EM is still in force. If the enactivist did endorse EM, she would (among other things) be claiming (a) that an extended cognitive system is an autopoietic system, and thus (b) that an extended cognitive system is itself (and does not merely contain) a living system. As far as I can see, (a) is debatable, but in any case (b) violates our highly plausible thought concerning organismic extension (or rather the lack of it). Since the enactivist cannot give up on the colocation of the cognitive system and the living system, she cannot endorse EM. To repeat, enacted minds are not extended minds.

There is, then, a *prima facie* case for the incompatibility of EM and enactivism. There are, of course, further moves that could be made. For example, one might try to unpack a genuinely enactivist notion of cognition in which there is no (explicit or implicit) *necessary* dependence on the phenomenon of autopoiesis, thereby creating proper conceptual space for the previously mentioned view that living systems form a subset of cognitive systems. On this account, being an autopoietic (living) system would not be necessary for a system to be cognitive. That would allow there to be a cognitive system that is not itself identical with some living system. Consistency with EM would thereby be established, since one could have an extended cognitive system linked to a nonextended living system. The plausibility of such an approach cannot be ruled out in advance. At the very least, however, I hope to have gestured in the direction of the sorts of sacrifices and challenges that the project of developing such a view would face, given the autopoietic heritage of enactivism as standardly conceived.

Aside from the parting of the embodied-embedded ways that I have identified (and note that I've said nothing about which of our two views might be correct), it also remains unclear, from what I've said about



162 Michael Wheeler

enactivism, exactly why that approach might be the road to imbuing material culture with something that Malafouris would welcome as vital materiality. I shall finish with a sketchy remark on this point. The key here is in Malafouris's claim that the material nature of the interactions between the potter's fingers, the wheel, and the clay means that materiality shapes the fundamental phenomenological structure of the situation. There is no doubt that this is broadly in line with enactivist theorizing, which holds experience to be bodily in character. Let's say that this is right, and that in the end it's the way in which the materiality of things, alongside our own embodiment, structures our fundamental phenomenological space, that gives us a concept of the vital materiality of material culture. Without further argument, it is at least uncertain that this phenomenological point *alone* could ever mandate the claim that things-beyond-the-skin *must* be counted parts of the cognitive system, as opposed to mind-external elements that play a causal role alongside mind-internal elements in structuring experience. More, it seems, remains to be said.

This chapter is reproduced with permission from the McDonald Institute for Archaeological Research: Wheeler, M. (2010). Minds, Things, and Materiality. In Renfrew C., and Malafouris L. (Eds), *The Cognitive Life of Things: Recasting the Boundaries of the Mind*. Cambridge: McDonald Institute for Archaeological Research Publications, pp. 29–37.

## References

- Churchland, P. M. (2005). Functionalism at forty: a critical retrospective. *Journal of Philosophy*, 102, no. 1, 33–50.
- Clark, A. (2008). Pressing the flesh: a tension in the study of the embodied, embedded mind? *Philosophy and Phenomenological Research*, 76, note 1, 37–59.
- Clark, A., and Chalmers, D. (1998). The extended mind. *Analysis*, 58, no. 1, 7–19.
- Dawkins, R. (1982). *The Extended Phenotype: the Long Reach of the Gene*. Oxford: Oxford University Press.
- Di Paolo, E. A. (2005). Autopoiesis, adaptivity, teleology, agency. *Phenomenology and the Cognitive Sciences*, 4, no. 4, 429–52.
- Mackay, W. E., Fayard, A.-L., Frobert, L., and Medini, L. (1998). Reinventing the familiar: exploring an augmented reality design space for air traffic control. In *Conference Proceedings on Human Factors in Computing Systems (CHI 1998)*. New York: ACM Press / Addison-Wesley, pp. 558–65.
- Malafouris, L. (2004). The cognitive basis of material engagement: where brain, body and culture conflate. In E. DeMarrais, C. Gosden, and C. Renfrew (Eds), *Rethinking Materiality: the Engagement of Mind with the Material World*. Cambridge: McDonald Institute for Archaeological Research, pp. 53–61.

# PROOF



Minds, Things, and Materiality 163

- Maturana, H. R. (1970). Biology of cognition. In Maturana and Varela, 1980, pp. 2–58.
- Maturana, H. R., and Varela, F. J. (1980). *Autopoiesis and Cognition: the Realization of the Living*, Vol. 43. Dordrecht: Boston Studies in the Philosophy of Science, D. Reidel.
- Menary, R. (Ed.) (Forthcoming). *The Extended Mind*. Aldershot: Ashgate.
- Putnam, H. (1967). Psychological predicates. In W. H. Capitan, and D. D. Merrill (Eds), *Art, Mind and Religion*. Pittsburgh: University of Pittsburgh Press.
- Rumelhart, D. E., Smolensky, P., McClelland, J. L., and Hinton, G. (1986). Schemata and sequential thought processes in PDP models. In J. L. McClelland, and D. Rumelhart (Eds). *Parallel Distributed Processing: Explorations In The Microstructure Of Cognition, Vol. 2: Psychological And Biological Models*. Cambridge, MA: MIT Press, pp. 7–57.
- Rupert, R. (2004). Challenges to the hypothesis of extended cognition. *Journal of Philosophy*, 101, no. 8, 389–428.
- Scott Turner, J. (2000). *The Extended Organism: the Physiology of Animal-Built Structures*. Cambridge, MA: Harvard University Press.
- Thompson, E. (2004). Life and mind: from autopoiesis to neurophenomenology. A tribute to Francisco Varela. *Phenomenology and the Cognitive Sciences*, 3, 381–98.
- Thompson, E. (2007). *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*. Cambridge, MA: Harvard University Press.
- Varela, F. J. (1979). *Principles of Biological Autonomy*. New York: Elsevier North Holland.
- Varela, F. J. (1991). Organism: a meshwork of selfless selves. In A. Tauber (Ed.), *Organism and the Origin of Self*. Dordrecht: Kluwer Academic Publishers, pp. 79–107.
- Varela, F. J., Thompson, E., and Rosch, E. (1991). *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MA: MIT Press.
- Weber, A., and Varela, F. J. (2002). Life after Kant: natural purposes and the autopoietic foundations of biological individuality. *Phenomenology and the Cognitive Sciences*, 1, 97–125.
- Wheeler, M. (2005). *Reconstructing the Cognitive World: the Next Step*. Cambridge, MA: MIT Press.





## 7

## Contributions of Mirror Mechanisms to the Embodiment of Cognition

*Arthur M. Glenberg*

### Introduction: historical reflection

Mirror neurons in the macaque are equally active when the animal initiates an action and when the animal observes another take the same or similar action. These neurons may be used in goal recognition (I recognize the goal of your action through resonance of that part of my motor system used by me to accomplish a particular goal), inferring intent, empathy, and, in general, greasing the wheels of social interaction. I describe in this chapter a new methodology for investigating mirror mechanism function: because some mirror neurons are part of the motor system they can be adapted through repeated action; and, because mirror neurons are multimodal, the effect of that adaptation can be shown in other cognitive tasks, thereby demonstrating the causal effect of the mirror mechanism on that other cognitive task.

The editor of this volume asked me to include some historical and human context for the research reported in the main body of this chapter. The next few pages do just that.

My personal history with the concept of embodiment is convoluted. What follows is a reconstruction from memory, but it is a story told so often that I cannot verify whether I am remembering events or remembering the tellings. In about 1990 or 1992, I ran into Larry Barsalou at the Art Institute in Chicago; we were both playing hokey from a meeting of the Midwestern Psychological Association. Larry related to me his doubts regarding empirical support for the idea that propositions, formed of abstract symbols, are the basis of cognition. I didn't really



# PROOF

*Contributions of Mirror Mechanisms to the Embodiment of Cognition* 165

understand him, but it did get me thinking a bit about the issue (see "Not propositions," Glenberg et al., 1999).

Not long after my conversation with Larry, I was invited to spend time at the Center for Interdisciplinary Research in Bielefeld, Germany, as part of a group studying mental models. As I was boarding the plane, Will Langston, a graduate student at the time, ran down the tarmac to give me a copy of a book to read on the plane (the tarmac part is probably a confabulation, the Langston part is not). The book was Lakoff's "Women, Fire, and Dangerous Things," and it astonished me. Here was a way of dealing with Barsalou's problem!

Fortunately for me, Tony Sanford was also a member of the group, and together we gave Lakoff a close read. I was much more convinced than Tony, but I still had some nagging questions. Could Lakoff's image schemas and conceptual metaphors really provide the grounding for cognition? I spent six months looking for a satisfying answer, and the result (with a lot of help later from Morton Gernsbacher, Bill Epstein, and others) was my paper published in *Brain and Behavioral Science* ["What memory is for"] (Glenberg, 1997). I proposed that cognition is grounded in action.

AQ1

Working with then graduate students David Robertson and Mike Kaschak, we tried to figure out how to test this claim. One outcome of this productive collaboration was a much clearer framing of the idea in the form of the Indexical Hypothesis (Glenberg and Robertson, 1999; 2000). A second outcome was our discovery of the ACE effect (Glenberg and Kaschak, 2002), which provided some of the most convincing evidence (at that time) of the relation between language and action.

AQ2

Around the same time, Larry Barsalou was encouraging me to pay closer attention to the cognitive neuroscience literature, which was a task that seemed overwhelming. But, because I had already learned to follow Larry's lead, I was primed to jump at an opportunity. That opportunity came when I was invited to participate in a symposium on Language and Action sponsored by the Max Planck Institute for Cognition (then in Munich, currently in Leipzig). One of the speakers was Giacomo Rizzolatti, and he talked about his research on mirror neurons. I was astonished again; Rizzolatti solved two problems for me. The first problem was the underlying neurophysiological and evolutionary basis for grounding language in action (see Rizzolatti and Arbib, 1999). The second problem was how to satisfy my wife's demand that I could spend an upcoming sabbatical anywhere, as long as it was in Italy.

AQ3

That sabbatical, half at the Max Planck in Leipzig (without the family) and half in Parma, Italy, was one of the highlights of my academic

166 Arthur M. Glenberg

life: I learned more than I thought possible, I formed lasting friendships and collaborations (e.g., Glenberg et al., 2008a; Rueschemeyer et al., in press), and my interest in mirror neurons grew.

Sato and Cattaneo were post-docs in the Department of Neuroscience at Parma. We completed some work using transcranial magnetic stimulation (TMS) to test the Indexical Hypothesis and the underlying neurophysiology of the ACE effect (Glenberg et al., 2008b). After we completed those experiments, the TMS equipment was unavailable, and I was returning to Wisconsin, so it wasn't clear how we were to proceed with other planned research. At one brainstorming session, Cattaneo proposed what would become the motor adaptation paradigm that generated the results reported below. On the strength of that idea, I could not begrudge Sato and Cattaneo the enormous amount they drank at my going-away party.

### Contributions of a mirror mechanism to embodiment

The defining feature of mirror neurons is that they are active both when an animal is engaged in a task and when the animal observes another engaged in the same or related task (Rizzolatti and Craighero, 2004). Thus, mirror neurons may indicate how motor resonance can be used to recognize the actions and intents of others. That is, when animal A observes animal B acting, animal A's mirror neuron system resonates to the extent that A has learned the action. This resonance provides the goal that A has in doing the action (see below), and this goal is attributed to B. If this hypothesis is correct, then mirror neurons are likely to play an important role in greasing the wheels of social interaction and cooperation. A human mirror neuron system, or mirror mechanism (MM), may play a similar role in social interaction (Gallese et al., 2004) and contribute to theory of mind (Gallese and Goldman, 1998) and language processes (Glenberg et al., 2008a; Rizzolatti and Craighero, 2007), and mirror neuron dysfunction may contribute to the symptoms of autism spectrum disorder (Dapretto et al., 2006).

What does a MM have to do with embodied cognition? One way of framing the embodiment hypothesis is to say that all cognition is grounded in, or makes use of, neural and bodily systems of perception, action, and emotion (see Glenberg, 2010). For example, Rueschemeyer et al. (in press; see also Saygin et al., 2010) used fMRI to demonstrate that comprehending language about visual motion often activates just those areas of cortex (MT/V5) used during the perception of visual motion. Similarly, Havas and colleagues (2007; 2010) showed that preventing



people from expressing emotions on the face slowed understanding of sentences describing emotional situations. And, most importantly for the current argument, Glenberg and colleagues (2002; 2008; see also Zwaan and Taylor, 2006) showed that language comprehension engages neural systems used in the control of action.

AQ4

In this context, a human MM provides the neurophysiological underpinning for the relation between cognition and action. It is not just a matter of being able to point to the brain for supporting evidence. Instead, the association of a human MM and cognition provides a fuller understanding of both the neurophysiology and the cognition. Furthermore, the association suggests new ways of testing both the embodiment hypothesis and the function of a MM. Finally, making the connection between embodiment and MM helps to move psychology toward biology (Glenberg, 2006), and, at the same time, the association helps to bridge the gap between mental processing and brain activity.

The remainder of this chapter is divided into three sections. The first presents a brief review of the literature on mirror neurons and mirror mechanisms in humans. This review highlights some of the promise of MM theory and some of the reasons for skepticism. One of the reasons for skepticism is the reliance on methods that are difficult to interpret. Next, I describe how the methodology can be used to investigate motor contributions to speech perception, and this is followed by a review of work using the methodology to investigate language comprehension. The results of these experiments demonstrate how cognition is thoroughly embodied: motor processes play an important role even in cognitive activity previously believed to be based on abstract processes.



### Mirror neurons and mirror mechanisms

Identification of mirror neurons in the macaque is straightforward (although not simple): using single cell recording, find neurons that respond both during execution of a particular action and upon observation of the same action. Indeed, mirror neurons have been found in the macaque premotor cortex (area F5 in particular, which is a likely homolog of the human Broca's area), inferior parietal lobule, and other areas. The predominate method for investigating the human MM is functional imaging. Unfortunately, this technique is limited in that a) it is correlational and thus difficult to use in the study of MM function, b) it depends on the inference that spatial overlap of fMRI blood-oxygen-level dependent (BOLD) signals indicates activity in the same neurons, and c) it is expensive and inappropriate for some populations (e.g.,



168 *Arthur M. Glenberg*

children) and some investigations (e.g., when movement is required). Because of these problems, some researchers remain skeptical about claims for a human MM (e.g., Hickock, 2009).

Nonetheless, the possibility that a human mirror mechanism (MM) serves action recognition, empathy, theory of mind, and language has generated tremendous excitement. For example, Ramachandran (2005) suggested that mirror neurons will “do for psychology what DNA did for biology.” As just noted, however, there is growing skepticism about the data supporting a human MM and MM function. I begin with a brief review of the data producing the excitement, and then turn to the reasons for caution.

Investigation of area F5 of the macaque premotor cortex produced the first data (or, at least, the first data associated with the name “mirror neurons”) regarding mirror neurons and the theory of action recognition (e.g., Gallese et al., 1996). The signature finding, based on single-cell recording, is that the same motor neuron fires with essentially equal frequency during production of an action and during perception of the same action (about one-third of mirror neurons) or similar actions, that is, actions with the same goal (about two-thirds of mirror neurons). Later work demonstrated that about 17 per cent of mirror neurons have auditory, as well as visual and motor, properties; that is, these neurons also respond to the sounds produced by the actions (Kohler et al., 2002).

Several experiments with macaques have attempted to show the role of mirror neurons in action recognition. Umiltà et al. (2001) demonstrated that mirror neurons are active when action on an object is completed behind a barrier, but not when the animal knows that there is no object behind the barrier. Furthermore, Umiltà et al. (2008) demonstrated that the same neuron is active when pliers operated by the animal close on a goal (e.g., a peanut), and the neuron is active regardless of whether the pliers operate by a hand-closing or a hand-opening movement. That is, it is the goal that counts, not the specific movements for achieving the goal.

Research on a potential human MM is more inferential, given the relative scarcity of single-cell recording (but see Mukamel et al., 2010). Thus, using fMRI, the operation of a MM is inferred when activity in a particular cortical area (as reflected in the BOLD signal) is similar during action recognition and action production, and when the cortical area is a likely homolog of an area in macaque cortex in which mirror neurons have been identified. Using this or a similar logic, it appears that the human MM responds more to actions that the perceiver can perform than to actions with which the perceiver is familiar through



vision alone (e.g., Calvo-Merino et al., 2006); the MM is more sensitive to social actions than similar nonsocial actions (Kilner et al., 2006); the MM responds to both visually perceived actions and the linguistic description of actions (Tettamanti et al., 2005); the MM plays a role in speech perception (D'Ausilio et al., 2009); activity in the MM is positively correlated with empathy (Gazzola et al., 2006) and negatively correlated with autistic behaviors (Dapretto et al., 2006).

Neuropsychological data have also played a role in investigating the function of a human MM. For example, Pazzaglia et al. (2008) demonstrated that patients with lesions that produce buccofacial apraxia have difficulty recognizing the meaning of mouth-produced sounds, but not hand-produced sounds, whereas just the opposite is found for patients with lesions that produce limb apraxia. Fazio et al. (2009) demonstrated that frontal aphasic patients had difficulty reordering pictures taken from a video of hierarchical human actions (e.g., opening a door), but not pictures taken from videos of equally complex physical events (e.g., a bicycle falling over). That is, lesions in Broca's area, the human homolog of macaque area F5, affect human-action recognition when there is no verbal content in the task.

Even with this evidence supporting the role of a human MM in action recognition, there are reasons to be skeptical. Dinstein et al. (2008) argue that the data used to support the claim of a MM, data based predominately on fMRI, are not up to the task. Namely, fMRI data cannot clearly demonstrate that the same neural populations are contributing to both action production and action recognition. Note that the same criticism applies to the many studies using transcranial magnetic stimulation (TMS), such as D'Ausilio et al. (2009). That is, TMS will disrupt neural processing at a particular location, but it is unclear whether the disruption is due to disruption of multimodal mirror neurons or to disruption of unimodal neurons in the same location.

Dinstein et al. (2008) suggest a solution for fMRI investigations: adaptation paradigms. Repeated presentation of an action stimulus for recognition adapts the BOLD signal. Then, will this adaptation due to recognition produce a reduced BOLD signal during action production? If so, that would provide relatively strong data in support of a MM that is active during both action recognition and production. Unfortunately, several experiments reviewed by Dinstein et al. failed to find the cross-task adaptation. A more recent study, Chong et al. (2008), does demonstrate cross-task adaptation in the BOLD signal. However, as discussed in Dinstein et al. (2008), the location of the adapted BOLD signal was unexpected on the basis of previous work putatively supporting a



170 Arthur M. Glenberg

human MM. Also, as discussed in Lingnau et al. (2009), the asymmetric nature of the adaptation is worrisome. In contrast to these less-than-convincing results, Kilner et al. (2009) demonstrated just the expected symmetric, cross-modal adaptation in an area of the brain (inferior frontal gyrus) associated with MM activity by other research. Hence, the fMRI adaptation paradigm may yet prove useful.

Nonetheless, there are several other problems with the fMRI protocols. First, most are inherently correlational. Thus, even demonstrating similar BOLD signals in the same cortical area during action recognition and production does not demonstrate a causal relation. Second, fMRI is expensive. Third, fMRI is inappropriate for some populations, for example children, and often inappropriate when action is required.

Hickock (2009) raises other serious issues. On his analysis of the data, there are eight problems with the mirror neuron theory of action recognition. Among these is that the data from monkeys are limited in their ability to demonstrate a causal relation between mirror neuron activity and action understanding because there have been no lesion studies. Hickock suggests that the extant data (notably, Umiltà et al., 2001; 2008) can be explained along the lines of motor priming. Similar reservations are raised by Lingnau et al. (2009) and Mahon (2008).

One conclusion is obvious: if MM data are ever to match MM excitement, there is a dire need for new methods. A good method should be capable of a) demonstrating causal relations; b) investigating action understanding as well as the myriad of other claims made for the MM; c) capable of being used with a large variety of populations, including children, neuropsychological patients, and autistics. And it would certainly be a bonus if the method were cheap and hence widely available to the scientific community. A new behavioral adaptation procedure has the potential to meet these requirements.

### Using behavioral motor adaptation to investigate MMs

The behavioral method takes advantage of a general property of cortical neurons, namely, that they adapt with use, along with three properties of mirror neurons: a) they are multimodal, sensorimotor neurons, b) they are specific to particular actions, and c) they differentially couple to biological stimuli. The procedure begins with a repetitive motor task that can be executed in several ways (e.g., repeatedly moving the fingers to cup the hand or to open the hand, or repeatedly moving objects from location A to location B, or from location B to location A). The point of the task is to differentially adapt the motor system

# PROOF

*Contributions of Mirror Mechanisms to the Embodiment of Cognition* 171

(for physiological evidence on use-induced plasticity, see Classen et al., 1998). Next, the participant engages in a cognitive task that is putatively unrelated to the motor adaptation, such as action recognition or language comprehension. The experimental question is whether the cognitive task is affected by the differentially adapted motor system.

AQ5 That is, does adapting the motor system in one way affect the cognitive task, and does adapting the motor system in the opposite way produce AQ6 a different affect on the cognitive task? If so, then we have evidence of a MM that causally affects cognition.

The first experiment is meant to secure this conclusion by ruling out alternatives, as explained below. In the task, participants moved beans, one at a time, from a wide-mouth container to a narrow-mouth container an arm's length away (Figure 7.1). Half of the participants moved the beans in a direction Away from the self, and half moved the beans in a direction Toward the self. Note that very similar muscles are used in the two conditions. Thus the actions are defined in terms of target location and specifics of the movement (e.g., a precision grip is used in the Away movement but not the return toward the body, and vice versa in the Toward movement).

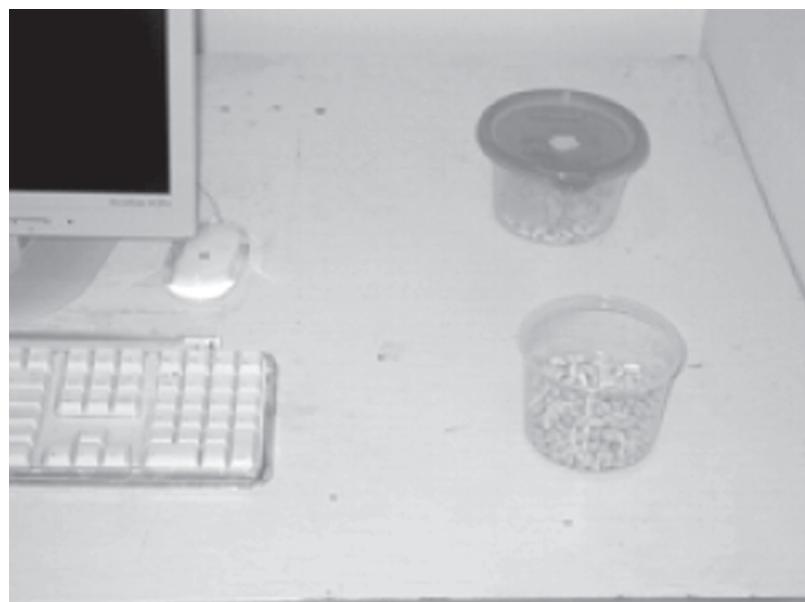


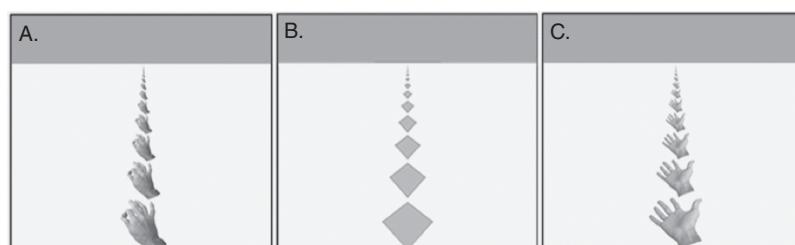
Figure 7.1 Containers illustrating the Away condition.

In the experiment, there were eight groups of participants. Participants in half of the groups moved the beans in the Toward direction and participants in half of the groups moved the beans in the Away direction. Orthogonally, the number of beans moved was 15, 45, 135, or 405. Moving 405 beans takes about 10–15 minutes and leads to measurable use-induced plasticity (e.g., Classen et al., 1998). The bean movements were done while the participant was blindfolded. The reason for the blindfold was to better secure the claim that any affects of the adaptation were due to motor adaptation rather than adaptation in the visual system.

Immediately after moving the beans (and removing the blindfold), participants engaged in a psychophysical task to measure bias in the perception of movement toward and away (Lewis and McBeath, 2004). Participants viewed on a computer monitor a screen tiled with stimuli that appear to be receding into the distance (Figure 7.2).

A second screen was presented one second later with the tiles slightly shifted, which induces an illusion of motion. A shift of 50 per cent of the inter-tile distance is completely ambiguous, whereas a shift of about 20 per cent or 80 per cent is perceived as movement: small percentage shifts lead to an illusion of movement away from the observer whereas large percentage shifts lead to an illusion of movement toward the observer. We used the MOBS procedure (Tyrrell and Owens, 1988) to adjust the percentage shift from trial to trial to locate the shift threshold. We defined this shift threshold as the percentage of the inter-tile distance at which people can no longer consistently judge the direction of movement as toward or away.

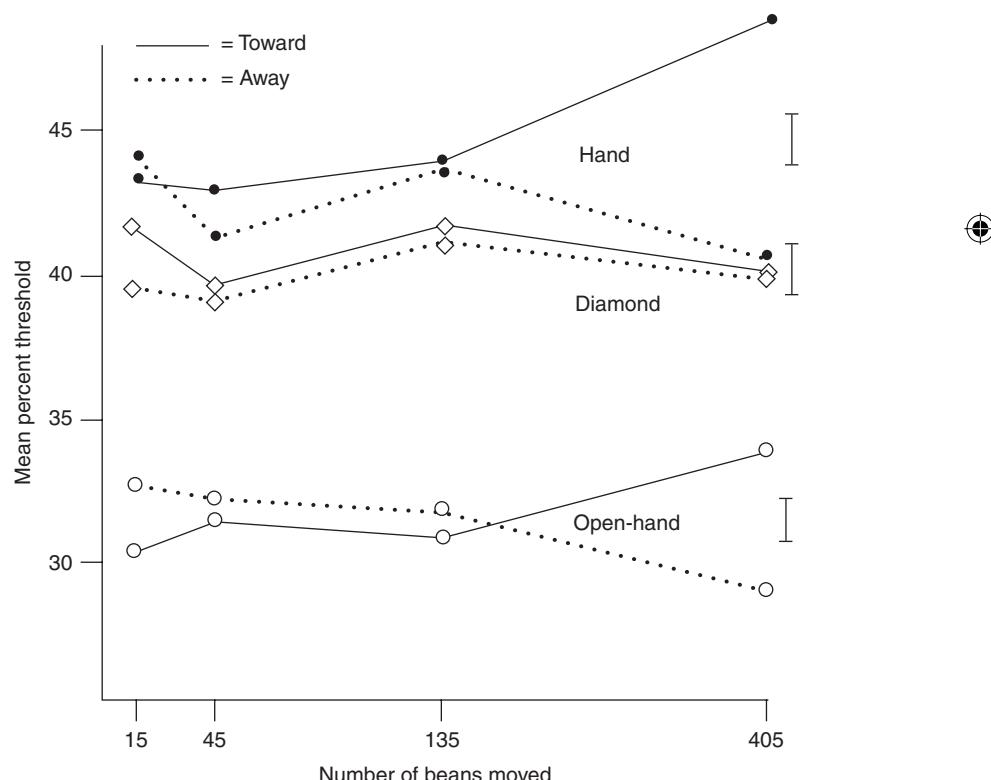
The threshold was measured for three stimuli, a hand holding a bean (the bean-hand), an open hand with palm up, and a diamond (see Figure 7.2). The type of stimulus was manipulated within-subjects, and



*Figure 7.2* The three tiles used in the action recognition threshold task.

the measures for the three stimuli were interleaved so that a participant might see a diamond shift, a hand holding a bean shift, and then an open hand shift. The point of the procedure is to determine whether the direction and number of bean movements affect the direction and size of visual threshold shifts.

The data are presented in Figure 7.3. As the number of Toward and Away bean movements increased, there was an increasingly large difference between the thresholds. This interaction effect was modulated by the biological verisimilitude of the visual stimulus to the hand that made the movements. That is, the interaction was largest for the Hand stimulus, next largest for the Open Hand stimulus, and not evident for the Diamond stimulus.



*Figure 7.3* Changes in the recognition thresholds as a function of number of beans moved, direction of movement, and visual stimulus tile.

# PROOF

174 Arthur M. Glenberg

The pattern of data is strongly consistent with adaptation of a MM. First, the adaptation is cross-modal: blindfolded motor adaptation affects visual perception. Second, the effect is direction-specific: bean movement in the Toward direction changes the threshold so that ambiguous movement is less likely to be perceived in the Toward direction (the threshold moves so that more “toward” information is needed to perceive movement in that direction), and vice versa for Away bean movement. Third, as expected from an adaptation mechanism, but contrary to interpretations based on task demands or expectancies, the perceptual aftereffect appears only after prolonged exposure to the adapting procedure, that is, only after moving 405 beans. Fourth, the effect is stimulus-specific; that is, it is strongest for perception of the biological stimulus most similar to the adapted action (the hand holding the bean), weaker for a related stimulus (the open hand), and not evident for the nonbiological stimulus (the diamond).

AQ8 The results are particularly noteworthy given a recent criticism of (fMRI) adaptation techniques in Lingnau et al. (2009). They suggest two conditions that would produce a convincing demonstration that a human mirror neuron system plays a role in action recognition. The first is that a common set of act-specific neurons must be activated by observation and execution. The second is that there is evidence for direct activation rather than priming based on inference. Evidence for this second condition could come, they suggest, from demonstrating adaptation of action recognition produced by prior action execution. It is this second condition that has been difficult to find in fMRI work (e.g., Chong et al., 2008; Lingnau et al., 2009; but see Kilner et al., 2009). Note that the data presented in Figure 7.3 meet both of these conditions. Act specificity is demonstrated by the differential effect of Toward and Away movement. The second condition is met because the action task preceded the measurement of change in action recognition.

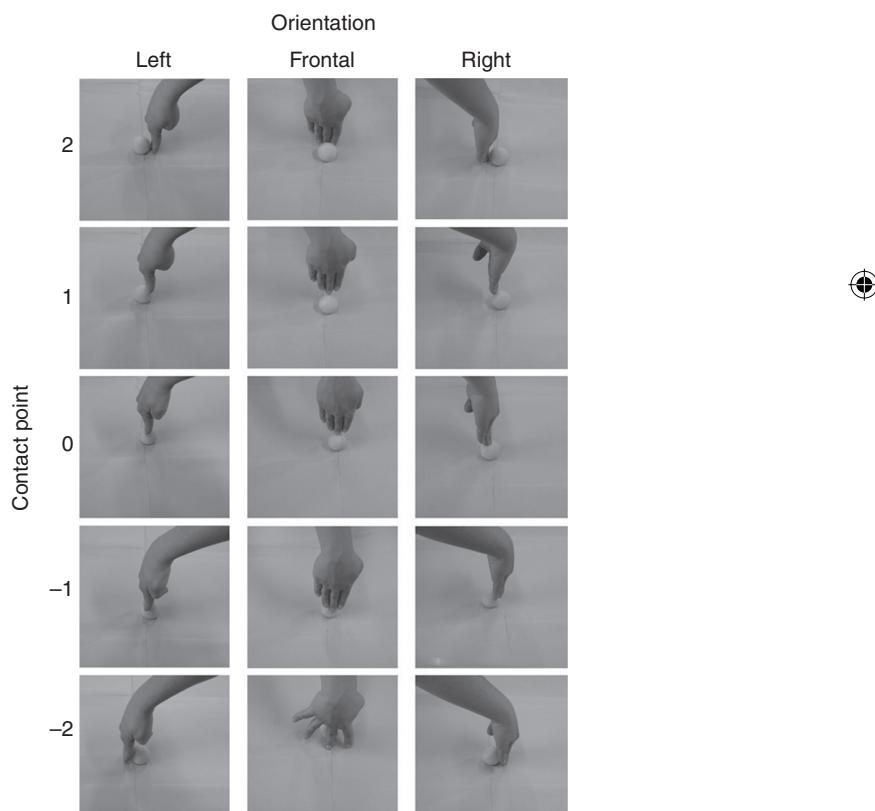
Cattaneo et al. (in press) developed a variant of the bean technique that not only appears to be more efficient, but also provides convincing new information that the motor adaptation technique can tap a MM. Their variant used chickpeas instead of cannellini beans...but that wasn't the important difference. Instead of using the arm to move beans from one container to another, Cattaneo's participants put their hand into a single jar of chickpeas and either pushed the chickpeas away, using the dorsal surface of the fingers, or pulled the chickpeas toward themselves by using the palm-side of the fingers. Participants performed one or the other action for just one minute. Following the

# PROOF

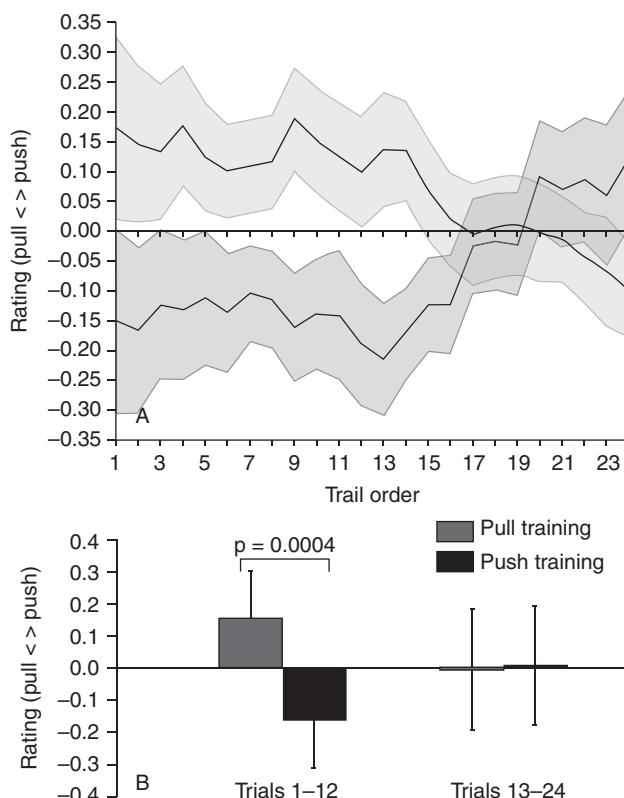
*Contributions of Mirror Mechanisms to the Embodiment of Cognition* 175

adaptation task, participants viewed pictures (for 500 msec each) of a hand interacting with a ball (see Figure 7.4) and judged whether the hand was pushing or pulling the ball. The procedure of pushing or pulling (manipulated within-subject) followed by judgments was repeated eight times (four blocks in each direction).

Figure 7.5 shows the results for the ambiguous stimuli (third row in Figure 7.4) as a function of order of presentation of the pictures. Once again, direction of adaptation affected the judgments. Importantly, the effect was quite short-lived: it is evident for the first 12 picture trials (about 30 seconds), but not after.

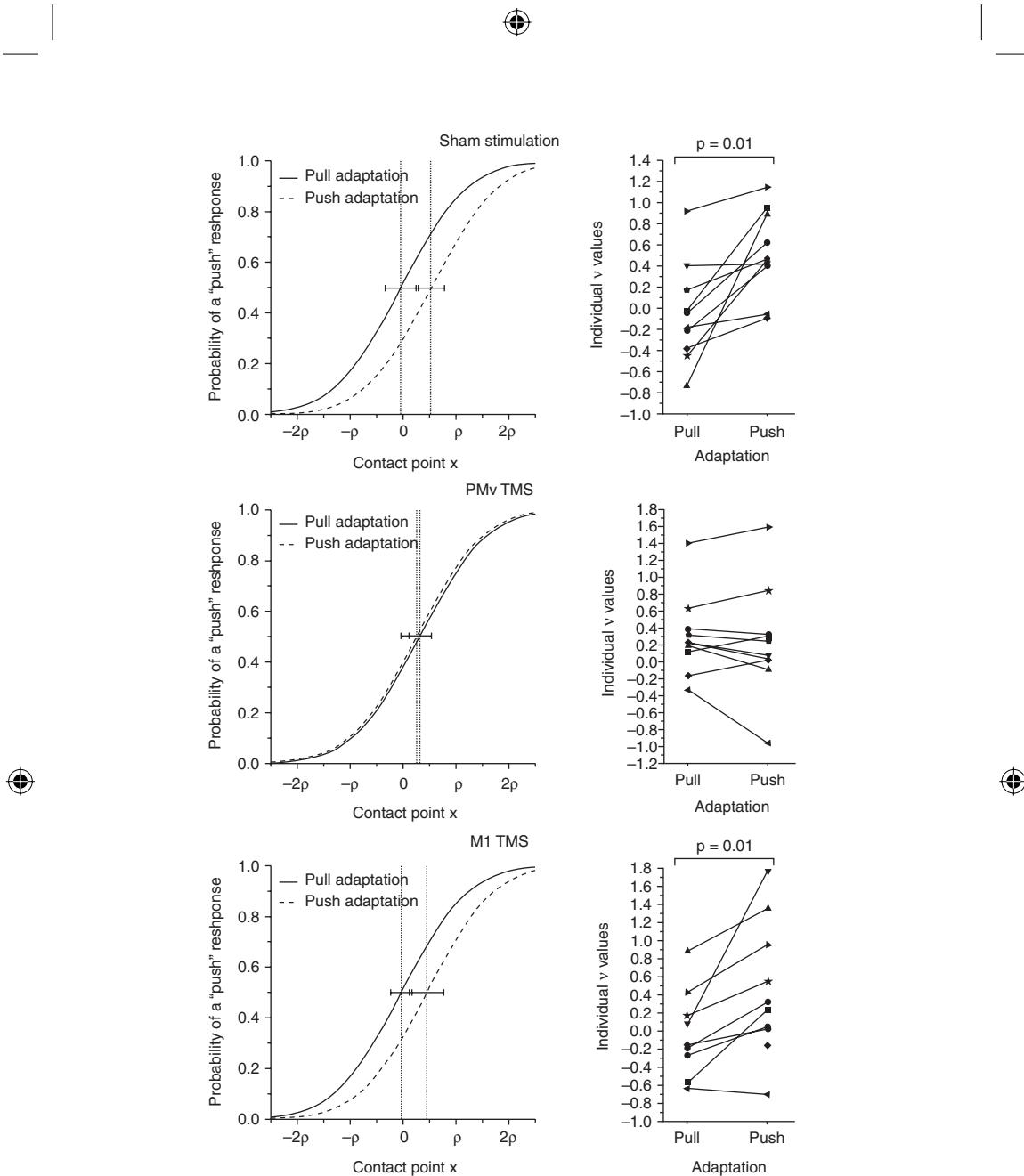


*Figure 7.4* The hand interacting with a ball. Positive contact points suggest movement of the ball away from the body; negative contact points suggest movement toward the body; contact points 0 are ambiguous.



*Figure 7.5* Time course of adaptation effect. A – Mean values of the ratings for the ambiguous stimuli (contact point  $x = 0$ ) sorted by order of appearance in the trials series. The shading represents  $\pm$  SEM. B – results of the ANOVA between mean ratings of the first half of trials and the second half. The error bars indicate 95% confidence intervals.

The real advance of the Cattaneo data comes from combining the procedure with TMS, in which a single TMS pulse was presented concurrently with the picture. The TMS pulse was in one of three conditions: pulse over left ventral premotor cortex (PMv), pulse over left hand motor cortex (M1), or a sham stimulus with the coil positioned midway between PMv and M1. The full psychometric functions (not just the responses to the ambiguous stimuli) are presented in Figure 7.6. Note first the effect of direction evident with the sham stimulation; the data at contact point 0 are a replication of the effects



**AQ9**

Figure 7.6 Left: grand-average of all the individual psychometric functions in the two adaptation conditions. The contact point  $x$  is expressed in units of participant's psychometric function slope  $\bar{v}$ . The dotted vertical lines represent  $\bar{x}$  (normalized distance between the functions) average values and horizontal error bars represent 95% confidence intervals. Right: individual values of  $\bar{v}$  for all 20 participants. The data for the three different TMS conditions are given. The  $p$ -values refer to significant Bonferroni corrected  $p$ -values of pairwise t-tests.

# PROOF

178 Arthur M. Glenberg

in Figure 7.5. Next, note that the effect of direction of adaptation is equally apparent with the TMS pulse directed at M1. However, when the TMS pulse is directed at PMv, the effect of adaptation direction is eliminated.

What is going on here? In previous work, Cattaneo and his colleagues demonstrated that single-pulse TMS has a particularly large effect on neurons that have been adapted. Thus, finding an effect of TMS (relative to the sham pulse) in PMv indicates that the adapted neurons are likely to be more highly represented in that area of cortex than in M1. Furthermore, other work has demonstrated that mirror neurons are more highly represented in PMv than in M1. Thus, the conclusion is that the motor task was particularly effective in adapting a MM.

One other aspect of Cattaneo's results, as well as the Glenberg et al. results in Figure 7.3, is important. In both cases, the adaptation effect is in the direction of habituation or a fatigue-like effect. In Figure 7.3, moving the beans Away changes the threshold so that more away-consistent information is required to judge the stimulus as moving away. Similarly in Figure 7.6, after pushing the chickpeas away, people require more unambiguous away information to judge the picture as illustrating an away action.

Cattaneo et al. offer a social hypothesis to explain the direction of adaptation. Consider first that adaptation in sensory systems improves perception by changing gain to avoid floor and ceiling effects (e.g., after the eyes adapt to the bright sun, it is easier to see small changes in brightness). Movement adaptation also adjusts gain to enhance sensitivity. Namely, after adaptation, we are more sensitive to perceiving others' actions when they are different from our own. Thus, it seems likely that MMs are embedded in a dynamic framework that is adjusted by both our own actions as well as those of others in the service of better (e.g., quicker) recognition of others' actions.

Before moving on, a brief recap is in order. Much previous work on human MMs has failed to convincingly demonstrate that a MM plays a role in action recognition. In fact, Hickock and Hauser (2010) suggest that both human and monkey mirror neuron findings have been misinterpreted in that the data really show a system that is used in action selection, not action recognition. In contrast, the motor adaptation paradigm produces very clear findings. The MM tapped by adaptation plays a causal role in action recognition. Furthermore, the direction of that adaptation suggests that one function is to heighten sensitivity to acts that are the opposite of our own.



### MMs in language: speech perception

A number of researchers have worked on determining whether there is a link between motor activity and speech perception. Such a link has long been suggested by the motor theory of speech perception (Galantucci et al., 2006; Liberman, 1957), and fMRI and TMS evidence (e.g., D'Ausilio, 2009) points to activation of speech motor areas during speech perception. However, evidence that the motor system plays a causal role in speech perception is rare, and evidence of exactly what that role may be is rarer still.

One of the rare instances of a demonstration of a causal link is provided by D'Ausilio (2009). Participants attempted to identify one of four consonants presented in white noise. Two of the consonants were tongue-related ("d" and "t") and two were lip-related ("b" and "p"). Some of the stimuli were preceded by a double TMS pulse, one pulse 100 msec before presentation of the stimulus and the other at 50 msec before presentation. Importantly, the double TMS pulse was directed at that part of the motor cortex that controls the lips or that part that controls the tongue. As is often the case, the double TMS pulse produced a facilitation, in this case both by speeding reaction time and reducing errors (identifying the consonant in the wrong class). The critical result was the interaction of TMS location and type of consonant. That is, TMS directed at the tongue area facilitated perception of the tongue-related consonants, whereas TMS directed at the lip area facilitated perception of the lip-related consonants. These results clearly demonstrate a MM that has a causal effect on speech perception.

But can we say more about the nature of that causal effect? Application of the motor adaptation paradigm to speech perception does just that. Sato et al. [submitted] asked participants to identify a speech stimulus as an instance of "pa" or "ta" (multiple instances from multiple speakers were used). In some trials the stimuli were presented in noise and in other trials they were presented alone. The identification task was presented in one of four contexts. For half of the participants, the speech stimuli were presented either after pursing the lips 150 times or with no motor adaptation (in counterbalanced order). For the other participants, the speech stimuli were presented after lifting the tongue to the back of the teeth 150 times or with no motor adaptation (in counterbalanced order). Pursing the lips was meant to adapt a MM that participates in controlling the lips, whereas lifting the tongue was meant to adapt a MM that controls the tongue.

AQ10



# PROOF

180 Arthur M. Glenberg

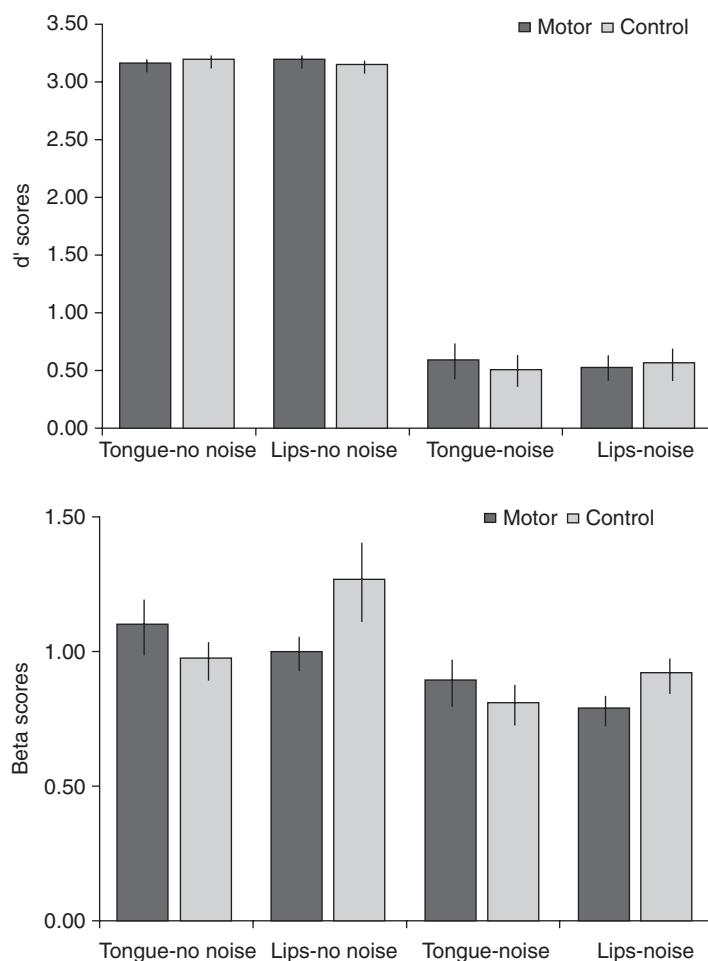
Sato et al. analyzed their data using signal detection theory to produce two dependent measures. The  $d'$  measure is an index of sensitivity, that is, how well the system can accurately discriminate between the syllables. The *beta* measure is an index of bias, that is, what a syllable is labeled. As expected, whether or not the syllable was presented in noise had a very large effect on  $d'$ . However, whether the block of trials was preceded by an adaptation task, and the nature of that adaptation task (tongue or lips), did not have a significant effect on  $d'$ .

In contrast, type of adaptation task did significantly affect *beta*, and the effect was the same both when the syllables were presented in noise and with no noise. *Beta* was computed using the probability of calling a syllable "pa" both when it really was a "pa" (i.e., a hit) and when it was a "ta" (i.e., a miss). Given this method of calculation, an increase in *beta* indicates the need for more evidence before the participant is willing to call a stimulus "pa."

Turning to the bottom panel of Figure 7.7, we can see how this plays out. After tongue motor adaptation, *beta* increases (relative to no adaptation). That is, participants needed more evidence to call a syllable a "pa," and less evidence to call a syllable "ta." In contrast, after lip adaptation, *beta* decreased (relative to no adaptation). That is, participants needed less evidence to call a syllable "pa" and more evidence to call a syllable "ta." These effects were found both when the syllables were presented in noise (right side of figure) and when no noise was used (left side of figure).

AQ11 Although speculative, Sato et al. [submitted] offer an interpretation of these data that helps to mediate between those who propose that the motor system plays a role in speech perception and those who are skeptical. First, based on the  $d'$  measure, the motor system plays no role in the neural coding of the stimulus in temporal lobe and other auditory structures. Second, based on the *beta* measure, the motor system does influence how that neural code is categorized, that is, whether it sounds like a "pa" or a "ta."

The Sato et al. interpretation fits with the results of D'Ausilio et al. (2009), although direct comparison is difficult because the use of four stimuli in D'Ausilio et al. precludes a simple signal detection analysis. When producing a particular consonant, commands are sent to motor cortex, although the exact nature of the sound that is produced will also depend on the current state of the articulators (e.g., where lips are after having produced the previous sound) and the need to produce the next sound. These three aspects of articulation (location of the articulators, current motor command, anticipated command) combine to produce



*Figure 7.7* D-prime and beta scores with (motor) and without (control) adaptation of the tongue and lips.

the phenomenon of coarticulation of speech sounds so that there is not a one-to-one correspondence between motor command and produced sound, nor is there a one-to-one correspondence between produced sound and perceived sound. In this situation, if the perceiver's motor cortex is primed, either by motor adaptation (as in Sato et al.) or by a double TMS pulse (as in D'Ausilio), the state of the motor system will bias interpretation of the stimulus toward a stimulus that the motor

182 Arthur M. Glenberg

system might have just produced. That is, we are biased to hear a stimulus similar to the one that our own motor system might have just produced. This sort of bias would a) produce the effects on *beta* observed by Sato et al., b) speed responding as observed by D'Ausilio et al., and c) reduce cross-category misses (e.g., calling a tongue-related stimulus a lip-related stimulus).

### MMs in language: comprehension

One of the most striking demonstrations of a MM in language comprehension was provided by Aziz-Zadeh et al. (2006). While fMRI data were being collected, participants read phrases that referred to the mouth (e.g., “biting the peach”), the hand (e.g., “grasping the scissors”), or the foot (e.g., “pressing the car brake”). Later, the same participants saw videos of actions similar to those described. Analyses showed a strong congruency: left premotor voxels most strongly activated by the mouth videos were most strongly activated by the mouth phrases; left premotor voxels most strongly activated by the hand videos were most strongly activated by the hand phrases; and left premotor voxels most strongly activated by the foot videos were most strongly activated by the foot phrases. This congruency was not evident in the right hemisphere, nor was it evident for metaphorical phrases (e.g., “grasping the idea,” “kicking off the year”).

These results are certainly consistent with the embodiment claim that language understanding involves a simulation using neural systems of perception, action (as in this case), and emotion. The data are also consistent with the claim that a MM affects language comprehension. Nonetheless, there are two reasons for skepticism. First, as with almost all fMRI data, the result is correlational. That is, there is no strong evidence that the premotor activations were causally involved in language comprehension. Second, and relatedly, the temporal relation between comprehension and observed voxel activation is unknown. Consequently, the voxel activation might reflect a postcomprehension process, such as motor or visual imagery, rather than comprehension itself.

Fortunately, the adaptation procedure addresses both reasons for skepticism. In an initial experiment (Glenberg et al., 2008a; experiment 1), participants moved 600 beans, one at a time, either away from themselves or toward themselves. (In this experiment, the participants were not blindfolded during the motor adaptation task. We have since replicated the findings using blindfolded movement.) Following

# PROOF

*Contributions of Mirror Mechanisms to the Embodiment of Cognition* 183

adaptation, participants determined whether sentences were sensible or nonsense by using the right index finger to make the sensible response and the left index finger to indicate nonsense. The sensible sentences all described a type of transfer. The transfer was either toward the participant (e.g., "Mike gives you a pen") or away from the participant (e.g., "You give Mike a pen"). In addition, the transfer was either of a concrete direct object (e.g., a pen) or of an abstract direct object (e.g., "Anna delegates the responsibilities to you"). The dependent measure was the time taken to indicate that a sentence was sensible.

The two leftmost bars in Figure 7.8 report the results. The height of the bar indexes the extent to which motor adaptation slows the sensibility judgment of sentences describing transfer in the same direction. For example, the first bar in the figure is, for abstract sentences, the average of a) time taken to read an Away sentence after bean movement Away minus the time taken to read an Away sentence after bean movement Toward, and b) the time taken to read a Toward sentence after bean movement Toward minus the time taken to read a Toward sentence after bean movement Away. Thus, the first bar indicates that, for abstract sentences, having moved the beans in the same direction as that described in the sentences slows sentence understanding by about

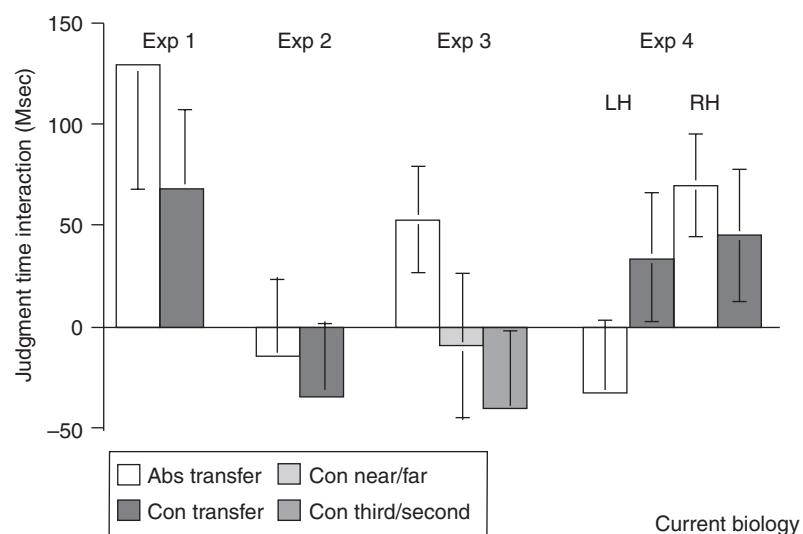


Figure 7.8 Motor adaptation and language comprehension. See text for explanation.

# PROOF

184 Arthur M. Glenberg

125 msec. The second (darker) bar indicates a similar effect for the concrete sentences. Thus, to the extent that the motor adaptation task taps a MM, we have demonstrated here that a MM causally affects language comprehension.

An alternative interpretation of the results in Figure 7.8 is that, during the adaptation task, the participants were saying to themselves "toward" or "away." This repetition might have led to semantic satiation, that is, loss of meaning of the term due to repetition, and perhaps it is this loss of meaning that produced the contrast effect. A second experiment tested this alternative.

Instead of moving 600 beans, participants engaged in 600 lexical decisions. In one condition, participants determined whether a visual stimulus was the word "toward," or whether the stimulus was a non-word anagram (e.g., "toward"). In the other condition, participants determined whether the visual stimulus was the word "away" or one of its anagrams. This lexical decision task should also have produced a semantic satiation. As is evident from the two bars in Figure 7.8 labeled "Exp 2," the lexical decision task did not produce the adaptation effect; apparently, motor activity is required to produce adaptation that affects comprehension.

Because the finding of an adaptation effect with the abstract transfer sentences is not intuitive, we replicated that effect in a third experiment. In addition, in the third experiment we used the concrete sentences to test two alternative explanations. Perhaps the effect was not due to a contrast of bean direction movement and direction of transfer, but instead reflects a contrast between bean direction movement and sentences that described objects near the grammatical subject (e.g., "You hold the coins in your hand") or away from the grammatical subject (e.g., "You see the coins at the bottom of the well"). The second alternative was that the contrast was between bean direction movement and use of first-person sentence subjects or second-person sentence subjects (e.g., "Andy smells the pizza on his plate"). The bars in Figure 7.8 labeled "Exp 3" show that the adaptation effect was replicated for the abstract transfer sentences (white bar). However, distance of the object from the subject of the sentence does not seem to matter (light gray bar), and grammatical person does not seem to matter (darker gray bar).

In the final experiment, we tested the degree to which the effect is exquisitely embodied; that is, will the adaptation effect hold only when the responses in the sentence judgment task are made with the hand used in the adaptation task? The bars in Figure 7.8 labeled "Exp 4" show

AQ12





that, indeed, the adaptation effect is larger when the same hand is used for adaptation and for responding “sensible” in the sentence task.

There is an apparent contrast between the data in Figure 7.8 and the data reported by Aziz-Zadeh et al. (2006). Namely, Aziz-Zadeh and colleagues report no mirror activation for the “abstract” metaphorical sentences. The data in Figure 7.8 show a large contribution of a MM to the comprehension of sentences describing metaphorical transfer of abstract direct objects. The difference might be due to differences in the tasks and measurement procedures. There is probably a simpler explanation, however. In imaging the situation described by a sentence such as “You delegate the responsibilities to Anna,” most people imagine themselves using a literal arm motion to effect the transfer (e.g., handing Anna a sheet of paper that describes the responsibilities). In fact, many sign languages use an outward-going arm to convey both literal and metaphorical transfer. In contrast, if asked to imagine themselves in a situation described by “kicking off the year” (an example of the metaphorical phrase used by Aziz-Zadeh et al.), it is unlikely that many imagine themselves kicking. Hence, understanding of the metaphorical phrases used by Aziz-Zadeh et al. may require a simulation (e.g., of a celebration) but not a simulation using the particular motor effectors investigated in the Aziz-Zadeh et al. experiment.



### Conclusions

This review leads to one secure and one tentative conclusion. The secure conclusion is that many cognitive processes are embodied. That is, the results using the adaptation paradigm demonstrate contributions of motor mechanisms to action recognition (Figures 7.3, 7.5, and 7.6), speech perception (Figure 7.7), and language comprehension (Figure 7.8). (Of course, it is almost certainly the case that these processes are embodied in other ways in addition to motor activity. As noted earlier, there is good evidence that language comprehension draws upon neural and bodily processes of perception (e.g., Ruschemeyer et al., in press) and emotion (Havas et al., 2007; 2010).) Given the extent of the data, it is not too radical to suggest that cognition may be embodied through and through. That is, no cognitive processes operate on pure, abstract symbols that have no grounding in neural and bodily systems of action, perception, and emotion.

The less secure conclusion is that mirror neurons and mirror mechanisms play a major role in the embodiment of cognitive processes. As described earlier, there are several reasons to be skeptical regarding the





186 Arthur M. Glenberg

contribution of mirror mechanisms to human cognition, and perhaps even macaque cognition, as suggested by Hickok and Hauser (2010). On the other hand, if the adaptation paradigm taps mirror mechanisms, then some of the reasons for skepticism are obviated. That is, the paradigm seems to demonstrate that mirror mechanisms play a causal role in cognitive processes. But does the paradigm really tap mirror mechanisms?

The original definition of a mirror neuron was based on single-cell recording. While powerful in demonstrating that a particular cell can be active during both action perception and recognition, the technique cannot be conclusive about function. When the study of mirror mechanisms moved to humans, the definition morphed so that it included anatomical locations as a major criterion. Thus, activity in a particular locus (e.g., Broca's area), particularly when the activity level is similar for both action production and action perception, was taken as critical.

The adaptation paradigm suggests another change in what should be taken as critical to the definition of a MM, namely, that the putative MM be shown to play a causal and functional role in some cognitive process such as action recognition. The difficulty is in tying this causal mechanism to the physiological mechanisms identified by other techniques such as single-cell recording and fMRI activity. Localization of adaptation, as demonstrated by Cattaneo et al. (in press, Figure 7.6), provides some evidence in this direction.

In the final analysis, it is likely that the solution will be some variant of what psychology and cognitive neuroscience have always depended on. Namely, does postulating a MM advance our thinking and understanding of phenomena? Or, in this case, is it an advance to postulate that MMs underlie the causal data found with the motor adaptation procedure? As usual, it is likely that only future data and theorizing will answer these questions with any certainty.

## References

- AQ13
- Aziz-Zadeh, L., Wilson, S. M., Rizzolatti, G., and Iacoboni, M. (2006). Congruent embodied representations for visually presented actions and linguistic phrases describing actions. *Current Biology*, 16, 1,818–23.
- Calvo-Merino, B., Grezes, J., Glaser, D., and Haggard, P. (2006). Seeing or doing? Influence of visual and motor familiarity in action observation. *Current Biology*, 16, 1,905–10.
- Cattaneo, L., Arfeller, C., Tabarelli, D., Sato, M., and Glenberg, A. M. (in press). *Social Cognitive and Affective Neuroscience*.



# PROOF

- Chong, T. T.-J., Cunnington, R., Williams, M. A., Kanwisher, N., and Mattingley, J. B. (2008). FMRI adaptation reveals mirror neurons in human inferior parietal cortex. *Current Biology*, *18*, 1,576–80.
- Classen, J., Liepert, J., Wise, S. P., Hallett, M., and Cohen, L. G. (1998). Rapid plasticity of human cortical movement representation induced by practice. *Journal of Neurophysiology*, *79*, 1,117–23.
- Dapretto, M., Davies, M. S., Pfeifer, J. H., Scott, A. A., Sigman, M., Bookheimer, S. Y., and Iacoboni, M. (2006). Understanding emotions in others: mirror neuron dysfunction in children with autism spectrum disorders. *Nature Neuroscience*, *9*, 28–30.
- D'Ausilio, A., Pulvermüller, F., Salmas, P., Buflari, I., Begliomini, C., and Fadiga, L. (2009). The motor somatotopy of speech perception. *Current Biology*, *19*, 381–5.
- Dinstein, I., Thomas, C., Behrmann, M., and Heeger, D. J. (2008). A mirror up to nature. *Current Biology*, *18*, R13–R18.
- Fazio, P., Cantagallo, A., Craighero, L., D'Ausilio, A., Roy, A. C., Pozzo, T., Calzolari, F., Granieri, E., and Fadiga, L. (2009). Encoding of human action in Broca's area. *Brain*, doi:10.1093/brain/awp118.
- Galantucci, B., Fowler, C. A., and Turvey, M. T. (2006). The motor theory of speech perception reviewed. *Psychonomic Bulletin and Review*, *13*, 361–77.
- Gallese, V., and Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, *2*, 493–501.
- Gallese, V., Fadiga, L., Fogassi, and Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, *119*, 593–609.
- Gallese, V., Keysers, C., and Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends in Cognitive Sciences*, *8*, 396–403.
- Gazzola, V., Aziz-Zadeh, L., and Keysers, C. (2006). Empathy and the somato-topic auditory mirror system in humans. *Current Biology*, *16*, 1,824–9.
- Glenberg, A. M. (1997). What memory is for. *Behavioral and Brain Sciences*, *20*, 1–19.
- Glenberg, A. M. (2006). Naturalizing Cognition: The Integration of Cognitive Science and Biology. *Current Biology*, *16*, R802–R804.
- Glenberg, A. M. (2010). Embodiment as a unifying perspective for psychology. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*, 586–96.
- Glenberg, A. M., and Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin and Review*, *9*, 558–65.
- Glenberg, A. M., and Robertson, D. A. (1999). Indexical understanding of instructions. *Discourse Processes*, *28*, 1–26.
- Glenberg, A. M., Robertson, D. A., Jansen, J. L., and Johnson-Glenberg, M. C. (1999). Not propositions. *Journal of Cognitive Systems Research*, *1*, 19–33.
- Glenberg, A. M., Sato, M., and Cattaneo, L. (2008a). Use-induced motor plasticity affects the processing of abstract and concrete language. *Current Biology*, *18*, R290–R291.
- Glenberg, A. M., Sato, M., Cattaneo, L., Riggio, L., Palumbo, D., and Buccino, G. (2008b). Processing abstract language modulates motor system activity. *Quarterly Journal of Experimental Psychology*, *61*, 905–19.
- Havas, D. A., Glenberg, A. M., and Rinck, M. (2007). Emotion simulation during language comprehension. *Psychonomic Bulletin and Review*, *14*, 436–41.
- Havas, D. A., Glenberg, A. M., Gutowski, K., Lucarelli, M., and Davidson, R. (2010). Cosmetic use of botulinum toxin-A affects processing of emotional language. *Psychological Science*, *21*, 895–900.

AQ14



# PROOF

188 Arthur M. Glenberg

- Hickock, G. (2009). Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *Journal of Cognitive Neuroscience*, 21, 1,229–43.
- Hickock, G., and Hauser, M. (2010). (Mis)understanding mirror neurons. *Current Biology*, 20, 593–4.
- Kilner, J. M., Marchant, J. L., and Frith, C. D. (2006). Modulation of the mirror system by social relevance. *Social Cognitive and Affective Neuroscience*, 1, 143–8.
- Kilner, J. M., Neal, A., Weiskopf, N., and Frith, C. D. (2009). Evidence of mirror neurons in human inferior frontal gyrus. *Journal of Neuroscience*, 29, 10,153–9.
- Kohler, E., Keysers, C., Umiltà, M. A., Fogassi, L., Gallese, V., and Rizzolatti, G. (2002). Hearing Sounds, Understanding Actions: Action Representation in Mirror Neurons. *Science*, 297, 846–8.
- Lewis, C. F., and McBeath, M. K. (2004). Bias to experience approaching motion in a three-dimensional virtual environment. *Perception*, 33, 259–76.
- Liberman, A. M. (1957). Some results of research on speech perception. *Journal of the Acoustical Society of America*, 29, 117–23.
- AQ15 Lingnau, A., Gesierich, B., and Caramazza, A. (2009). Asymmetric fMRI adaptation reveals no evidence for mirror neurons in humans. *Proceedings of the National Academy of Sciences*. www.pnas.org/cgi/doi/10/1073/pnas.0902262106.
- Mahon, B. Z. (2008). Action recognition: Is it a motor process? *Current Biology*, 18, R1,068–9.
- Mukamel, R., Ekstrom, A. D., Kaplan, K., Iacoboni, M., and Fried, I. (2010). Single-neuron responses in humans during execution and observation of actions. *Current Biology*, doi:10.1016/j.cub.2010.02.045.
- Pazzaglia, M., Pizzamiglio, L., Pes, E., and Aglioti, S. M. (2008). The sound of actions in apraxia. *Current Biology*, 18, 1,766–72.
- Ramachandran, V. S. (2005). MIRROR NEURONS and imitation learning as the driving force behind “the great leap forward” in human evolution. *Nova, Science Now* website.
- Rizzolatti, G., and Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, 27, 169–92.
- Rizzolatti, G., and Craighero, L. (2007). Language and mirror neurons. In G. Gaskell (Ed.), *The Oxford handbook of psycholinguistics*. Oxford, UK: Oxford University Press, pp. 771–86.
- Rueschemeyer, S.-A., Glenberg, A. M., Kaschak, M., Mueller, K., and Friederici, A. (in press). Top-down and bottom-up contributions to understanding sentences describing objects in motion. *Frontiers in Cognition*.
- Sato, M., Grabski, K., Glenberg, A. M., Brisebois, A., Basirat, A., Menard, L., and Cattaneo, L. (in press). Articulatory bias in speech perception: Evidence from use-induced motor plasticity. *Brain and Language*.
- Saygin, A. P., McCollough, S., Alac, M., and Emmorey, K. (2010). Modulation of BOLD Response in Motion-sensitive Lateral Temporal Cortex by Real and Fictive Motion Sentences. *Journal of Cognitive Neuroscience*, 22, 2,849–90.
- AQ16 Tettamanti, M., Buccino, G., Saccuman, M. C., Gallese, V., Danna, M., Scifo, P., et al. (2005). Listening to action-related sentences activates fronto-parietal motor circuits. *Journal of Cognitive Neuroscience*, 17, no. 2, 273–81.

# PROOF



*Contributions of Mirror Mechanisms to the Embodiment of Cognition* 189

Tyrrell R. A., and Owens, D. A. (1988). A rapid technique to assess the resting states of the eyes and other threshold phenomena – The modified binary search (MOBS). *Behavior Research Methods Instruments and Computers*, 20, 137–41.

AQ17

Umiltà, M. A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C., et al. (2001). I know what you are doing: A neurophysiological study. *Neuron*, 31, 155–65.

Umiltà, M. A., Escola, L., Intskirvell, I., Grammont, F., Rochat, M., Caruana, F., Jezzini, A., Gallese, V., and Rizzolatti, G. (2008). When pliers become fingers in the monkey motor system. *Proceeding of the National Academy of Sciences*, 105, 2,209–13.

Zwaan, R. A., and Taylor, L. J. (2006). Seeing, acting, understanding: motor resonance in language comprehension. *Journal of Experimental Psychology: General*, 135, 1–11.





## 8

## The Neural Systems Involved in Motor Cognition and Social Contact

*Sébastien Hétu and Philip L. Jackson*

### Introduction: not all contacts are first contacts

I had just received my badge, the conference program, and the complimentary laptop sleeve; I turned around and stood face to face with John, who was also waiting in line to register for the conference. We exchanged a short word of greeting and we shook hands wholeheartedly. When this encounter was over, I bumped into Mary; we both expressed our surprise and then hugged each other. Right after this hug, I nodded politely to Paul passing behind Mary, noticing that he wanted to engage in a conversation. Had Mary been accompanied by her spouse, perhaps the hug would have been toned down to a warm handshake. Had I not hugged Mary, I might have had more trouble avoiding Paul, whom I suspected to be reviewer #3 of my recently rejected paper. This type of familiar scene is usually fluid and many of these social interactions feel automatic. However, if we analyze further the simple action of the greeting, the initial social contact with others, we learn that incalculable cognitive operations subserved by a number of neurofunctional systems are at play. But at the source of this cascade of cerebral processing often lie the actions of the others.

The complexity of the human brain is such that the mere perception of someone's actions often leads to the decoding of a specific social intent, as when I see Paul miming an eating gesture to someone at the other side of the conference hall and understand that he does not want to eat alone. Such social understanding of other people's behavior is essential in adapted social contacts. The previous example does not imply a simplistic automatic decoding mechanism from a perceived action to social cognition: there is a leap between the representation



of an action and a social interaction. We recognize that a number of cognitive processes are necessary to reach beyond the simple motor representation of an observed action, and we also acknowledge that this basic route from the representation of the other's movement to its social representation can even be sidestepped.

To reflect this integration of motor and cognitive processes, inspired by James's ideomotor theory (James, 1890), Dewey's Coordination (Dewey, 1896), and others, the contemporary term *motor cognition* has been coined (Jeannerod, 2001). This integration of motor representation with higher level computation is subserved by two main processes: 1) motor resonance: the mapping of the others' behaviors onto our own motor repertoire, which can also be seen as the coupling of perception and action (Prinz, 1997; Sperry, 1952), and 2) mentalizing: the inference mechanism allowing the exploration of others' perspective and mental state (Frith and Frith, 2006). These processes are cortically embedded in two neural systems briefly presented in Figure 8.1 and which will be the focus of the present chapter: the **resonance system** and the **mentalizing system**. The relative importance of each of these systems to motor cognition and social interactions will be discussed in this chapter, based notably on recent neuroimaging literature. In the first part we will show that motor resonance can be influenced by a number of variables that are subdivided into three categories: other-related variables, self-related variables, and self-other interaction variables. We will also present evidence to the effect that this resonance mechanism is functionally plastic. Because I can understand that John, my colleague and friend, who is raising his fist, is not threatening me but rather showing me that he knows that I got an important grant that he also wanted, there must be a system able to use and complement the information arising from the motor analysis of my friend's movement. Thus, the second part of this chapter will put forward evidence that the mentalizing mechanism is necessary to make possible the decoupling from the actual perceived information in order to explore hypotheses based on stored information. The targets of such hypotheses are usually other humans, but our ability to anthropomorphize can extend the reach of this skill to any concrete or abstract concept. Moreover, the fact that the target is active and conscious, and interacts with the observer (as opposed to James's candle in the child–candle example), adds computational requirements beyond simple perception–action coupling. The last section of this chapter will open the debate on the extent to which these two mechanisms – resonance and mentalizing – are sufficient to support human social contact and to provide flavor to human interactions.



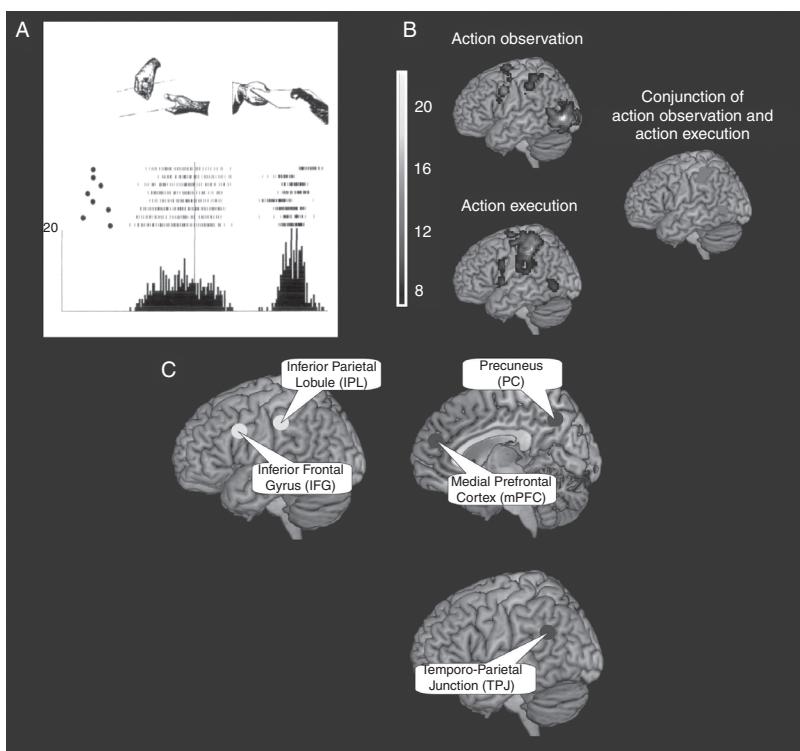


Figure 8.1

From mirror neurons to motor resonance and mentalizing. Mirror neurons are neurons that fire during the production and the observation of movements. Neurons possessing such properties were first discovered with single-cell recording studies in the premotor cortex (area F5) of nonhuman primates by Giacomo Rizzolatti's laboratory (di Pellegrino et al., 1992; Gallese et al., 1996). (A) Single-cell recording of a mirror neuron. The upper part schematically represents the two behavioral conditions (on the left, the experimenter grasping a piece of food while the monkey observes and, on the right, the monkey grasping the piece of food) while the lower part shows eight consecutive trials and the relative response histogram of the neuron. Note that this neuron discharge occurs both when the monkey observes and when it performs a grasping movement. Even if mirror neurons in humans have yet to be directly observed, as single-cell recording is seldom used in humans,

AQ1

scientists have utilized functional magnetic resonance imaging (fMRI) and conjunction analyses to try to identify the presence of structures that were active during both the perception and the production of actions. (B) Example of such conjunction analyses in which the maps of brain activity during action observation and action execution are superimposed to identify regions that are active during both processes (Hétu et al., unpublished data). Such patterns of activity have led to the motor resonance terminology, in which the observer's brain "resonates" with the model's (i.e., some brain regions active in the person performing the action are also active in the person observing him). In addition to the resonance system, the mentalizing system is also thought to be involved during social interaction. The mentalizing system is involved in the inference mechanism allowing the exploration of others' perspective and mental state (Frith and Frith, 2006). (C) On the left, the fronto-parietal resonance system (also known as the mirror neuron system) in humans. The respective poles of this system are the inferior frontal gyrus (IFG) and the inferior parietal lobule (IPL) (Cattaneo and Rizzolatti, 2009; Fabbri-Destro and Rizzolatti, 2008; Rizzolatti et al., 2009). Note that the ventral premotor cortex (vPMC) is also often mentioned but is not shown here. On the right is the mentalizing system. The brain areas of the mentalizing system are often considered as the precuneus (PC), temporo-parietal junction (TPJ) and the medial pre-frontal cortex (mPFC) (Van Overwalle, 2009). The yellow/blue dots and lines are used to provide a rough localization of the regions and are thus not representative of the exact boundaries of these areas.

This chapter does not pretend to address the full complexity of social interactions and the different neurocognitive systems that support it; our objective is to depict how a simple, and to all appearances automatic, action (such as a greeting) is represented through a number of distinct yet interacting cerebral systems and sparks information beyond what is perceived and acted. The form of the greetings I exchanged with John, Mary, and Paul was the product of a succession of motor acts and perceptions of these acts. Understanding other people's sensory-motor states thus has clear adaptive advantages, and understanding emotional signals associated with those states is especially important in the formation and maintenance of social relationships.

### **Motor resonance (from others to self)**

The representation of a complex action can often be conveyed through a single and simple sensory input. For instance, seeing another individual

reaching for a glass of water is enough to evoke the action of drinking or even thirst; hearing the sound of a saw cutting something is enough to evoke the idea of someone holding a saw and trying to cut a piece of wood. The term *Shared Representation* has been proposed to reflect how the cerebral systems at play during the observation of a behavior are very similar to those involved during the production of that behavior (Jeannerod, 2001). Initially reserved for motor acts, this term has been used (not without criticism) to reflect that this sharing of cerebral representations is also observed for other concepts such as touch (e.g., Keysers et al., 2004), pain (e.g., Jackson et al., 2005; Saarela et al., 2007), and emotions (Carr et al., 2003; Wicker et al., 2003). The extent of this shared representation (in terms of similar brain activity) is thought to reflect the level of resonance of the observer. Thus, the level of motor resonance in an individual is thought to be reflected by the extent of the overlap between the cerebral representation of the produced and the observed action.

While phylogenetic and ontogenetic development trajectories suggest that motor resonance takes root in primitive behavior, as it can be observed at the neural level in very young babies (Fecteau et al., 2004) and monkeys (Rizzolatti et al., 1996), it remains true that the simple resonance between the observer's and the model's brains can extend beyond the motoric component and serve a function beyond its reproduction. When we interact with other individuals and observe their movements, we usually do not automatically mimic their behavior but, rather, we adjust our own actions in response to what the other person is doing. Indeed, if, while looking at Paul pouring coffee in my mug, I was reproducing the same movements, the results would be inefficient to say the least. Thus, during social behavior, motor resonance and its underlying shared motor representation need to be integrated with and modulated by other cognitive processes such as memory and executive functions in order to produce an effective and socially adapted response. Motor resonance by which someone gets access to part of the mental representation of another person's action can be modulated by a number of variables, often associated with social interactions, that we here divide into three categories: *other-related variables* relating to the nature of the stimulus/target; *self-related variables* concerned with the individual characteristics of the observer, such as gender and experience; *self-other interaction variables* that taint the relationship between the observer and the target, such as the social bond between them and the context. My interactions with Mary and Paul necessarily depend on who they are, who I am, and the nature of our relationships. In the

AQ2



following section, we will provide evidence that the motor resonance system can be influenced by such variables.

### The other: how the stimulus/target modulates motor resonance

*The nature of the other.* Perhaps the most important variable that affects the extent of the resonance between an observer and a target is the target itself. What or whom the observer is looking at, what is being perceived, whether the target is animate or not is the first source of variability of motor resonance. If the cerebral system subserving motor resonance preferentially responds to human movements, this could indicate that it specializes in the processing of this type of visuomotor stimuli as opposed to other types of motion, such as mechanical motion. If this is the case, it would suggest that the system evolved specifically for human–human interactions. Many studies on this topic have compared human and robot motion to investigate the selectivity of the resonance system. Until now, both behavioral and imaging studies have produced conflicting results. Some have shown that human participants respond only, or preferentially, to human motion as compared with robots. For instance, Kilner et al. (2003) demonstrated that, while watching incongruent human actions interfered with movement production, watching the same actions performed by a robot did not (see also Brass et al., 2000, for an example of positive effect of congruency of observed movements). However, others have found that neural activation in premotor and parietal regions, areas involved in motor resonance, is similar during simple human and robot action observation (Gazzola et al., 2007a). Note that this similarity disappeared with the repetition of movements, notably due to a reduction of response to the robots, supporting a notion discussed in a subsequent section, but which is too often ignored, that motor resonance and its underlying cerebral representation are plastic and subject to change.

In addition to studies on human vs. robot motion, some neuropsychological studies with patients with cerebral lesions have pushed the exploration of the perception of animacy, with the use of point-light walkers (Heberlein et al., 2004) and animated geometrical figures (Heberlein and Adolphs, 2004), and the findings generally point to limbic regions, as opposed to motor regions, as playing an important role in the processing of such stimuli. This exemplifies that who or what we observe performing an action will modulate brain activity in a number of regions, and this modulation suggests that other, and likely higher level, processes are



at play, and not just direct motor matching, especially when the target presents with some form of ambiguity, as in the examples above, or when the ambiguity comes from our inexperience with the target, as in the case of meeting for the first time a friend's colleague at a conference.

*The actions of the other.* During a normal day, we perceive several different types of movements. Some of these involve objects, such as when I look at my fellow scientist Paul grabbing his badge. Some of these have a meaning but do not involve objects, such as when I see Paul waving at one of his students. Still other movements have no meaning at all, such as when I observe the balancing motion of Paul's arms while he walks toward his student.

Several attempts have been made to identify how various types of movement can modify our brain response when we observe another person move. The influence of meaningfulness on neural response was tested in a study on the yawning-contagion effect, showing that looking at meaningless movements of the mouth produced increased resonance responses when compared with yawns (meaningful mouth movements; Schurmann et al., 2005). Increased activity during meaningless compared with meaningful actions was also found in a study by Hétu and colleagues (Hétu et al., 2010), which found that regions within the resonance system (such as the inferior frontal gyrus (IFG) and inferior parietal lobule (IPL)) were most active during the observation of meaningless actions. Still, others have found opposite results, with meaningful actions producing greater brain activity in the resonance system than meaningless ones (Knutson et al., 2008; Villarreal et al., 2008). This variability in brain imaging findings is probably related to the difficulty in defining and manipulating specifically one at a time the many parameters that characterize the nature of an action (e.g., its kinematics, the presence of an object or a specific goal). Nevertheless, it seems that, when participants in an experiment perceive a change in the nature of the movement observed, its corresponding cerebral representation can also change.

Pushing forward this notion that the nature of the observed action can influence the activity of the resonance system, some work has focused on one specific type: socially relevant movements. If the resonance system is mainly involved in social interaction, it should respond preferentially to movements used for social communication (e.g., waving to say hello) than to nonsocial movements (e.g., grasping a pen). Indeed, this is what Villarreal et al. (2008) demonstrated as they compared communicative/symbolic gestures (e.g., thumbs up) with pantomimes of object-related actions (e.g., hammering motion without the hammer) and found that the response in the IFG was greater for communicative

# PROOF



*The Neural Systems Involved in Motor Cognition and Social Contact* 197

gestures. Still, one should remain careful before concluding that the resonance system is particularly involved in observation of motor stimuli with social content, as this distinction between communicative and object-related hand movements had not been found in an earlier report (Montgomery et al., 2007). A more pragmatic conclusion would be that adding a social connotation to a movement can in some cases change the cortical representation of the observed action (see, for example, Emmorey et al., 2010).

Scientific evidence arising from brain imaging as well as behavioral approaches thus suggests that the very nature of what we observe can influence how the resonance system processes these visual inputs. What or whom we are observing and the type of movement performed seem to modulate the functioning of the system responsible for the motor resonance response. This is logical, as I should not “resonate” in the same way to the motion of the hands of my watch when looking to see if I’m late for the keynote lecture than to the movements of Mary pointing towards an empty seat beside her when I enter a crowded conference room.

## Resonance in the eye of the beholder – self-related factors



We posit that the resonance system is not a simple automatic response to perceiving another person’s behavior, but, rather, this perception is integrated with the observer’s own repertoire, through shared neurocognitive networks. The content of this repertoire, which includes specific experiences and memories and which changes over time (i.e., adaptive mechanism and plastic system), is thus bound to affect the resonance mechanism.

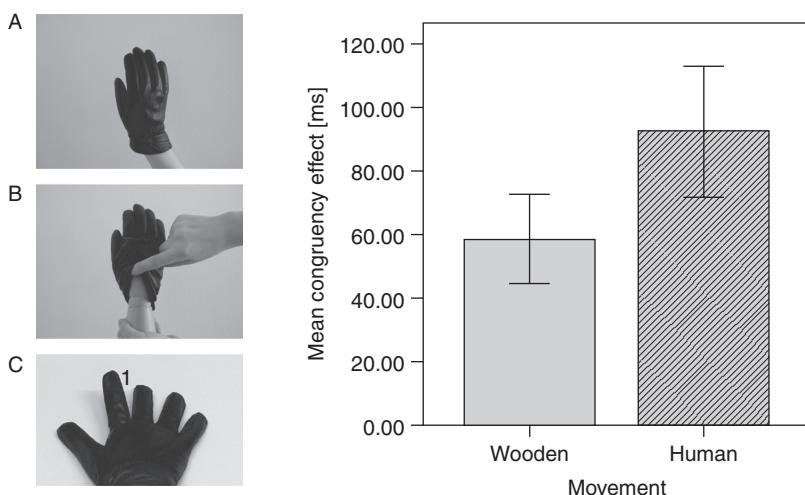


*What we know.* As presented in the previous section, the visual attributes of the observed individual are factors that can modulate motor resonance. However, there is a ton of information that can have an impact beyond the visual presentation of stimuli. For instance, simply knowing whether the source of an action is human or not (in two visually identical conditions) could influence the resonance system. A study by Liepelt and Brass (2010) recently showed that a motor priming effect (whereby reaction times of motor responses are faster when the movement used to respond is also observed) thought to reflect the level of resonance was larger in a group of participants who observed what they thought was a gloved human hand (animate object) performing an action as opposed to a group that looked at what they thought was a gloved wooden hand (inanimate object) performing the same action (Figure 8.2). Note that



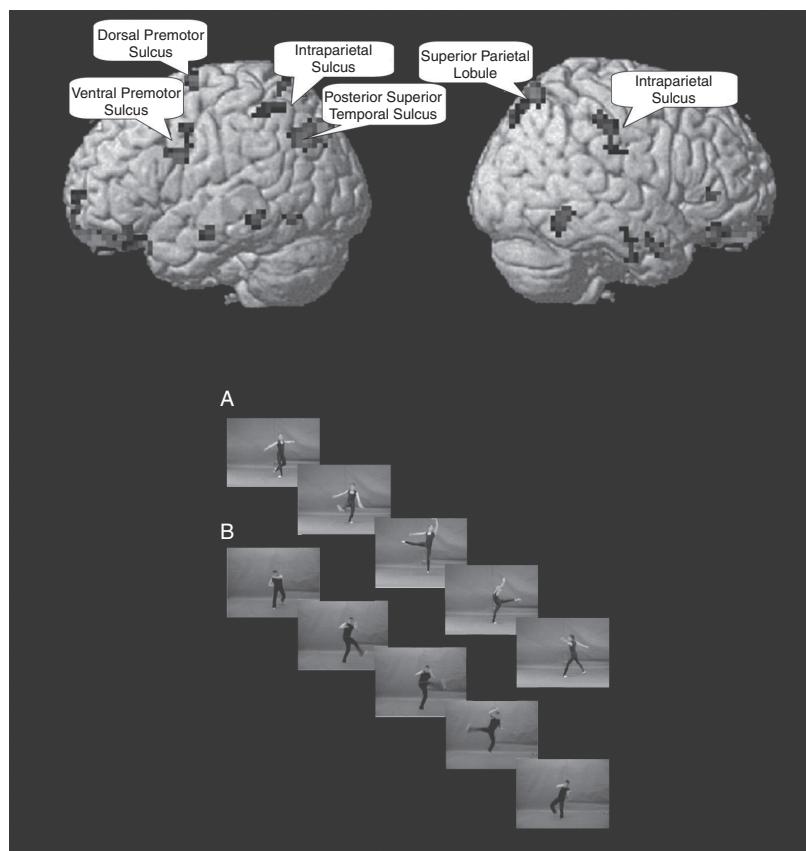
during the task the stimuli were visually the same, but before the experiment the authors manipulated the participants' belief as to the nature of the observed hand by showing one group a human hand and another a wooden hand in the glove. Hence, our previous knowledge/belief or other a priori factors may play an important role in how we react to movements made by another person.

Personal experiences or the level of expertise attained by an individual could be considered as such a priori factors, which could influence how the resonance system reacts to observed movements. This has been tested by Calvo-Merino and collaborators in a clever series of functional magnetic resonance imaging (fMRI) studies with expert dancers. In a first study, this group showed that expert dancers of different dance types (classical ballet or capoeira) showed greater activation in the whole resonance system, including premotor and parietal as well as temporal regions, when they observed the movements they were trained in, compared with the other dance type (Calvo-Merino



*Figure 8.2* In this experiment, the stimuli observed looked the same for all groups, as displayed in (C). However, prior to the experiment, the subjects were shown that the glove was put over either a human hand (A) or a wooden hand model (B). The graph on the right shows that the motor priming of the hand model was 34 ms larger in the group that thought the gloved hand contained a human hand compared with the group that thought it contained a wooden model. Error bars represent standard deviations of the mean values (Liepelt and Brass, 2010).

et al., 2005) (Figure 8.3). In a second study, they showed that the motor (kinesthetic) component of the expertise was more important than its visual counterpart. Indeed, male and female dancers frequently see every move from their dance type and are good at recognizing them



**AQS** *Figure 8.3* In an experiment that compared expert capoeira and ballet dancers with nonexpert participants, the pattern of activation shown on the left and right hemispheres of the [MNI template] displays the regions that are most related to the expertise of the observer. The motor expertise was expressed as the group (expert capoeira dancers vs. expert ballet dancers vs. nonexpert participants) by condition (previously learned vs. new movements) interaction. Expertise is related to changes in activity within the ventral, dorsal premotor, and intraparietal sulcus as well as the posterior superior temporal sulcus in the left hemisphere and the superior parietal lobule and intraparietal sulcus in the right. Examples for ballet (A) and capoeira (B) stimuli are shown in the lower part (adapted from Calvo-Merino et al., 2005).

200 Sébastien Hétu and Philip L. Jackson

(including highly gender-specific ones), but they are only proficient at performing the moves attributed to their gender. What the authors found was more activity in premotor, parietal, and cerebellar regions when dancers viewed moves from their own as opposed to the opposite gender (Calvo-Merino et al., 2006), supporting a highly motoric encoding contingent upon the participants' experience. A recent study has, however, found evidence that a mostly visual experience can also influence the response of the resonance system in health professionals. Fiorio and colleagues (Fiorio et al., 2010) showed that the corticospinal excitability (reflected by the motor evoked potentials in target muscles produced by TMS) of professionals who treat patients affected by dystonia (a movement disorder), and who are familiar with its motor signs, such as excessive co-contractions, was equivalent when they observed normal and dystonic movements. However, participants who were naïve to the dystonia condition showed hyperactivation when observing this condition as opposed to normal movements. The authors interpreted these findings as an influence of experience and knowledge causing a reduction (disengagement) of the normally strong response to pathological movement.

*Why we observe.* During our daily life we observe others in action for a wide range of motives: one can look at another conference attendee's movements in the hallway to predict the direction he will take and avoid social contact by avoiding him, but later during the conference one can also look at this same person's movements to interpret his or her mood and decide whether it would be a good time to greet him or her.

AQ6 In a much simpler version of this example, one study by Buddel and colleagues (Budell et al., 2010) used fMRI to examine the different patterns of cerebral activity when participants observed facial expressions of pain but for different reasons. Participants had to either rate the amount of pain experienced by the model (pain rating task) or determine whether facial movements were greater around the eye region or around the mouth region (movement task). As the stimuli were the same in both tasks, the only difference lay in the reason the participants had to watch (the "why"). Interestingly, activation within the inferior frontal gyrus (IFG), a region classically associated with motor resonance,

AQ7 was modulated by the task. Indeed, Buddel and colleagues showed that activity within the IFG was increased during the pain rating task as compared with the movement task and that the IFG activity was correlated with the pain ratings made by the participants. This suggests that the resonance system can be influenced not only by the motor but



also by the emotional dimensions of a stimulus, and that the level of activity in this system can change depending on the observer's objective. How information is processed within these regions is no doubt constrained by influences coming from frontal regions linked to executive functions as well as information from the limbic system, including memory and emotions.

*Who we are.* Another class of individual factors concerns the characteristics of the observer *per se*. Some early findings suggest that the gender of a participant can influence the level of response within the shared representation system (e.g. Cheng et al., 2006) and beyond (Proverbio et al., 2010). One potential confounding factor in earlier studies was the potential interaction between a gender effect and an effect due to the type of material used in the experiments (gender of the models; see Simon et al., 2006, for an example of an effect of the model's gender on the representation of pain). One study controlled for this and tested the change in the mu rhythm in male and female participants when they observed actions performed by either an androgynous, a male, or a female hand. The mu rhythm is a response in a specific bandwidth (typically 7 and 15 Hz depending on the source) of the electrical signal of the brain, typically measured at the scalp, that has been shown to be modulated in the context of action observation and could be a correlate of motor resonance (see Pineda and Hecht, 2008, for a discussion on the mu rhythm and its link to cognition). The authors showed that females had significantly stronger suppression of the mu rhythm than males during the observation of the moving hand conditions, but there was no indication that watching a same-sex or different-sex hand had an effect on the amount of suppression in female or male participants. Results from this study could be interpreted as showing how individual characteristics (in this case sex) can influence the neural response to an observed action.

Individual differences in the response to action observation (not related to sex) have also been identified in a recent transcranial magnetic stimulation (TMS) study (Hetuet et al., 2010). For instance, some individuals show strong effector-specific responses (increases in corticospinal excitability for muscles involved in the observed movement and not for muscles that are not used in the movement) while others do not display such preference, and this pattern of response is the same within an individual for various types of movements (movements made with proximal vs. distal upper-limb muscles).

*Problems at the periphery.* Experience can also be modulated through changes in the peripheral nervous system of individuals. These changes

AO8



can be sudden, as in the case of an amputation, but in other circumstances, rather than sudden changes, some individuals have birth defects at the periphery, which provide them with a different experience from that of most people. One case study with two aplastic participants, born without arms and hands, provides interesting information on the effect of experience (Gazzola et al., 2007b). This study examined, using fMRI, the response of the resonance system when these individuals observed hand, foot, and mouth actions. What they found was that activity in this system was similar to what was found in normally developed participants when the patients watched hand actions. This finding suggests that one does not need to have executed an action (or be able to) to activate the resonance system.

Many deaf people are exposed to a completely different set of movements than hearing people through the learning of sign language. To investigate whether lifelong experience with sign language in deaf individuals can influence the response to observing different types of communicative gestures, Emmorey and colleagues (Emmorey et al., 2010) had deaf signers and hearing individuals with no experience with sign language watch pantomimes and signed action verbs (ASL language) while recording their hemodynamic activity. Strikingly, while in hearing participants both pantomimes and signed verb movements produced increased activity in the shared motor representation network (i.e., fronto-parietal network), neither were associated with increased activity in deaf participants. Hence, it seems that extensive experience with manual communication can profoundly modify the neural response to action observation, even to the degree of completely abolishing the usual resonance response.

*Problems in the central nervous system.* Serino and colleagues (Serino et al., 2010) elegantly showed that motor impairments can influence how perceived actions are processed. They asked hemiplegic patients (i.e., patients unable to move one side of their body following a stroke), stroke patients without hemiplegia, and healthy controls to try to recognize pantomimes of object-directed movements or communicative gestures made with the right and left upper limbs. Movements were presented as point-light displays. This ensured that only the kinematics of the movements were available to the participants. Whereas healthy controls and stroke patients without hemiplegia were equally able to recognize movements made by right or left upper limbs, hemiplegic patients' performance decreased when they observed movements made by an upper limb of their affected side (e.g., patient with left hemiplegia

watching a movement of the left arm). Furthermore, this lateralized impairment was only present if movements were shown in the upright (real) position and not when stimuli were flipped on the horizontal axis (top becoming bottom). Indeed, for all participants such flipped movements were significantly more difficult to recognize, probably because they are not easily recognized as human movements. Therefore, the impairment seems to be limited to movements considered to have been produced by humans and not just random point-light movements, suggesting that the deficits are highly specific to the perception of human actions. These results highlight the tight link between motor-represented and visually represented movements, as impairments in one modality can lead to impairments in the other (motor deficits producing perceptual deficits in this case). For a more extensive discussion on how to investigate this network in neurological patients, see Andric and Small (2010).

*Transient changes in the system.* Cheng and colleagues (Cheng et al., 2007) found that motivational factors can modulate the response to observed behaviors. Using hunger as a motivational variable, the authors scanned healthy hungry or satiated participants while they watched reach to grasp motions towards food or non-food-related objects. Results showed that both types of reach to grasp motions activated the fronto-parietal network in satiated and hungry participants. However, only hungry participants had increased hemodynamic signal in the IFG, superior temporal gyrus, and superior parietal cortex during the movements made towards food when compared with movements made towards non-food objects. Furthermore, activity in the IFG of hungry participants was correlated with hunger/motivation to eat as measured on a visual analog scale. Outside the action–perception network, activity in regions that have been proposed to regulate goal-directed behaviors according to motivational needs of the organism (e.g., the orbitofrontal gyrus, the fusiform gyrus, and the hypothalamus) was increased during the observation of reach to grasp of food as compared with objects in hungry participants only. The change in pattern of brain activity related to the change in hunger state cannot be attributed to changes in the resonance system only, as other factors, such as attention to the stimuli, were not controlled for, but it seems clear that we can respond differently to the same stimuli depending on our physiological state.

Overall, the fact that cerebral changes in the observer linked to motor practice, motor deficits, or even transient motivational state can influence the response of the shared representation network during action

AQ9

observation further emphasizes the plastic nature of this more-than-perception-action coupling process. Direct testing of the plasticity of this system (e.g., Catmur et al., 2007) has even shown short-term malleability by training healthy young participants to produce one movement while watching another movement. This conflicting training over a period of several blocks of many trials (864 trials in total) was successful at reversing the “normal” pattern of response typically shown during action observation. More precisely, before training, participants showed increased excitability in the primary motor cortex (M1; measured with TMS) for the index finger when watching index finger movements and for little finger when watching little finger movements. After training in which participants executed little finger movements while watching index finger movements (and vice versa), the pattern of response during action observation was inverted: increased excitability for the little finger was observed when watching index movements. Participants who practiced and observed congruent movements did not show this reversal effect. If such a simple sensorimotor experience is able to modify the pattern of cerebral responses to action observation, it is very likely that repeated social interactions (in which we are necessarily observing many actions) also affect this motor matching process.

### The interaction between self- and other-related variables: the context

So far, the discussion has mainly addressed the changes in brain response related to what is observed or who is observing an action. While isolating these two components is essential to provide optimal control over experimental variables, real life situations are bound to create an interaction between these variables such that one individual will not respond the same way to all types of stimuli and all types of stimuli will not provoke the same response in all individuals.

Parent-child interactions are, from a child’s point of view, usually the first ones to emerge, and, from the parents’ point, arguably the strongest (emotionally speaking). Lenzi and colleagues used this interaction as a basis for their recent fMRI study examining the cerebral response to seeing and then imitating their own child’s (6 to 12 months old) facial expression compared with imitating another child’s (Lenzi et al., 2008). The findings show that the motor resonance system (IPL, IFG, vPMC), as well as the amygdala and insula, were more strongly activated by the interaction with their own child. This suggests that the resonance system is participant to differential levels of activity depending on the

nature of the relationship between a stimulus and the participant. In this case, the additional involvement of limbic regions is not surprising, due to the emotional bond between a mother and her child.

Along the same idea of relationship between an observer and a model, a new wave of studies on action observation have started to examine the cerebral basis that could underlie how individuals can perceive and respond differently to movements made by in-group members as compared with out-group members. This social psychology concept of in-group vs. out-group has been tested along different dimensions, including race and ethnicity, religion, gender, and even supporters of sport teams. In one example, Molnar-Szakacs and colleagues (Molnar-Szakacs et al., 2007) showed to Euro-American individuals several types of movements made by Euro-American or Nicaraguan models and used TMS to assess the modulating effect of action observation on primary motor cortex (M1) excitability. Movements were either neutral (i.e., modified American Sign Language signs) or culturally significant hand movements (i.e., movements of either Euro-American or Nicaraguan culture). Results showed that observation of movements made by Euro-American models, when compared with movements made by Nicaraguan models, was associated with a greater increase in M1 excitability in participants (Euro-Americans) when compared with movements made by Nicaraguan models, regardless of the type of movement. This prompted the authors to suggest that the way the brain responds during action observation is sensitive to in-group/out-group distinction. In this case, it is plausible that observing a member of our own cultural group moving enhances the response of the resonance system, which can in turn influence the primary motor cortex to facilitate the production of motor behaviors.

Using a similar approach, Désy and Théoret (Desy and Theoret, 2007) found that, instead of having an in-group bias during action observation, their participants responded (in terms of M1 excitability) more to the movements made by out-group models. In this case, in-group vs. out-group was represented by the skin color of the participants corresponding or not to the color of the models. These conflicting results may come from methodological differences between the two studies. Whereas Molnar-Szakacs et al. used stimuli in which the faces of models were visible, the stimuli used by Désy and Théoret only showed the hands of the models. As previous results had shown that seeing faces similar to one's own produced increased responses when compared with dissimilar faces (Keenan et al., 2001), it is possible that the presence of the model's face in the study by Molnar-Szakacs may have contributed

206 Sébastien Hétu and Philip L. Jackson

to the in-group bias they found. Also, while both females and males showed an in-group bias in Molnar-Szakacs's study, only the female participants of Désy and Théoret's study showed a higher resonance response to out-group models. This last point highlights the fact that action observation is not a stereotyped process in which all movements are identically represented in the observer's brain. Instead, results from Molnar-Szakacs and Désy and Théoret suggest that in-group/out-group differentiation, which takes into account characteristics of both the observer and the observed stimulus, can influence the level of activity within the resonance system. It is important to note that, altogether, these studies show that the activity in this system can change depending on many factors, but they do not confirm the etiology of this activity; they do not provide information as to whether the increase is due to a simulation mechanism, a change in the level of attention, or some other process active during social interaction.

The previous section has highlighted that factors linked to whom/what we observe, who we are, and the relation between these two can modify the activity within the resonance system. Some may argue that motor resonance has been mainly studied using nonsocial movements and therefore its involvement in social interaction remains to be shown. However, the fact that activity within brain regions thought to underlie motor resonance (i.e., the shared representation system) is modulated by social factors such as who we are, whom we are looking at, and our relationship suggests that, even if motor resonance cannot explain in itself the way I act when I meet Paul and Mary, it is likely that information processed within this system is used by other higher order systems during my many social interactions. One such system involved primarily in social context is the mentalizing system. It will be the focus of the next section of this chapter.

### Controlling shared representation, from motor resonance to motor cognition

The second part of this chapter will focus on the other end of the motor representation continuum: how top-down processes play a key role in providing the most accurate representation of other people's actions. One of the suggested processes has been mentalizing (the inference mechanism allowing the exploration of others' perspective), which makes possible decoupling (Frith and Frith, 2003) from the actually perceived actions to explore alternate hypotheses based on a comparison with stored information (memory). There are a number of

# PROOF

*The Neural Systems Involved in Motor Cognition and Social Contact* 207

excellent reviews on mentalizing and related processes (Burnett et al., 2010; Pelphrey and Carter, 2008; Van Overwalle, 2009) and we refer the reader to these, but we deemed it necessary to briefly address here the works of people who are investigating these processes from a motor representation standpoint.

AQ10

Though we have demonstrated thus far that the level of implication of the shared representation network is subject to modulations that can be self-related, other-related, or related to the interaction between these two sets of factors, it remains unclear whether awareness is necessary for these changes to occur. In fact, most of the findings described so far likely operate without the participant's full awareness and can be considered automatic. It is, however, possible in some cases to deliberately modulate the level of activity within the shared representation system. If the shared representation system leads the observer to evoke the mental representation of observed behavior, thereby facilitating subsequent behavior, there must be some parallel mechanism to prevent unnecessary and constant mirroring of other people's behavior. It has been proposed that the control of the shared representation is what ties in resonance mechanisms with mentalizing function. In a series of studies, Brass and collaborators have found common regions of activity in the anterior fronto-medial cortex and the temporo-parietal junction for tasks thought to measure the control of the motor resonance (through the inhibition of the tendency to imitate; e.g., lifting the index finger while observing a middle finger) and mental state attribution tasks (Brass et al., 2009). A subsequent study found that the participant's ability to control the influence of the resonance mechanism during imitation was related to behavioral and neural (anterior fronto-medial cortex) correlates of mentalizing (Spengler et al., 2009). The authors argue that such a mechanism could contribute to the bridging of shared representation with social cognition.

AQ11

In another study attempting to investigate this gap between the motor component of an action and its social relevance, Spunt et al. (2010) used videoclips of human actions in ecologically valid contexts. Using fMRI, the authors examined changes in neural activity while participants identified either the mental state of the model or his/her overt physical behavior. As predicted, the observation of action was associated with activation of the resonance system, but this activity was not modulated by the degree to which the participants attributed a mental state to the models. Instead, this latter variable was associated with changes in the cortex of the medial prefrontal, posterior cingulate, and temporal pole regions. These data add to the argument that other systems must be at

208 Sébastien Hétu and Philip L. Jackson

play, on top of those subserving the resonance mechanism, in order to infer about the mental state of other individuals and thus provide social relevance to motor behavior.

If we accept that different systems are at play when establishing a social contact, some more basic and more automatic, others more complex and controlled, it remains that observation of certain types of behaviors is likely to recruit more basic systems and that these systems remain more efficient at processing certain types of information. The more complex systems might be an evolutionary advantage for socialization, but other types of interaction with conspecifics might still require a primary involvement of more basic processes. For instance, producing movements in a competitive context might require that we rely more on reactive mechanisms than intentional ones, as suggested by Welchman and colleagues (Welchman et al., 2010). They showed that ballistic movements could be produced up to 10 per cent faster when executed in reaction to an opponent's move as opposed to being produced intentionally. This effect was not moderated by the fact that movement observed and executed did not match, or by whether the opponent was a human or a computer. Thus, when the speed of a movement is important, it appears that a reactive situation has advantages over volition-based ones (but see Kokal et al., 2009, for evidence that both means of response are necessary in collaborative actions). A similar pair of variables is likely at play during the observation of pain in others (e.g., Yamada and Decety, 2009), which, evolutionarily speaking, must have been associated with a need to flee or an alert signal. However, through socialization, a secondary prosocial empathic response has also emerged (Decety and Jackson, 2004). It seems, however, that, in the social context, the great advantage of engaging this latter system is to move from a reactive stance to a predictive and proactive one. The reader is referred to other work for a more extensive discussion on empathy (de Vignemont and Singer, 2006; Decety and Jackson, 2004; Decety et al., 2007; Fan et al.).

AQ12

### Conclusion: is motor cognition necessary for social contact?

When we meet and greet another individual, we perceive the immediate actions of the other, and the shared representation system plays an important part in this. However, in order to reach beyond the here and now, especially in new or unfamiliar situations, we will likely also put our mentalizing potential at play to complete our understanding of the

other's mood and intention, and adapt our response appropriately. Still, while the involvement of the mentalizing system in social interaction is widely accepted, because not all social contacts involve observed movements *per se* (for example, when we talk on our cellphone), the exact role of the resonance system is still debated.

The motor theory of social cognition, suggesting that the resonance mechanism and its shared representations play a central role in action understanding (Rizzolatti and Sinigaglia, 2010), language (Arbib, 2009), and ultimately social contacts (Oberman et al., 2007), has profoundly influenced our comprehension of human interactions. For example, deficiencies in the resonance mechanism are now frequently mentioned as being part of a wide range of neuropsychopathologies with social impairments, such as schizophrenia and autism spectrum disorder (ASD). However, if tenants of the motor theory have for a long time been the dominant speakers in the field of action observation and social interactions, new voices have lately arisen which have shaken the very core of the motor theory for social contacts. Indeed, one of the central claims of the motor theory supporters, namely, that action understanding is supported by the resonance mechanism, is now openly questioned (Heyes, 2009; 2010; Hickok, 2009; Jacob, 2008; 2009; Jacob and Jeannerod, 2005; Vivona, 2009).

Hickok mentions that actual evidence for the role of the mirror neuron system, and thus of motor resonance, in action understanding is still insufficient and could even contradict the motor theory of social cognition (Hickok, 2009). For example, Hickok makes the case that, in humans, there often are double dissociations between motor production deficits and action understanding deficits and that lesion and disruption studies focusing on regions of the resonance system (e.g., vPM/IFG) offer meager support for the role of these structures in action understanding. Lastly and importantly, Hickok emphasizes that several of the studies that tried to demonstrate the role of motor resonance in action understanding have used tasks that do not directly assess action understanding, such as delayed matching-to-sample tasks. Performance in such discrimination tasks, according to Hickok, "did not depend on understanding the meaning of the actions, but could be performed based on configural information alone" (p. 1,238).

In addition to the concerns raised by Hickok, Jacob (2009) openly questions the proposition that the motor resonance response directly evolved in humans to serve as the primary mechanism for action understanding. Instead, Jacob suggests that the resonance response observed in humans may be a by-product of an evolutionary pressure

# PROOF

210 Sébastien Hétu and Philip L. Jackson

that favored individuals with structures that underlie the conceptual nature of actions. Indeed, Jacob mentions that such structures would probably be active during both the production and the perception of actions. For example, brain regions active during both the observation and the production of grasping motion could constitute the neural basis of the *concept* of grasping. Having structures that hold conceptual representation of actions may have given an evolutionary advantage by helping individuals make predictions about others' behaviors, because observing a movement activates the concept (as opposed to the motor command) of this movement, making it possible for the individual to react in a proper way to others' actions. As so, motor resonance is considered not as the basis of action understanding but as a manifestation of the activity of neural structures containing conceptual representations of actions, and it is the ability to have and use these concepts that enables us to infer the intentions of others when perceiving their actions. Jacob (2008) further suggests that the role of the resonance mechanism is not to infer the intention of others from the activation of the representation of the motor command within the observer's brain but, rather, to infer the representation of the motor command used by the other individual from the representation of the other's prior intention within the observer's brain. This process, Jacob adds, could allow the observer to predict the next motor act of the other, conferring a predictive role (as opposed to understanding the immediate state) on the resonance mechanism during action observation.

As mentioned previously, one of the arguments for the role of motor resonance in social cognition is that the "ability" of certain neurons to be active during both action production and perception has evolved for the specific purpose of action understanding, and thus through a specific adaptation process. However, recent data, such as the seminal work of Catmur and colleagues (Catmur et al., 2007), which demonstrated that motor resonance responses can be modified through sensorimotor associative learning, questions this assumption. These findings have led some researchers to suggest that the resonance response and the specific neurons that underlie it (i.e., mirror neurons) are a product of associative learning (Heyes, 2010), in which experiences of actions being produced and perceived in a highly correlated manner enhance the probability of resonance response patterns in neurons (motor neurons that discharge during perception of motor act). As mentioned by Heyes (2009), because these sensorimotor experiences are most common during social interaction, the associative learning origin of mirror neurons implies that the motor resonance response is in [some way]

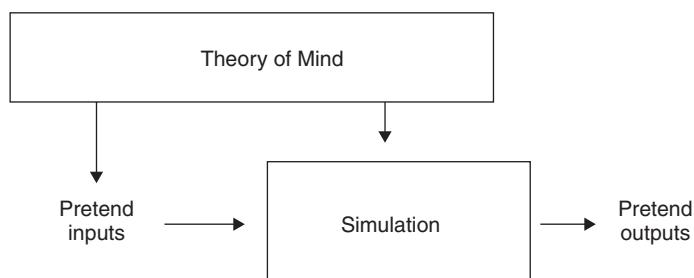
AQ13

a product of social interaction. Importantly, Heyes (2009) also mentions that, even if motor resonance is probably not specifically aimed at action understanding, it is well suited to contribute in part to several social cognition functions, leading her to say that “the mirror neuron system is a product, as well as a process, of social interaction” (p. 581).

In line with this reduced role of motor resonance during social interaction, Saxe (2005) previously suggested a model to explain errors people make when reasoning about other people. In her model motor resonance is not at the core, but is partially involved. It is, instead, the mentalizing functions that represent the main component of the model, and it is through the interaction between the mentalizing functions and the simulation functions (which are often associated with the motor resonance mechanism) that such errors are better explained (Figure 8.4).

The question as to whether, or to what extent, the resonance system is involved in social contact still remains. In an effort to better define the specific role of the resonance and mentalizing systems in understanding the goals of people we interact with and thus try to respond to this question, Van Overwalle and Kris Beatens recently conducted an impressive meta-analysis focusing on the fMRI literature on both systems (Van Overwalle and Baetens, 2009). They conclude that “the mirror and mentalizing systems are two distinct systems each specialized in the processing of observed sensory or verbal information about other persons but based on different inputs. The mirror system

AQ14



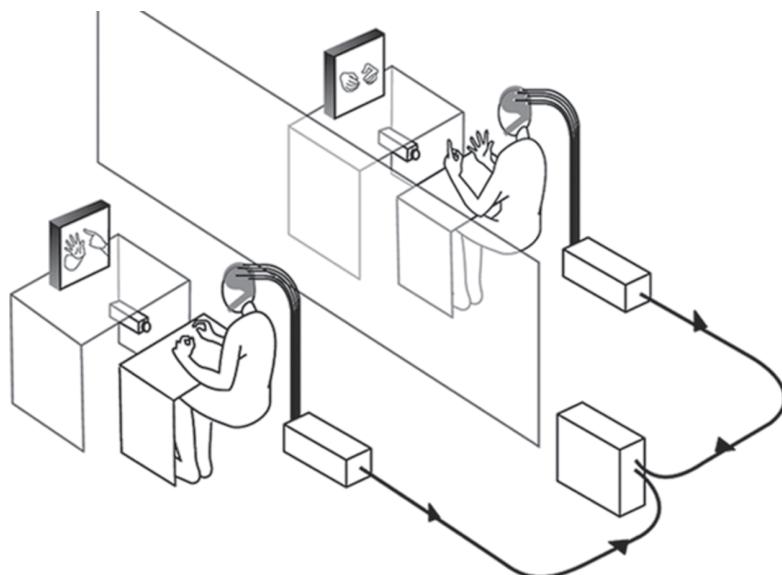
*Figure 8.4* One of the models proposed by Saxe (2005) to explain how humans make mistakes when thinking about the content of another person's mind. In this model, mentalizing/theory of mind and motor resonance/simulation functions interact. Mentalizing/theory of mind is here necessary to adjust the resonance/simulation process. Pretend inputs: what the observer perceives in the other. Pretend outputs: what the observer thinks the model is thinking (taken from Saxe, 2005).

212 Sébastien Hétu and Philip L. Jackson

is recruited when moving body parts are observed. The mentalizing system is recruited when no such input is available" (p. 579). We could add that the mentalizing system can also intervene if the observer is not paying attention to the motor output of the other. Results from this meta-analysis further point to separate systems that do not work together. However, as mentioned by the authors, research on action observation and/or social cognition has until recently mainly tried to use "pure" tasks in which each concept (e.g., action observation; social cognition) is separately studied.

One can argue that this highly focalized approach is bound to produce results showing that only one system is active during a specific type of task or is recruited when processing a specific type of stimuli. Such "artificial" tasks are necessarily much too reductive to help in addressing the core question: during everyday social interactions, when both types of stimuli are available, such as when someone raises his fist and shouts at you, is the resonance mechanism active? Therefore, future work is needed using ecological paradigms in order to better "see" what the brain actually does when I meet Paul and Mary at a conference. Such a paradigm could take advantage of novel research tools, such as hyperscanning, where two individuals' neural activity can be simultaneously recorded (Dumas et al., 2010; Lindenberger et al., 2009; Montague et al., 2002) (see Figure 8.5 for an example).

The discovery of mirror neurons in the brain of nonhuman primates in the 1990s has sparked a huge interest in how shared representations (motor-perceptive in this particular case) can be used by humans. Twenty years have passed since this discovery, during which it seems that the scientific world has seen both the rise of a "motor theory of everything social" and utter rejection of any implication of the resonance response during social contact. Sadly, such polarized positions have forced many to take position "for" or "against" motor resonance in social cognition. The truth (as scientists we keep searching for it) probably lies between these two points of view. As we have highlighted in the first part of this chapter, because the resonance system is plastic and can be influenced by variables linked to who we are, whom we observe, and the nature of our social relations it would seem odd to completely exclude its role during social contact; even more so as movements have such importance during human interactions. However, we have also highlighted that the resonance system alone cannot support social cognition function, and that other brain structures, such as those involved in mentalizing, are necessarily at play when we socially interact. In fact, from the current state of knowledge on social cognition, it seems that, even if the



**Figure 8.5** Schematic representation of the set-up used in the Dumas et al. (2010) paper which allows simultaneous scanning of the brain activity of two participants with EEG while they perform a motor interaction task.

resonance system is neither sufficient nor necessary to explain social interaction, the very nature of social interactions, in which movements have such an important place, can only lead us to consider a social cognition without any motor cognition to be improbable.

### Acknowledgements

The authors are grateful for the contribution of the Canadian Institute for Health Research made through a salary grant to PLJ and a scholarship to SH.

### References

- Andric, M., and Small, S. L. (2010). Functional imaging of putative human mirror neuron systems in neurological disease. *Exp Neurol*, 221, no. 1, 5–9.
- AQ15 Arbib, M. A. (2009). Mirror system activity for action and language is embedded in the integration of dorsal and ventral pathways. *Brain Lang*.
- Brass, M., Bekkering, H., Wohlschlager, A., and Prinz, W. (2000). Compatibility between observed and executed finger movements: comparing symbolic, spatial, and imitative cues. *Brain Cogn*, 44, no. 2, 124–43.

- Brass, M., Ruby, P., and Spengler, S. (2009). Inhibition of imitative behaviour and social cognition. *Philos Trans R Soc Lond B Biol Sci*, *364*, no. 1,528, 2,359–67.
- Budell, L., Jackson, P., and Rainville, P. (2010). Brain responses to facial expressions of pain: emotional or motor mirroring? *Neuroimage*, *53*, no. 1, 355–63.
- AQ16** Burnett, S., Sebastian, C., Cohen Kadosh, K., and Blakemore, S. J. (2010). The social brain in adolescence: Evidence from functional magnetic resonance imaging and behavioural studies. *Neurosci Biobehav Rev*.
- Calvo-Merino, B., Glaser, D. E., Grezes, J., Passingham, R. E., and Haggard, P. (2005). Action observation and acquired motor skills: an fMRI study with expert dancers. *Cereb Cortex*, *15*, no. 8, 1,243–9.
- Calvo-Merino, B., Grezes, J., Glaser, D. E., Passingham, R. E., and Haggard, P. (2006). Seeing or doing? Influence of visual and motor familiarity in action observation. *Curr Biol*, *16*, no. 19, 1,905–10.
- Carr, L., Iacoboni, M., Dubeau, M. C., Mazziotta, J. C., and Lenzi, G. L. (2003). Neural mechanisms of empathy in humans: a relay from neural systems for imitation to limbic areas. *Proc Natl Acad Sci U S A*, *100*, no. 9, 5,497–502.
- Catmur, C., Walsh, V., and Heyes, C. (2007). Sensorimotor learning configures the human mirror system. *Curr Biol*, *17*, no. 17, 1,527–31.
- Cattaneo, L., and Rizzolatti, G. (2009). The mirror neuron system. *Arch Neurol*, *66*, no. 5, 557–60.
- Cheng, Y., Meltzoff, A. N., and Decety, J. (2007). Motivation modulates the activity of the human mirror-neuron system. *Cereb Cortex*, *17*, no. 8, 1,979–86.
- Cheng, Y. W., Tzeng, O. J., Decety, J., Imada, T., and Hsieh, J. C. (2006). Gender differences in the human mirror system: a magnetoencephalography study. *Neuroreport*, *17*, no. 11, 1,115–9.
- Decety, J., and Jackson, P. L. (2004). The functional architecture of human empathy. *Behav Cogn Neurosci Rev*, *3*, no. 2, 71–100.
- Decety, J., Jackson, P. L., and Brunet, E. (2007). The cognitive neuropsychology of empathy. In T. F. F. P. W. Woodruff (Ed.), *Empathy in Mental Illness and Health*. New York: Cambridge University Press.
- Desy, M. C., and Theoret, H. (2007). Modulation of motor cortex excitability by physical similarity with an observed hand action. *PLoS ONE*, *2*, no. 10, e971.
- de Vignemont, F., and Singer, T. (2006). The empathic brain: how, when and why? *Trends Cogn Sci*, *10*, no. 10, 435–41.
- Dewey, J. (1896). The reflex arc concept in psychology. *Psychological Review*, *3*, 357–70.
- di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., and Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study. *Exp Brain Res*, *91*, no. 1, 176–80.
- Dumas, G., Nadel, J., Soussignan, R., Martinerie, J., and Garnero, L. (2010). Interbrain synchronization during social interaction. *PLoS One*, *5*, no. 8.
- Emmorey, K., Xu, J., Gannon, P., Goldin-Meadow, S., and Braun, A. (2010). CNS activation and regional connectivity during pantomime observation: no engagement of the mirror neuron system for deaf signers. *Neuroimage*, *49*, no. 1, 994–1,005.
- Fabbri-Destro, M., and Rizzolatti, G. (2008). Mirror neurons and mirror systems in monkeys and humans. *Physiology (Bethesda)*, *23*, 171–9.
- Fan, Y., Duncan, N. W., de Grecq, M., and Northoff, G. Is there a core neural network in empathy? An fMRI based quantitative meta-analysis. *Neurosci Biobehav Rev*, *35*, no. 3, 903–11.

# PROOF

Fecteau, S., Carmant, L., Tremblay, C., Robert, M., Bouthillier, A., and Theoret, H. (2004). A motor resonance mechanism in children? Evidence from subdural electrodes in a 36-month-old child. *Neuroreport*, 15, no. 17, 2, 625–7.

AQ17 Fiorio, M., Cesari, P., Bresciani, M. C., and Tinazzi, M. (2010). Expertise with pathological actions modulates a viewer's motor system. *Neuroscience*.

Frith, C. D., and Frith, U. (2006). The neural basis of mentalizing. *Neuron*, 50, no. 4, 531–4.

Frith, U., and Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philos Trans R Soc Lond B Biol Sci*, 358, no. 1, 431, 459–73.

Gallese, V., Fadiga, L., Fogassi, L., and Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, 119 (Pt 2), 593–609.

Gazzola, V., Rizzolatti, G., Wicker, B., and Keysers, C. (2007a). The anthropomorphic brain: the mirror neuron system responds to human and robotic actions. *Neuroimage*, 35, no. 4, 1, 674–84.

AQ18 Gazzola, V., van der Worp, H., Mulder, T., Wicker, B., Rizzolatti, G., and Keysers, C. (2007b). Aplastics born without hands mirror the goal of hand actions with their feet. *Curr Biol*, 17, no. 14, 1, 235–40.

Heberlein, A. S., and Adolphs, R. (2004). Impaired spontaneous anthropomorphizing despite intact perception and social knowledge. *Proc Natl Acad Sci U S A*, 101, no. 19, 7, 487–91.

Heberlein, A. S., Adolphs, R., Tranel, D., and Damasio, H. (2004). Cortical regions for judgments of emotions and personality traits from point-light walkers. *J Cogn Neurosci*, 16, no. 7, 1, 143–58.

AQ19 Hetu, S., Gagné, M., Jackson, P. L., and Mercier, C. (2010). Variability in the effector-specific pattern of motor facilitation during the observation of everyday actions: implications for the clinical use of action observation. *Neuroscience*.

Hétu, S., Mercier, C., and Jackson, P. L. (2010). *Effect of visual perspective and transitivity/meaning on the action observation network*. Paper presented at the 6th Annual Meeting of the Organization for Human Brain Mapping.

AQ20 Heyes, C. (2009). Where Do Mirror Neurons Come From? *Neurosci Biobehav Rev*.

AQ21 Heyes, C. (2010). Mesmerising Mirror Neurons. *Neuroimage*.

AQ22 Hickok, G. (2009). Eight Problems for the Mirror Neuron Theory of Action Understanding in Monkeys and Humans. *J Cogn Neurosci*.

Jackson, P. L., Meltzoff, A. N., and Decety, J. (2005). How do we perceive the pain of others? A window into the neural processes involved in empathy. *Neuroimage*, 24, no. 3, 771–9.

Jacob, P. (2008). What do mirror neurons contribute to human social cognition? *Mind and Language*, 23, no. 2, 190–223.

Jacob, P. (2009). The tuning-fork model of human social cognition: a critique. *Conscious Cogn*, 18, no. 1, 229–43.

Jacob, P., and Jeannerod, M. (2005). The motor theory of social cognition: a critique. *Trends Cogn Sci*, 9, no. 1, 21–5.

James, W. (1890). *Principles of Psychology*. New York: Henry Holt.

Jeannerod, M. (2001). Neural simulation of action: a unifying mechanism for motor cognition. *Neuroimage*, 14, 1 Pt 2, S103–9.

Keenan, J. P., Nelson, A., O'Connor, M., and Pascual-Leone, A. (2001). Self-recognition and the right hemisphere. *Nature*, 409, no. 6, 818, 305.

Keysers, C., Wicker, B., Gazzola, V., Anton, J. L., Fogassi, L., and Gallese, V. (2004). A touching sight: SII/PV activation during the observation and experience of touch. *Neuron*, 42, no. 2, 335–46.

# PROOF

216 Sébastien Hétu and Philip L. Jackson

Kilner, J. M., Paulignan, Y., and Blakemore, S. J. (2003). An interference effect of observed biological movement on action. *Curr Biol*, 13, no. 6, 522–5.

Knutson, K. M., McClellan, E. M., and Grafman, J. (2008). Observing social gestures: an fMRI study. *Exp Brain Res*, 188, no. 2, 187–98.

AQ23 Kokal, I., Gazzola, V., and Keysers, C. (2009). Acting Together in and beyond the Mirror Neuron System. *Neuroimage*.

AQ24 Lenzi, D., Trentini, C., Pantano, P., Macaluso, E., Iacoboni, M., Lenzi, G. L., et al. (2008). Neural Basis of Maternal Communication and Emotional Expression Processing during Infant Preverbal Stage. *Cereb Cortex*, 19, 1,124–33.

Liepelt, R., and Brass, M. (2010). Top-down modulation of motor priming by belief about animacy. *Exp Psychol*, 57, no. 3, 221–7.

Lindenberger, U., Li, S. C., Gruber, W., and Muller, V. (2009). Brains swinging in concert: cortical phase synchronization while playing guitar. *BMC Neurosci*, 10, 22.

Molnar-Szakacs, I., Wu, A. D., Robles, F. J., and Iacoboni, M. (2007). Do you see what I mean? Corticospinal excitability during observation of culture-specific gestures. *PLoS ONE*, 2, no. 7, e626.

AQ25 Montague, P. R., Berns, G. S., Cohen, J. D., McClure, S. M., Pagnoni, G., Dhamala, M., et al. (2002). Hyperscanning: simultaneous fMRI during linked social interactions. *Neuroimage*, 16, no. 4, 1,159–64.

Montgomery, K. J., Isenberg, N., and Haxby, J. V. (2007). Communicative hand gestures and object-directed hand movements activated the mirror neuron system. *Soc Cogn Affect Neurosci*, 2, no. 2, 114–22.

Oberman, L. M., Pineda, J. A., and Ramachandran, V. S. (2007). The human mirror neuron system: A link between action observation and social skills. *Soc Cogn Affect Neurosci*, 2, no. 1, 62–6.

Pelphrey, K. A., and Carter, E. J. (2008). Brain mechanisms for social perception: lessons from autism and typical development. *Ann N Y Acad Sci*, 1,145, 283–99.

AQ26 Pineda, J. A., and Hecht, E. (2008). Mirroring and mu rhythm involvement in social cognition: Are there dissociable subcomponents of theory of mind? *Biol Psychol*.

Prinz, W. (1997). Perception and action planning. *European Journal of Cognitive Psychology*, 9, 129–54.

AQ27 Proverbio, A. M., Riva, F., and Zani, A. (2010). When neurons do not mirror the agent's intentions: Sex differences in neural coding of goal-directed actions. *Neuropsychologia*.

Rizzolatti, G., and Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations. *Nat Rev Neurosci*, 11, no. 4, 264–74.

Rizzolatti, G., Fadiga, L., Gallese, V., and Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Brain Res Cogn Brain Res*, 3, no. 2, 131–41.

Rizzolatti, G., Fabbri-Destro, M., and Cattaneo, L. (2009). Mirror neurons and their clinical relevance. *Nat Clin Pract Neurol*, 5, no. 1, 24–34.

Saarela, M. V., Hlushchuk, Y., Williams, A. C., Schurmann, M., Kalso, E., and Hari, R. (2007). The compassionate brain: humans detect intensity of pain from another's face. *Cereb Cortex*, 17, no. 1, 230–7.

Saxe, R. (2005). Against simulation: the argument from error. *Trends Cogn Sci*, 9, no. 4, 174–9.

# PROOF

- [AQ28] Schurmann, M., Hesse, M. D., Stephan, K. E., Saarela, M., Zilles, K., Hari, R., et al. (2005). Yearning to yawn: the neural basis of contagious yawning. *Neuroimage*, 24, no. 4, 1,260–4.
- Serino, A., De Filippo, L., Casavecchia, C., Coccia, M., Shiffrar, M., and Ladavas, E. (2010). Lesions to the motor system affect action perception. *J Cogn Neurosci*, 22, no. 3, 413–26.
- Simon, D., Craig, K. D., Miltner, W. H., and Rainville, P. (2006). Brain responses to dynamic facial expressions of pain. *Pain*, 126, no. 1–3, 309–18.
- [AQ29] Spengler, S., von Cramon, D. Y., and Brass, M. (2009). Control of shared representations relies on key processes involved in mental state attribution. *Hum Brain Mapp*.
- Sperry, R. W. (1952). Neurology and the mind-body problem. *American Scientist*, 40, 291–312.
- [AQ30] Sputz, R. P., Satpute, A. B., and Lieberman, M. D. (2010). Identifying the What, Why, and How of an Observed Action: An fMRI Study of Mentalizing and Mechanizing during Action Observation. *J Cogn Neurosci*.
- Van Overwalle, F. (2009). Social cognition and the brain: a meta-analysis. *Hum Brain Mapp*, 30, no. 3, 829–58.
- [AQ31] Van Overwalle, F., and Baetens, K. (2009). Understanding others' actions and goals by mirror and mentalizing systems: A meta-analysis. *Neuroimage*.
- [AQ32] Villarreal, M., Fridman, E. A., Amengual, A., Falasco, G., Gerscovich, E. R., Ulloa, E. R., et al. (2008). The neural substrate of gesture recognition. *Neuropsychologia*, 46, no. 9, 2,371–82.
- Vivona, J. M. (2009). Leaping from brain to mind: a critique of mirror neuron explanations of countertransference. *J Am Psychoanal Assoc*, 57, no. 3, 525–50.
- [AQ33] Welchman, A. E., Stanley, J., Schomers, M. R., Miall, R. C., and Bulthoff, H. H. (2010). The quick and the dead: when reaction beats intention. *Proc Biol Sci*.
- Wicker, B., Keysers, C., Plailly, J., Royet, J. P., Gallese, V., and Rizzolatti, G. (2003). Both of us disgusted, in My insula: the common neural basis of seeing and feeling disgust. *Neuron*, 40, no. 3, 655–64.
- [AQ34] Yamada, M., and Decety, J. (2009). Unconscious affective processing and empathy: An investigation of subliminal priming on the detection of painful facial expressions. *Pain*.



## 9

Action and Cephalic Expression:  
Hermeneutical Pragmatism*Jay Schulkin and Patrick Heelan***Introduction**

Our age is one in which we hope to preserve “reason” without over-exaggerated “rationalism.” Our epistemic stances are related to a particular social and historical context and community; the community shares this context through the functioning of a common spoken and written language and other public systems of representation and communication, such as the use of numbers, diagrams, pictures, gestures, and so on. In these are encoded the horizons of real and possible facts, networks of facts, circumstances, outcomes, and practices relevant to the “world”/“lifeworld” and the goals of the community in question. Earlier chapters in the book have set out the biological, social, psychological, neural, and philosophical perplexities. The running theme is a sense of philosophy being continuous with science without pernicious scientism.

The pursuit of factual truth, say, by experimentation, understands that conceptual and theoretical hypotheses are necessary, but that regularly many are tried and rejected. Thus, the experimental truth is characterized by an appropriate sense of prudence and humility. Human knowledge intends the common good of the community. What binds its members regularly as a community are the practices and rituals of common life, which they hope constantly to improve and renew.

Our knowledge is supported by diverse cognitively related physical systems, especially in the brain, which have a long gestational developmental trajectory from which emerge embodied cognitive processes, including habits of self-corrective inquiry and trial-and-error learning as well as many other less noble competing processes. A normative goal for embodied cognitive processes is to create public human scales or

AQ1

# PROOF

Action and Cephalic Expression 219

genres – for example, theoretical, descriptive, analogical, mathematical, or aesthetic – to provide and foster contexts of insight for ourselves and others, and to improve our relationships with fellow citizens of our community. What we share with one another is – or can be – shared pervasively within our community; that is, it can be communicated directly to anyone using the media of communication available in the community's lifeworld (Dewey, 1925/1989; Merleau-Ponty, 1962).

A note on the terms "world," "lifeworld," and "reality:" traditionally, in classical science, the term "world" means "what exists independently of culture and history." Since the work of Thomas Kuhn (1962), however, science has come to be seen as a historical human culture; this is the way the writers of this paper have come to see it. Consequently, in this paper, the term "world" means "lifeworld." The shared experience of life is tied to a "lifeworld" or "reality" in which we humans participate. Our participation in our "lifeworld" defines who we are as a community that shares knowledge and what our human life means. The writers of this paper are trying to bridge both phenomenology and pragmatism. The terms "world" and "lifeworld" generally cover both the phenomenology of human transcendental and existential consciousness of the "lifeworld," which is *a priori* to human worldly action, and the goal-directed pragmatism of direct *a posteriori* human worldly action embedded in a shared sense of experience.

In this chapter, we begin with a discussion of a shared sense of experience, followed by a discussion of "embodied" action, followed by a discussion of "concept," "hypothesis," or "theory" formation, and "inquiry." Problem-solving and statistical inferences are grounded in human physical action in the world. We also take note of the function of cognitive neuroscience, of anticipatory cephalic systems linked to cortical expansion and expectations. Finally, we end the chapter with a sense of hermeneutical pragmatism and self-corrective inquiry.

## Shared sensibility and action: pragmatism and hermeneutics

The classical pragmatists (e.g., Peirce, James, Dewey, Mead) went to great lengths to distinguish their sense of experience from that of the seventeenth and eighteenth-century empiricism of Locke and Hume, which was not about human activity in the lifeworld, but about the association of sensations. For James, as for other pragmatists, experience was active, structured by ideas and tested by actions, and aimed functionally at broadening lifeworld horizons by informing

# PROOF

220 *Jay Schulkin and Patrick Heelan*

them with new ideas; experience was not just a passive reception of a narrowly focused world of, say, sensory impressions. James's radical empiricism [James, 1912/1996] is a kind of functionalism based on the search for adaptive behaviors through instrumental action, that is, an inquiry as the search to secure stability amidst ever-changing physical circumstances.

AQ2 James understood that traditional empiricism was too passive, and that is why he associated himself with something he called "radical empiricism." "Radical empiricism" moves beyond simple sensation; as James put it, "no one ever had a simple sensation by itself" (James, 1890/1952, vol. 1, p. 234).

AQ3 Pragmatists took inquiry to stem from a breakdown of a form of life (see also Heelan and Schulkin, 1998; Heidegger, 1927/1962) that had been replete with functional coherence. Frustrated with the doubt and insecurity that follows this breakdown, inquiry begins as the search to recover the sense of understanding our lost place in the world. For Peirce, "abduction" with its associated inductive and deductive activities was the generative process of new ideas that aimed to restore a new, secure and settled form of life and new stable patterns of action.

Dewey shared the view that the origin of inquiry is in the precariousness of human existence and the countless searches to recapture life's lost equilibrium. In the pursuit of inquiry, theories were to guide actions, and feedback from actions was to correct theories. In Dewey's naturalistic evolutionary vision of life, the twin cardinal poles of human action were strife and resolution.

As adaptive systems, we function largely in anticipation of events, embodied in habits of expectation coded hermeneutically in cephalic systems that are mobilized automatically to address the current occasion. In fact, as James noted, "we actually ignore most of the things before us" (1890/1952, vol. 1, p. 284). Attention is pivoted on what matters (Desimone et al., 1984; Pashler, 1998). Competition between needs and resources, between information capabilities and environmental context, is the recurring rule in a radical experience which engages the world, not as a spectator viewing a sensation and then making judgment, but when attention and judgment are required as part of the doing, part of the functionally engaged experience.

James (1890/1952) understood that perceptual systems, and attention to context and circumstance, constitute the functional background of the all-pervasive shared social space, and include diverse practical systems of action "ready to hand" for action (Heidegger, 1927/1962). One result in modern terms is a virtual world of which we are an embedded

# PROOF



Action and Cephalic Expression 221

AQ4 part (Clark, 1998; Menary, 2010; Noe, 2004; Rowlands, 2010; as also described in many chapters in this book).

A central message of classical pragmatism is that inquiry is a functional mode of action in the world of doing, in contrast with the cognitively sterile view that answers are to be researched in a museum of settled knowledge. To pursue this metaphor, since science and philosophy are not museum pieces, they share, as it were, a common interactive space and can influence each other across disciplinary boundaries without one absorbing or being reduced to the other.

Hermeneutics is oriented towards meaning and towards the things – signs, symbols, actions – that can be construed as having meaning. These comprise, for example, relics of past events, social institutions, religious myths and rituals, cosmological and natural phenomena, and cultural artifacts, not just in the domain of the arts and sciences, but, most particularly, in the spoken word and written texts. All of the things just mentioned enter into public awareness endowed with some – though usually ambiguous – meaning. Many of these symbolic vehicles of meaning are clearly human artifacts, but others are natural or cosmological phenomena that seem at first sight to belong totally to a nonhuman realm. On deeper scrutiny, however, the meanings they carry also turn out to have a constitutive hermeneutic dimension capable of calling into question any hard and fast distinction between nature and culture.

Inquiry for Husserl and Heidegger – as, incidentally, also for Peirce and Dewey – begins when some real expectation based in past experience fails and we are curious as to why, and look for an answer that will enable us to fulfill our failed expectation, or, failing this, to go around the problem, or alternatively to transform the context, reassessing our goals if need be (see Heelan and Schulkin, 1998). For Husserl (1969; 1973), eidetic phenomenological analysis explores the invariant boundaries of an imagined experience that is subjected to imagined variations of approach. Both see inquiry as connected with a breakdown of intelligibility – for Heidegger when action fails in the world, for Husserl when the noetic structure of the imagination fails. Husserl's approach is more logical, conceptual, and abstract, while Heidegger's is more existential, historical, and action-oriented. It is, then, to Heidegger's philosophy that we will turn almost exclusively for an account of hermeneutical philosophy.

For the limited purpose of this inquiry, it is useful to enter Heidegger's philosophy through *Being and Time*, first drawing down from his analysis of equipment to lay out the methodology of the hermeneutical circle



# PROOF

222 *Jay Schulkin and Patrick Heelan*

(Heidegger, 1927/1962). Here, he illustrates the genesis and process of any inquiry by focusing on what happens when, in the middle of a task, a tool – say, a hammer – breaks. A hammer, a tree, a text, an atom, all are recognized by their characteristic function in the lifeworld; each may fail on some occasion – perhaps even systematically – to fulfill its expected function, and such a failure would initiate a process of inquiry.

To start with, the philosophical background: the breakdown of the task initiates inquiry within the lifeworld. The human inquirer is Da-sein – “There-to-be” – the embodied understanding of “to be” or “to exist.” Inquiry is awakened when Da-sein poses a directed question (*Vorsicht*) to a problem or query (*Vorhabe*), which, like all directed questions, already implicitly contains an outline of a search and discovery strategy that aims at finding a solution to the problem or an answer to the query. The question so construed in this case is not yet articulated; only later will it achieve an adequate expression in apophantic discourse as (what philosophers of science call) an “explanation.”

There follows an active dialogue between *Vorsicht* and *Vorhabe*, accompanied by actions seeking practical fulfillment in the awareness that the sought-for understanding (*die Sachselbst*) has presented itself and made itself manifest to the inquirer (*Vorgriff*). If, on first trial, the sought-for understanding is absent, something nevertheless has been learnt, and the search resumes, dipping again into the available resources of *Vorhabe*, *Vorsicht*, and *Vorgriff*. This circle of inquiry, called the “hermeneutical circle” – sometimes called the “hermeneutical spiral” – is repeated until a solution presents itself within a new cultural praxis in the lifeworld. Only at that time is it in order to express the solution linguistically in statements, that is, in apophantic form.

Let us now apply the hermeneutical circle to the problem of the broken hammer. In order to finish the job on hand, we draw on our background understanding of the lifeworld, and ask ourselves, perhaps for the first time in our lives, what kind of thing is a hammer, what specifications does it have. The question itself suggests strategies for searching and finding. A dialogue is initiated that has theoretical and practical dimensions. First, we aim at defining the theoretical specifications of a hammer, then we look for or construct something that fulfills these specifications, and, when we have it, we try it out. Does it work? If it does the job – if our first attempt at defining the theoretical specifications is fulfilled in practical experience – we are aware of the presence of what we were looking for, something with which to finish this job.

# PROOF

Action and Cephalic Expression 223

We may still need a new hammer, but, for the moment, the job can go on. If, however, the trial fails, a thought may intervene: "We may have to revise our goals!" But no! We revise the theoretical specifications in the light of the previous outcome and try again, modifying the conditions of the experimental trial if necessary. This phase may be repeated several times until we have a theory and practice that work. If we experience nothing but failure, we reassess our options, for example, getting the job done in a different way; hiring a carpenter, or turning to a different technology, or, finally, we could just fold our tent for the time being.

This philosophic-methodological process of inquiry is hermeneutical because it is a search for explanatory meaning – meaning to be fulfilled in a new cultural praxis in the lifeworld. The process has a cyclic pattern which is repeated over and over, from the general background understanding of the lifeworld, to proposed theory, to trial experience, or, summarily, the circle of background, theory, and experience. Every rational inquiry then moves in a forward spiral aiming at fulfillment in a solution made present in experience as governed by a conscious, if revisable, goal.

Indeed, when expectations are thwarted, a broad array of learning occurs through new problem-solving search principles. This is close in scope to Peirce's (1878) and Heidegger's view of inquiry and the development of new solutions to problems. Of course, inquiry is more than this. An important discovery was that not only is there a set of learning equations that is not coupled to contingencies, but time of occurrence is not an axiomatic factor in learning *per se*, but rather for predicting events [Rescorla and Wagner, 1972]. The gap of information, that is, the measure of the ability of the starting point to communicate with its environment – including the observer – is both relative and absolute, and leads to a possible understanding of curiosity and the joys that one might find in the search for and attainment of instruments of media of communication (technically called "information"). The relationship of the communications media (or information gap) with curiosity has been articulated through the use of what is often called "information theory's entropy coefficient" and is about the measurable ability of a communications medium to carry messages (Loewenstein, 1994; 1996). An emphasis on visceral input is something vital to curiosity, learning, and inquiry, an embodied sense within the lifeworld. At the heart of the pragmatists or a hermeneutical perspective was a disruption from expected outcomes resulting in new inquiry, new searches and investigations (Heelan and Schulkin, 1998).

AQ5

## Hypotheses, background practices, and action

Indeed, we know that human organization is replete with anticipatory cognitive systems, most of which encompass the vast cognitive unconscious (Berthoz, 2002; Rozin, 1976; Schulkin, 2007); action sequences are well orchestrated, embedded in successful survival for both short and longer-term expression (Berridge, 2004; Gallistel, 1980; 1992).

AQ6 When Peirce (1868) noted correctly that “it was impossible to know intuitively that a given cognition is not determined by a previous one,” he was close to recognizing that cognitive resources figure in the organization of action, something that, some 30 years later, Dewey would note in his critique on the “reflex arc” (see Parrott and Schulkin, 1993), now greatly appreciated (Berthoz and Petit, 2006; Jackson and Decety, 2004; Lakoff and Johnson, 1999). The “ideo-motor” conception was something that James grappled with, and is also very much tied to modern behavioral neuroscience (Berthoz, 2002; Lundtz and Jeannerod, 1997; 1999; Schulkin, 2007).

AQ7 Of course, action is often habitual routine. Memory, attention, and other cognitive resources are minimized during diverse routines; cognitive capabilities are then recruited elsewhere in the ongoing action; it is the breakdown that helps generate further action and cognitive resources to learn, attend, and construe new resolutions and new forms AQ8 of adaptations (e.g., Peirce, 1878; 1892; Tennenbaum et al., 2011) and part of the clarity to determine the diverse consequences of the new set AQ9 of hypotheses determinant in human action.

AQ10 Part of his purpose was to establish warranted beliefs and legitimated habits of action, routines towards goals that serve diverse purposes; science, like conduct, is established by expectations. These expectations are not reduced to simple sensations that legitimate the action; detecting noticeable difference in sensation is one thing, grounding all of epistemological legitimacy is quite another. Action is rich in cognitive resources (e.g., Peirce, 1899/1992); the automatic perception of events, and the orchestration of action vital for diverse social behaviors (Bandura, 1997; Bargh and Chartland, 1999; Knoblich and Schanz, 2006). Indeed, we have reenvisioned the motor regions of the cortex with regard to the codification of action as we have expanded our notion of cognitive systems; there are many diverse cognitive systems that underlie the organization of perception and action (Rizzolatti and Arbib, 1998; Swanson, 2000). As Susanne Langer remarked, “the brain is an organ, and like all organs, it is built into a greater living whole, an

# PROOF



Action and Cephalic Expression 225

organism that helps to sustain by its special functions" (Langer, 1962, p. 29); the "living whole" (Schulkin, 2009).

Cognitive capabilities and adaptation are distributed across many levels and systems and are not strictly hierarchical (Barton, 2004; 2006); James understood that cephalic function is oriented to "intelligent action" (James, 1890/1952). Habits well orchestrated, cognitively enriched, as most pragmatists would argue, are the articulated features of intelligent action: the healthy habits that serve psychobiological adaptation. Habits are expressed by automatic cephalic expression. James would characterize them in his work as an arsenal of cognitive capability, strongly marked by motor expression, and readily expressible in suitable environments. Mind and brain are joined at the hip, cohabitating a shared terrain of shared practices and experiences (Merleau-Ponty, 1962).

Bodily sensibility underlies bodily intentional behavior (Damasio, 1994; Merleau-Ponty, 1962), tied to motor regions of the brain which we know are both computational and inherently tied to organization of action and to diverse practices that pervade human life and "ready at hand," to use a Heideggerian phrase.

Cognitive systems are pervasive in our understanding of the organization of action (e.g., Gallistel, 1980). Importantly, James and others (e.g., Dewey, 1896; Lashley, 1951; see also Prinz, 2003) understood cognitive systems as endemic to motor control. Cognitive systems underlie, and can be anticipatory of, movement. Reenvisioning the motor systems is to suggest that there is no absolute separation of the brain motor systems from the cognitive systems (see also Berthoz and Luc-Petit, 2006; Gallistel, 1992; Jackson and Decety, 2004)

But they are often richly visceral and anchored to action and to practices in which a background of joined and shared practices is all-pervasive, a diverse array of shared stuff; this takes place in a context in which memory systems are not strictly within the head, in which the predictions of others and statistical inferences underlie human expression.

## Abduction and statistical inferences

C. S. Peirce understood that "thinking is a species of the brain and cerebration is a species of the nervous action" (1880 Logic chapter 1, Indiana [vol. 4]). The logic of inquiry comprises three modes or stages: abduction or the genesis of an idea, induction informed by specific instances, and deduction of logical consequences from general principles. Of these, abduction – sometimes called "retroduction" – is the most



# PROOF

226 Jay Schulkin and Patrick Heelan

interesting from the point of view of hermeneutics, which is depicted in the Figure 9.1.

Abduction or hypothesis in context, constrained by context and ecology, was something Peirce appreciated; problem-solving in context drove his whole thought process. Our sensory capacity is keen a priori to detect objects that afford sustenance or harm (Gibson, 1966; 1979; Gigerenzer, 2000). Fast forms for detecting information (both detecting a “code” and detecting the “decoded meaning”) can change the structure; diverse forms of heuristics, fast ways to solve problems both specific and general, evolved in the evolution of our brain.

Advances in statistical reasoning changed the landscape and offered a chance to embrace probability as inherent in biology and medicine; a sense of objectivity long held within physics now changed when statistical theories questioned the inherent probability of events, moving away from the Newtonian view (Kuhn, 1962; 2000; Galison, 1988). The metaphors in physics changed towards a probabilistic sensibility that was sensitive to the intrinsic diversity – therefore, uncertainty – in decoding the measure number produced by measurement. In the social/historical sciences, such as economics, it was embraced and understood to provide a logical apparatus for distinguishing the singular character of a particular conceptual exemplar, for instance, a green and sour apple, from the generic character of the conceptual exemplars when applied in general inferences and reasoning, as, for example, “An apple a day keeps the doctor away!” In the former case apples are taken as being distributed statistically by color and taste, while in the latter case no such distinction is made. Thus, in the former case, probabilistic relationships are relevant and ground statistical inductions. A logical apparatus in the form of statistical relationships was needed for computational activities, and took material shape.

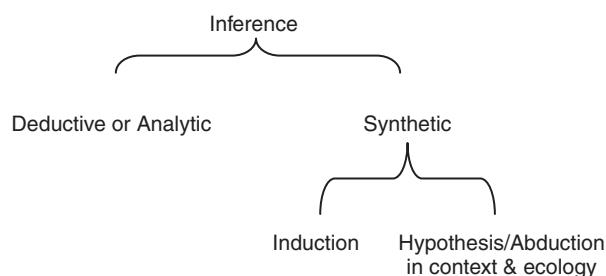


Figure 9.1



The rise of statistical reasoning, however, took on a pejorative sensibility in the hands of those who *a priori* sought deduction models, such as those dominated by Euclidean geometry and mathematics, and by formal logic and Newtonian physics. Nevertheless, in state records, statistics were used in cataloging diseases, viruses, and plague events that impacted human and animal populations, as well as plants and fruits. Although it was seen as a lesser science, a science of chance, nevertheless “taming chance” by formal logic and mathematical models became a wellspring of activity at the heart of critical reason and is a fundamental feature of cognitive architecture at the heart of human experience.

### Demythologized problem-solving

Our cephalic and adaptive responses aim at conserving and storing what we can. To do so, we evolved a set of cephalic adaptations or toolboxes of problem-solving capabilities in terms of naturalistic predilections, embedded in adaptive systems and diverse forms of reasoning and arguments. The term “cephalic structures” includes the “toolbox” of cognitive problem-solving, and, therefore, hermeneutical “applications” of such structures are “cephalic” in addition to the neurological



AQ17 structures.



Central to these concerns is the vital sense of the meaning that is communicated in human expressions. Valuation is a fundamental part of our experience and cephalic expression. One needs a framework in which our appraisals are to be understood: the values that inhere in our judgments about events [Dewey, 1939/1963; Moreno, 1995, 1999/2003; AQ18 AQ19 Neville, 1974; Weissman, 2000]; the social cooperative behaviors as an extension of cephalic capabilities. In other words, valuational judgments are at the heart of our appraisal systems, the cognitive capacity that reflects our interests, and the cultural milieu in which we live (Dewey, 1939/1963; Levinson, 2006). As Jonathan Moreno so nicely put it, “the pragmatic naturalist understands that the knower and that which is known are in the same matrix, just as the inquirer is within nature and is one of its entities along with the object of nature, not outside of nature or fundamentally disconnected from the object” (Moreno, 1999/2003).

Heuristics, orientations that seek less than perfection as a starting point, are close to the pragmatist position; Gerd Gigerenzer and his colleagues (Gigerenzer and Brighton, 2009; Gigerenzer and Goldstein, 1996; [Gigerenzer et al., 1999] following [Herbert Simon, 1969]), look for ways in which we adapt, and satisfy ourselves with the best we can do

AQ20



# PROOF

228 *Jay Schulkin and Patrick Heelan*

with mechanisms that serve us well; we are rooted in a world of social practices and filled with valued objects; hypothesis and experiment begin within this context.

Thus we come prepared to associate a number of events linked by causal building blocks in cephalic structures by worldly events. "Ecological rationality," or readily available heuristics well grounded in successful decision-making (Goldstein and Gigerenzer, 2002; see also Gibson, 1966; 1979; Clark, 1983; 1998; 1999), places decision-making and the use of statistical features within cephalic predilections about numbers and representations of frequencies in real contexts: statistical reasoning grounded in action, in hypothesis formation, in contexts and terrain.

Peirce was prescient and well placed to overview what became dominant themes in the philosophy of science: the status of propositions, reasoning about inferences, a philosophy of induction, a formal representation of inferences to good adaptive explanations given the data, and an orientation to hypothesis generation, something he called, variously, abduction or retrodiction or idea or hypothesis. He provided a logical place for the relationships between induction, deduction, and abduction. This concept can be found in all his works from 1860 to the end of his life in the first part of the twentieth century. In fact, the kernel of his whole thought was abduction, how ideas emerge, their explanatory range, and their long-term consequences.

Pragmatism was regarded as the method to determine consequences writ larger, a move away from nominalism to diverse forms of realism, towards the realism of Scotus and Reid about kinds of objects and our knowledge of them. For Peirce, pragmatism is knotted to "abduction" (Peirce, 1878); abductive events are at the heart of inquiry, abduction is about hypothesis creation. Abduction, in other words, is about hypothesis formation and testing; the grounding is in objects and action, reference and a natural sense of objects.

The richer the hypothesis, the larger the implication: the deeper the hypothesis, the better the understanding. A logic of discovery was the driving force behind Pierce's pragmatism and the emphasis on abduction and hypothesis formation in his notion of inquiry (Hanson, 1958/1972; 1971; Misak, 1991; 2004). Abduction is "preparatory;" induction is a "concluding step" grounded in natural outcomes (Houser et al., 1997; 1998; Peirce, 1878).

The science of discovery is demythologized, the discovery process, pointing towards heuristics, a toolbox of problem-solving devices and an alternative to rationalism and deduction, endless empiricism and

AQ21

# PROOF

Action and Cephalic Expression 229

induction by introducing a term for hypothesis generation: abduction. To use a more modern term, “search engines” underlie inquiry (Dennett, 1987); they can be well grounded or not, but a “lens” serves as an “orientation towards searches using the lens” (Anderson, 1995; Hanson, 1958/1972; Heelan and Schulkin, 1998).

AQ22 Cephalic Heuristics: Numbers and Natural Frequencies: what those who attended would have learned is that evolution did tag, as Peirce suggested, cognitive adaptations towards objects that makes their ready use and understanding available, a veritable set of cephalic heuristics (Gigerenzer, 2000; 2003; 2007); we arrive at some ideas rather quirkily.

AQ23 AQ24 AQ25 Natural frequencies are more readily understood in diverse contexts; for example, for individual researchers with low numeracy capabilities as well as for those who are numbers-oriented. For example, physicians and patients are better able to grasp decisions about risk when the facts are presented as natural frequencies vs. probabilities; the depiction of distribution of events is more transparent, more figurative, less abstract, and perhaps easier to understand. Ecological validity or primacy is linking cephalic capabilities towards enhanced problem-solving (Gigerenzer and Goldstein, 1996, Figure 9.3).

Anchoring our decision-making to useful heuristics with a cephalic orientation about number and frequencies is very useful. The way problems are framed influences the way judgments and decisions are made; by keeping numerical track, the decision-making is demythologized by making it transparent and readily understandable. Keeping track of

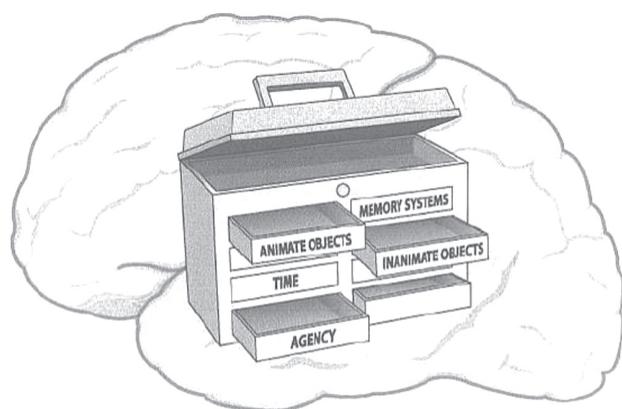


Figure 9.2

# PROOF

230 Jay Schulkin and Patrick Heelan

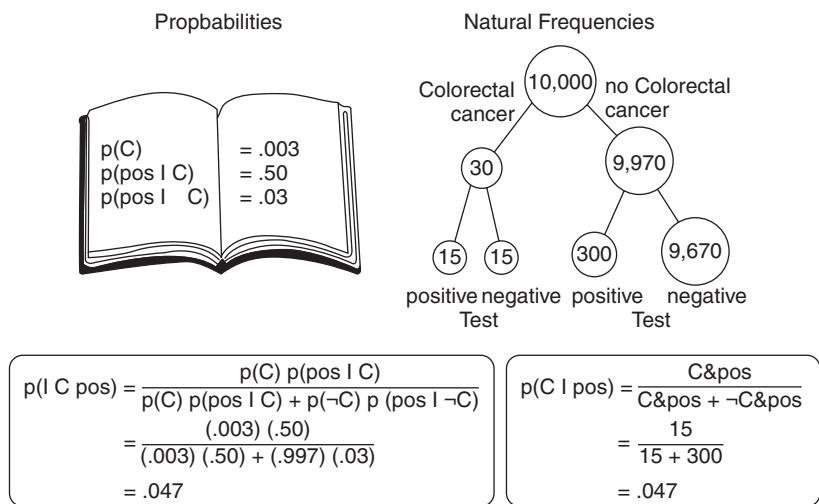


Figure 9.3

events is expanded by the scaffolding of the age of information, our age: the computer revolution, an expansion clearly emboldened by the cataloging and quantification of pockets of information (Hobart and Schiffman, 1998).

Abduction in context, constrained by context and ecology, was something Peirce (1883; 1878; 1889/1992) appreciated; problem-solving in context drove his whole thought process. Our sensory capacity is keen to detect objects that afford sustenance or harm (Gibson, 1966; Gigerenzer, 2000). Thus, since fast forms for detecting information can change the structure and lead to diverse forms of heuristics, fast ways to solve problems, both specific and general, developed in the evolution of our brain and the evolution of our cultural milieu.

We come prepared by causal cephalic functions to associate and link a number of events to other worldly occurrences. "Ecological rationality," or readily available heuristics well grounded in successful decision-making (Goldstein and Gigerenzer, 2002), as the classical pragmatist understood this, places numbers and representations of frequencies in real contexts as a comfortable cephalic choice in decision-making and the use of statistical features.

The sense of numbers pervades everyday forms of reasoning, as the Socratic self-disclosure of the slave suggested (Plato, 1985). It is inherent in the way we structure our world, from space to time to predictions

# PROOF

Action and Cephalic Expression 231

about dangerous or opportunistic events. As Lakoff and Nunez (2000) propose, mathematics involves the following:

1. The embodiment of mind. The detailed nature of our bodies, our brains, and our everyday functioning in the world structures human concepts and human reason, and among these concepts are mathematical concepts and mathematical reason.
2. The cognitive unconscious. Most thinking is unconscious – not repressed in the Freudian sense but simply inaccessible to direct conscious introspection. We cannot intuit – look directly at – how our conceptual systems are embodied in low-level thinking processes; this includes how we think mathematically.
3. Metaphorical thought. For the most part, human beings conceptualize abstract concepts in concrete terms, using ideas and modes of reasoning grounded in the sensory-motor system. The mechanism by which the abstract is comprehended in terms of the concrete is called conceptual metaphor. Mathematical thought also makes use of conceptual metaphor, as when we conceptualize numbers as points on a line.

The cognitive apparatus is oriented towards objects – carving out objects, responding to them, drawing reliable inferences, and assessing their “cognitive significance” (Hempel, 1965), in the context of

AQ29 grounded reason [Kornblith, 1987; 1993; 1994] Godfrey-Smith, 2002; 2008). Reliable inferences are made in a context of what our commitments are [Quine, 1969; 1974; Heelan, 1983/1988; Heelan and Schulkin,

AQ30 AQ31 1998]. Projectable features occur in a framework of understanding, predicting, and cephalic expression; when embedded in adaptation, they cast the problems of induction in a new light – namely, in a context of expectations that matter to us (Gigerenzer, 2000; 2007; Goodman, 1955; 1978) and of practices that are ready at hand for us and pervade the human experience of our knowing.

AQ32 The grounding of induction in human cognitive expectations is put into a context of action (Dewey, 1920/1948; [1925/1989]; 1929/1960;

1938/1973; Peirce, 1878; 1883; 1892); epistemic concerns are grounded in human decision-making and action (Gigerenzer, 2000; Gigerenzer

AQ34 and Stelten, 2001; Hempel, 1965; Levi, 1967; 2004), and in naturalized epistemology (Dewey, 1910; 1916; [1920/1948]; Quine, 1953/1961; 1969;

AQ35 AQ36 1974); knowledge is rooted in our evolving place in nature, in which inferences to a best explanation, or at least a good and satisfactory explanation, are made given time, circumstance, facts, and grounded

232 Jay Schulkin and Patrick Heelan

possibilities. Mythologies of grandiosity are replaced with humility and at times unexpected satisfaction that an idea and the weight of the evidence might be so positively linked, amidst the toil of experimental and theoretical labor, to bring an idea to test and fruition.

What we have discovered is that context and background, along with the scaffolding to practice that expanded memory and capacity

(Donald, 1991; 2004; see also Clark, 1997; 1983; Johnson, 1987/1990; 2007; Noe, 2004; Schulkin, 2009), pervade human competence and performance.

Worlds of cognitive predilection and ecological and social viability converge in everyday expectations and practice (Goldman, 1999; Mead, 1928/1972; [1932/1980]; 1934/1972). Memory reaches the boundaries of the literal brain, both central and peripheral to the social historical milieu; our cognitive capacity is enriched by the contours of our broader historical and ecological context; meaning and memory are strictly beyond the central nervous system to the outer world in which one is adapting, constantly forging new and meaningful connections

(Dewey, 1929/1962; 1929/1966; [1939/1963]; 1938/1973).

Inferences to better explanations must follow something like Peirce's deduction, induction, and hypothesis. Peirce held that "settlement of opinion is the sole object of inquiry" (Peirce, 1877); of course that is misleading, since inquiry is broader, but ideally, as normative goal, finding a level of agreement about the investigation holds for most forms of human activity (e.g., law).

Indeed, our theories are constrained by the things we encounter in the environments with which we are trying to cope, and our ideas as a consequence are shaped by these environments, since adaptation requires us to forge a plan of action and implement strategies based on a coherent representation of these worldly environments.

With respect to perception, Peirce says that seeing is never (simply) seeing; it is always relative to a background, a "seeing as." To this, Dewey and Mead would add the sense of a rich social milieu, constituting what A. Schutz called "the lifeworld" (Schutz and Luckman, 1973).

## Embodied action and objects

The sense of objects is rooted in bodily sensibility (Johnson, 1987/1990; Merleau-Ponty, 1942/1967). Objects are not detached; knowledge is accumulated via transactions with others (Clark, 1999; Dewey, 1925/1989; Schulkin, 2004; Thomas, 2001; Varela et al., 1991). Our cognitive functional adaptations reflect diverse forms of bodily sensibility (e.g., Barsalou, 2003) that are captured across anatomical networks in the

# PROOF

Action and Cephalic Expression 233

brain (Martin, 2007) amidst diverse cultural practices that are endlessly shared [and quite ready] at hand (Heidegger, [1927/1962]).

Sensorimotor features reflect a deep part of our cognitive capabilities (e.g., Dewey, 1925/1989; Lakoff and Johnson, 1999). Cognitive systems are endemic to the organization of perception, and perception is at the heart of action (e.g., Prinz and Barsalou, 2000). But we step into a world of action, of practice, of embodied participation in the doing of things, of participating with others, of shared ways of being in the world (Heidegger, 1962).

The concept of embodiment has several meanings (e.g., Gibbs, 2006; Johnson and Rohrer, 2007; Rohrer, 2001; Wilson, 2002), but the one that I have in mind is the non-Cartesian account: representations of objects, such as whether something is animate or inanimate, are not isolated but endemic to the organization of action and human experience. Representations do not cut us off from the world of objects; indeed, cognitive systems are vehicles for engagement with others (Dewey, 1925/1989; Mead, 1934/1972; Schulkin, 2004). This is particularly apparent with regard to the consideration of agency and human action.

Dewey (1925/1989) and more recently Lakoff and Johnson (1999), among other investigators (e.g. Gibbs, 2006; Wilson, 2002), have suggested that cognitive systems are rooted in action (Barton, 2004; 2006; Jackson and Decety, 2004; Schulkin, 2007). Thinking, that is, has to be understood in the context of action, of transacting with others (Dewey, 1925/1989); cognitive systems are embedded in the organization of action (see also Dewey, 1910/1965; James, 1890; 1917; Johnson, 1987/1990; 2007; Schulkin, 2004).

The emphasis is on embodied cognitive systems (Barsalou, 2003; Wilson, 2002), and the sensorimotor systems are themselves knotted to cephalic machinations across all regions of the brain (see also, e.g., Barton, 2004; 2006; Dewey, 1896; 1938/1973; Schulkin, 1992; 2007).

In other words, cognitive systems are not just a cortical affair; they are endemic to cephalic function.

Moreover, distinctions between concepts in nature and culture blur as inquiry progresses (Clark, 1999; Dewey, 1925/1989). Concepts of nature have their roots in biology, and concepts of culture derive from our cognitive capacity. We are as condemned by, or embedded in, one as we are the other. The study of culture is the study of variation amid some common themes (e.g., Atran et al., 2005, Cassirer, 1944/1978; Langer, 1937; 1962). Mind is rooted in adaptive problem-solving (Clark, 1999; Parrott and Schulkin, 1993), but broadened to envision the cohabiting of our



234 *Jay Schulkin and Patrick Heelan*

living and shared experiences with others (Jaspers, 1913/1997; Merleau-Ponty, 1962); these are not abstract but up close and personal; the cultural context is all-pervasive in the human knowing; it is not surprising [that key] categories are our sense of something as animate (Aristotle, 1962), something alive, something alive with intentions, purposes and expectations; something alive with a culture of shared practices with a history.

AQ49 Such problem-solving, extended into social bonding – ways of conceiving social relationships, recognition of facial and bodily gestures – serves to guide us in the world. And anthropological inquiry reveals the myriad cognitive systems that are operative in the human condition (e.g., Boyer, 1990). Our cognitive abilities are both specific and general (Rozin, 1976; 1998). Cognitive flexibility is a central feature of our problem-solving proclivities (Mithen, 1996; Rozin, 1976).

In other words, core orientations that go into self-corrective inquiry include an orientation toward the observation of objects, the abduction or the genesis of ideas, the tying of ideas about the world to causation, and the development of diverse tools that expand one's observations (e.g., Gigerenzer, 2000; Hanson, 1958/1972; Heelan and Schulkin, 1998; Peirce, 1892; Sellars, 1962). Inferences take place within an orientation to events, and our orientations toward objects are part of the adaptive specializations of animals that create a world for action. Abduction for Peirce was the process of creating a hypothesis from the fallout of disputed hypotheses, within an investigatory form of life rooted in action guided by empirical consequences [see Peirce et al., 1998].

### Hermeneutical pragmatism, meaning, and the lifeworld

Meanings are not fully complete unless incorporated in a linguistic utterance used to affirm or deny some content that finds itself fulfilled in public experience. Perception relates to the perceptual field of the lifeworld. By "perceptual experience" we do not mean simply sensory inputs, but the public recognition of the existence of objects in the space(s) and time(s) of the lifeworld that are understood and categorized through sensory and bodily interactions, and which are the referents of a public ostensive/descriptive language (Heelan, 1983/1988).

Fulfilled and perceptual meanings are not just private mental representations of something, a referent, but are, in fact, by intent identical with the referent that is presented in experience and give access to the ontic and ontological character of that referent under the aspect of what is in truth on this occasion given to understanding. They include,



# PROOF

Action and Cephalic Expression 235

but are not exhausted by, whatever can be reached by a reflective and hermeneutical study of the constitution of fulfilled meanings. Husserl, for instance, typically focused on how "objects" (contents) of knowledge are "constituted" (presented to communal knowers) within "noetic" contexts of meaning (directed by a communal vector of inquiry). Heidegger referred to such objects as "ontic beings" disclosed perspectively to the "circumspective care" of the human inquirer, his term for the human being as coping with the lifeworld. Merleau-Ponty studied them as the "flesh" (or embodiment) of things revealed by perception through the forms of embodied human life in the world (Merleau-Ponty, 1962).

To the extent that language and other public expressive signs are the only means through which we articulate our public world and come to understand one another, the meanings that these signs convey are construals of human cultural communities and cannot be attributed to nonhuman sources except by metaphor. Taking meanings to be cultural does not mean that there is no truth, but that the truth possessed, even scientific truth, is always mediated by human language and culture, which are a part of human history.

Knowledge is handed down by the medium of linguistic and expressive inscriptions and the cultural forms of life in which they find fulfillment. Phrases, however, that once meant one thing come to mean another with the passage of time, for language and culture change. As historians of science know, this is as true for natural science as it is for literature and politics. Of special interest, then, are the circumstances of continuity and change in the historical transmission of scientific meanings via the media of language, mathematics, laboratory praxes, and the culture of the scientific community. Meanings originating at one (linguistic, historical, cultural, geographic) local site are received/interpreted/fulfilled at a distant local site as different meanings.

Like a hammer, or any piece of equipment, a text can be used successfully for several meaningful cultural purposes. As in the case of the hammer, for each useful purpose there are criteria as to how well it performs for this purpose in human life. The uses are not arbitrary, for nothing but nonsense would be gained by arbitrary use, but this does not imply that there is just a single correct meaningful use. Once again, as in the case of the hammer, there may be a priority of users set by a firm cultural tradition – hammers for construction workers, scientific results for scientific research communities – but no one use need go unchallenged either by logic or by experience, nor should any one use become the sole property of just one interested group. Hermeneutic inquiry acknowledges the existence of traditions of interpretation that

# PROOF

236 *Jay Schulkin and Patrick Heelan*

give to today's readers and inquirers a culturally privileged version (shaped to the goals of the linguistic and cultural environment of the community with special "ownership" rights in the subject matter) of past sources. Within the sciences, such traditions of interpretation are at the basis of what Kuhn (1962; 2000) called "paradigms."

In addition to meanings construed on the basis of a common tradition of interpretation (with its presumption of continuity), there can be other legitimate meanings that are independent of any presumption of the existence of a continuity of meaning with the source through a common tradition of life, action, and interpretation. Such discontinuities of meaning within the sciences are exemplified by what Kuhn called "revolutions" in which old "paradigms" are replaced by new ones.

In the work of hermeneutics, however, a radically new meaning need not expel the old, because each, though different, may be a valid historical and cultural perspective. Indeed, despite some sense of discomfort, we often find in the sciences the old flourishing side by side with the radically new, quantum mechanics with Newtonian mechanics (though these are formally incompatible with one another), statistical thermodynamics with phenomenological thermodynamics, and so on. Each, through its own empirical processes of testing and measurement, is in dialogue with confirming or disconfirming data.

In keeping with this transformed background, we can say with a certain presumption of broad agreement that perceptual experience is not to be taken as a curtain that cuts human inquirers off from a real world of generally imperceptible entities, but the milieu within which by interaction we come to understand the world of experience itself and the furniture it contains. There is general agreement that the entities of science are "real" – to be counted among the furniture of the world – but whether they are perceptual entities – "intuitable" by the human imagination – is still disputed. Also, there is widespread agreement that experience is active, embodied, and engaged with public cultural realities, and that it is not passive, or merely the private content of individual minds.

Experience is ever more deeply penetrated by theoretical understanding, from which people learn to adapt with growing success and flexibility to a changing environment. The experiential emphasis of all inquiry rests, then, on current and prospective shared knowledge of diverse perspectives through language, action, perception, and culture. This new background has changed the focus of philosophical interest in scientific inquiry away from the context of justification and prediction to the context of (let us call it) scientific culture. This is the context

AQ51



# PROOF



Action and Cephalic Expression 237

of discovery, interpretation, laboratory practices, historical change, and the influence of science-based technologies on general culture.

Perception relates to the perceptual field of the lifeworld. By "perceptual experience" we do not mean simply sensory inputs, but the public recognition of the existence of objects in the space(s) and time(s) of the lifeworld that are understood and categorized through sensory and bodily interactions, and which are the referents of a public ostensive/descriptive language.

Heidegger referred to such objects as "ontic beings" disclosed perspectively to the "circumspective care" of the human inquirer as *Da-sein*, his term for the human being as coping with the lifeworld and immersed in the "ontological" history of Being. Merleau-Ponty studied them as the "flesh" (or embodiment) of things revealed by perception through the forms of embodied human life in the world (Merleau-Ponty, 1962) with diverse social practices linked to pragmatism and ready at hand (Heidegger, 1962).

To the extent that language and other public expressive signs are the only means through which we articulate our public world and come to understand one another, the meanings that these signs convey are construals of human cultural communities and cannot be attributed to nonhuman sources except by metaphor. Taking meanings to be cultural does not mean that there is no truth, but that the truth possessed, even scientific truth, is always mediated by human language and culture, which are a part of human history.

Knowledge is handed down by the medium of linguistic and expressive inscriptions and the cultural forms of life in which they find fulfillment. Phrases, however, that once meant one thing come to mean another with the passage of time, for language and culture change. As historians of science know, this is as true for natural science as it is for literature and politics. Of special interest, then, are the circumstances of continuity and change in the historical transmission of scientific meanings via the media of language, mathematics, laboratory praxes, and the culture of the scientific community.

Meanings originating at one (linguistic, historical, cultural, geographic) local site are received/interpreted/fulfilled at a distant local site as different meanings. These latter are adopted from traditions of interpretation, or constructed or reconstructed in keeping with the responsibilities, constraints, and presumptions of rational hermeneutical inquiry (the hermeneutical circle), which require that each find fulfillment in local experience at the reception site (Heelan, 1983/1988; Heelan and Schulkin, 1998). One of these constraints is the extent of



# PROOF

238 *Jay Schulkin and Patrick Heelan*

the linguistic and cultural resources available to the distant reader. One of the presumptions is that there is no single meaning that is relevant to and fulfillable by all readers of such a text; there are many, and they depend on the linguistic and cultural resources as well as on the cultural ambience of the reader [Nickles, 1995]. Like a hammer, or any piece of equipment, a text can be used successfully for several meaningful cultural purposes. As in the case of the hammer, for each useful purpose there are criteria as to how well it performs for this purpose in human life. The uses are not arbitrary, for nothing but nonsense would be gained by arbitrary use, but this does not imply that there is just a single correct meaningful use.

The phenomenological principle is that the local lifeworld is the primary horizon of perception within which reality is given to individual humans. This principle takes issue both with rationalist criteria (the privilege of theory over practice) and the empiricist criteria (the privilege of measured data over cultural perception). One of the sources of paradoxes for scientific modernity is the tension in common usage between theory and practice, measurement and perception. This tension is resolved only when it is understood that reality (in a philosophical analysis) is never given absolutely in a unique way, but is always given locally in a local space for a local community for local purposes and, therefore, in spaces with possibly different metrical structure.

Thus, human meaning and understanding are neither in the head nor strictly in the environment; they are in the interaction between the two (Clarke, 1983; 1998; [1999]; Dewey, 1925/1989; Lakoff and Johnson, 1999; Noe, 2004). But they are real. As Peirce [1871] states it, "This realistic theory is thus a highly practical and common-sense position" (p. 488). The worlds we inhabit and adapt to already contain well-worked practices that pervade and scaffold us into a world we embody (Clark, 1996/1997; Heidegger, 1962; Peirce, [1889/1992]).

To relieve the brain from having to store so much information, we have two choices available to us. First is the Gibsonian route: meaning comes in the information that we directly receive from the world in which we live, through common practices (Heidegger, 1962), emboldened by expanding memory systems (Donald, 1992; 2004; Clark, 1997; 1997).

We derive meaning through our senses; we cannot attribute meaning without the sensory input, the external information. But Gibson (1966; 1979) emphasized action patterns that are released by invariant properties of the world in which we evolved. He was short on empirical data, however. Moreover, the mind is not barren; the senses themselves

AQ52

AQ53

AQ54

AQ55

AQ56

AQ57

AQ58

# PROOF



Action and Cephalic Expression 239

AQ59 are within cognitive systems. J. J. Gibson unloaded too much from the mind; but he had an important point: we are delivered ready at hand a diverse array of information coding and existential orientation that reflect ways in which we can engage different horizons of the lifeworld as we come to address them depending on time and context.

We are prepared to learn some kinds of events more easily than others (Garcia et al., 1974; Rozin, 1976). To simplify our cognitive abilities, we learned to tap into the meanings and stability of "natural kinds" AQ60 (see Carey, [1985/1987]; 2009; Kornblith, 1993). Therefore, there need not always be a lot of information decoding and processing. In this sense, the decoding of meaning occurs in the reception itself (e.g., Gibson, 1979; AQ61 Clark, 1996, 1997; Noe, 2004; Wheeler, [2005]). The world – the lifeworld AQ62 as lived – is already coherent; it has inherent regularity or [flusters of] regular patterns that provide the coding coherence which is decoded by the cephalic system and converted into meaning by extended memory systems that traverse the human landscape. The same holds for social kinds. Stable entities are pervasive, and the meaning is out there coded in these stable entities for which there is less need for processing (Clark, 1997; Heelan and Schulkin, 1998).



Inquiry geared toward settled routines is embodied in pervasive social practices. The events are intersubjective, and the social practices are embodied in the communities in which we reside and participate. The knowledge is coded in coherent clusters of meaning; the meaning is pre-packaged for use, linked to a social milieu of the experiences of others.



Indeed, the importance of understanding the experiences of others can be found in the idea that the attribution of intentions to others is not just epistemological; there is something about wanting to understand the experience of others. It is part of how we obtain an orientation to our social circumstances. It is our social evolution, and the utter profundity of the social discourse, that determines the content of our thoughts (Mead, 1934/1972; Sabini and Schulkin, 1994).

## Vision and social contact

We are such visual animals; we look at each other, watch closely what others do. In our species this trend contributed to our ability for joint contact, keeping tag of what we are both watching, of combining actions into coherent social organization (Premack and Premack, 1995; AQ63 Tomasello, 2004).

AQ64 Joint attention or gaze-following is a fundamental adaptation demonstrated in a number of primates (Emery, 2001 (macaques); Perrett et al.,



# PROOF

240 Jay Schulkin and Patrick Heelan

1982; Rolls, 1996) and several studies have shown that diverse regions of the brain in primates, including neocortical and amygdala, are linked to gaze-following. The temporal cortex and temporal pole and amygdala, long linked to meaning [Milner and Goodale, 1995; Perrett et al., 1992; 1982; Rolls, 1998] are importantly involved in social contact and discernment of social milieu.

Key abilities include discerning the wants and the desires of others, along with cognitive adaptations, such as recognizing the kinds of objects that are useful or affordable and those that should be avoided [Gibson, 1966; Rozin, 1976], coupled with a wide array of inhibitory capacities. All those abilities contribute to social cooperative behaviors.

This cognitive competence, so social in its roots, also reflects rapid brain growth during critical periods in our evolution, which perhaps figured in the dramatic expression of our social intelligence (Humphrey, 1976; Levinson, 2003; 2006; Mithen, 1996; 2006). Cultural variation in decision-making, for instance, is anchored to core cephalic predilections. Perhaps this resulted in a brain oriented to change rather than to stability. The world is social for our species; the self is embedded in the larger linguistic/social community.

The visual cortex is, of course, essential for this in us; and projections to a region of the amygdala, linked to the formation of social attachment

*Table 9.1 Components of Human Social Interaction (adapted from Levinson, 2003; 2006)*

Component	Description
Attribution of Intention ("mind-reading")	Crucial precondition, but is by itself nowhere near the abilities needed to generate human interaction.
Mirror Worlds	Mental computations that allow us to simulate the other simulating us.
Gricean Intentions	Intentions that drive behaviors whose sole function is to have an effect by virtue of having their intentions recognized. Allows for high level communication.
Cooperative Human Interactions	There is no point of getting into Schelling Mirror Worlds without presuming cooperation.
Observable Practices	Taking turns, sequencing templates, and repairing templates. These are ethological proclivities to account for multimodal communication.



and social aversion, are key neural connectivities. So our cephalic systems are bulging with visual input, transformed into vital pieces in the organization of action, the organization of meaning (Barton, 2004).

Early evolution showed a trend towards the expansion and design of the visual system (Barton, 2003; Dunbar, 2003; Gross, 1999; Hubel and Wissel, 2005; Hubel et al., 1977), and the evolution of the primate brain visual system (Barten, 2004; Van Essen, 2005; Van Essen et al., 1992). The degree of corticalization of function is related to the anatomy of the visual systems; and what stands out is the visual system in our species (Barton, 2003; Dunbar, 2003; Van Essen, 2005; Van Essen et al., 1992).

An elaborate neural set of structures widely distributed throughout the brain is linked to keeping track of others, watching what they do, and getting a foothold in the world of approachable and avoidable events (Frith and Wolpert, 2003; Jacob and Jeannerod, 2005; Schulkin, 2000). And, when the volume of visual primary cortex is measured in terms of the whole of the neocortex with regard to social behavior, they are clearly linked (Dunbar and Shultz, 2007). A wide variety of evidence links the degree of social interaction with neocortical expansion (Dunbar, 2003; Falk, 1983). Diverse models of group size have been linked to neocortical enlargement (Dunbar and Shultz, 2007; Zhou et al., 2005; see also Deaner et al., 2007) and cognitive competence (Byne and Corp, 2004) across distributed cognitive systems (Barrett et al., 2007; Barton, 2004; Schulkin, 2007) embedded in rich social practices that deliver a shared sensibility with others; visual architecture is just one part of a rich visual system that underlies visual experience. The emphasis, for us, is on the shared experiences and practices; they are all-pervasive. Our rooted orientation is embodied; no Cartesian separation from a buzzing world in action and perception and oriented to action.

### Brain, bodily motion, and faces

Of course, a reasonable strategy is to discern in mammalian nonhuman primates which regions of the brain are fundamental for knowledge – social and otherwise (e.g., Baron-Cohen et al., 1999; Brother, 1994; Hauser, 1999). More specifically, one reasonable and tractable question is how the brain responds to different objects (e.g., Desimone et al., 1984; Rolls et al., 1997). For example, we now know some of the regions of the brain (i.e., temporal cortex) and how these brain regions respond to faces (e.g., Desimone et al., 1984; Gross, 1992; Perrett et al., 1998) or to different kinds of object motion (Martin, 2007; Oram and Perrett, 1996).





242 Jay Schukin and Patrick Heelan

Different visual processing pathways are thought to mediate the identification of the location of an object, the form of an object, and the direction or the motion and trajectory of an object (Ungerleider and Haxby, 1994; Ungerleider and Mishkin, 1982; see also Miller and Goodale, 1995). A ventral visual pathway to the temporal lobe is known to represent a number of objects (e.g., faces, houses, chairs, etc. (Ishai et al., 1999) in the coding of semantics in the brain (Rolls, 2000)).

Thus, regions of the temporal cortex compute and decode faces, bodily posture, and motion (e.g., Perrett and Mistlin, 1990). Facial and bodily responses are great sources of coded information (e.g., Darwin, 1872/1965; Ekman, 1973; Young, 1998). Neurons within the temporal lobe are responsive to the individual attention of others, the direction of their eyes, and the direction of their motion (e.g., Perrett et al., 1992; Rolls and Treves, 1998), which can be immediately decoded.

For example, in one study, neurons within the superior temporal sulcus were recorded in macaques. They were responsive to both human and monkey faces (Rolls et al., 1997) to bodily posture, and to motion direction (Perrett et al., 1992). The firing pattern was more significantly elevated when the gaze of another was direct than when it was not. The direct gaze is typically a threatening gesture in nonhuman primates (e.g., Hinde, 1970), and therefore this region of the brain is probably linked to the fear information decoding and processing system in diverse regions of the brain (Rolls and Deco, 2002).

### Imagining and doing: cephalic capabilities

Both the cognitive revolution and the emphasis on biological adaptation have fostered the view that there is a diverse array of cognitive mechanisms in learning (Rozin, 1976) and in memory (Squire and Zola, 1996; Tulving, 1983/1993; 2002). Semantic processing reflects different cognitive decoding systems as well as neural mechanisms (e.g., Squire and Zola, 1996). More than one system appears to underlie semantics in the nervous system (e.g. Martin, 2011).

Moreover, from an anatomical point of view, representations of known objects, like most functions in the brain, are not simply localized in one part of the brain, but are distributed across the neural axis (e.g. (Rolls, 1998)). But the representations are not random; they reflect the underlying organization of the nervous system as well as its evolution and function. One important insight into the brain is that semantic processing reflects regions of the brain linked to action and

AQ79

AQ80



AQ81



# PROOF

Action and Cephalic Expression 243

perception which involve decoding of the representations (Jeannerod, 1999; Martin, 1998).  
AQ82

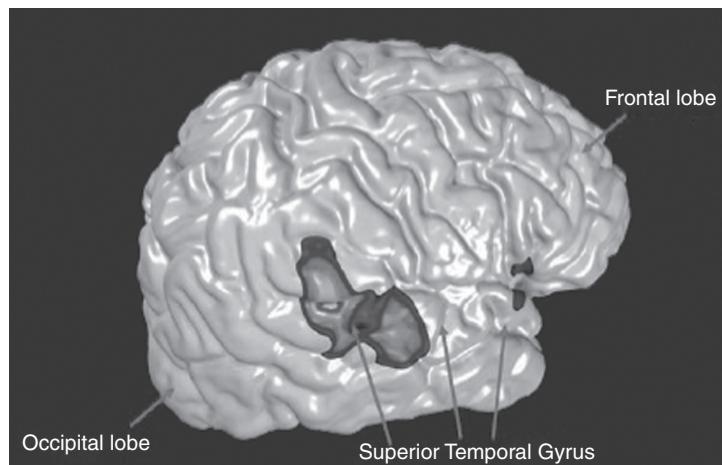
Consider briefly some illustrative basic facts about neural structure and function with regard to object and social perception. Experiments by Roger Shepard and his colleagues (Shepard and Cooper, 1982; Shepard and Metzler, 1971) demonstrated that, whether a visual rotation is imagined or an object is actually looked at, the time period needed to do so reflects the size of the object.

Moreover, we now know that similar neural circuits are also activated when the object is imagined or viewed (Farah, 1984; Kosslyn, 1994). Imagining is the process of creating brain stimulation internally similar to what would be created from external stimulation. In other words, the neural structures active in imagining objects appear similar to those structures active when looking at them (Kosslyn, 1980; 1994). For example, using positron emission topography (PET) to measure regional blood flow, it has been observed that the brain regions that are active when one imagines the object are the same as those active when looking at it (Kosslyn, 1994). These regions include the left angular gyrus, dorsal lateral prefrontal cortex, right superior parietal lobe, area 17 of the left hemisphere, fusiform gyrus in both hemispheres, and the right inferior parietal lobe (Kosslyn et al., 1993). These regions of the brain play differential roles in determining visual objects and their meaning, and also in imagining those objects.  
AQ83

Consider another example, in the auditory system. In a functional magnetic resonance imaging (fMRI) study, subjects were presented with auditory words through headphones; in another experiment, the same individuals were asked to identify the words with silent lip-reading (Calvert et al., 1997). The temporal cortex was activated during both the heard speech and the lip-reading. Other regions of the brain linked to visual motion were activated – the extrastriate cortex and inferoposterior temporal lobe (Calvert et al., 1997).

It is found that a piece of music activates many of the regions linked to auditory perception and to music (Zattore et al., 2002). Indeed, regions of the auditory cortex are clearly activated when subjects are asked to imagine music or other auditory stimuli (Zatorre and Halpern, 2005; Zattore et al., 2002).  
AQ84  
AQ85

Perhaps one is in a better position to understand the genius of Beethoven; deaf for many years, he must have heard music to compose the way he did; think of the cognitive complexity, the richness of the later parts of Beethoven's life. Also, it makes it somewhat easier to understand that music to one's ears may not be heard by one's neighbor  
AQ86



AQ87 *Figure 9.4* Parts of the auditory cortex that respond to imagining or hearing music (Zatorre et al., 2002).

in the same way. Of course, Beethoven is one thing, the rest of us quite another; but the recruitment of brain regions is generic.

In studies using PET to measure blood flow or neural activation, subjects were asked to imagine grasping objects (e.g., Decety et al., 1994; Stephan et al., 1995). Significant activation of regions of the brain concerned with movement was apparent. For example, Brodmann's Area 6 in the inferior part of the frontal gyrus of both cortical hemispheres was active when subjects were asked to imagine grasping an object. The anterior cingulate and the parietal cortex were also activated. In addition, both the caudate nucleus of the basal ganglia and the cerebellum were activated. In further studies using neuromagnetic methods to measure cortical activity, the primary motor cortex was active both when subjects observed simple movements and when the subjects performed them (Hari et al., 1998). Motor imagery is replete with cognitive structure and is reflected in the activation of neural circuitry (see also Rizzolatti and Arbib, 1998; Decety et al., 1994).

There is an emphasis on the body's appropriation of objects and their use (Dewey, 1925/1989; Lakoff and Johnson, 1999; Merleau-Ponty, 1942/1967). The body is a vehicle for meaning and knowledge, replete with meaningful cognitive structure for knowing what is around and what to attend to, learn from, and respond to. This sense of body knowledge is well represented in the brain and, when decoded, contributes



to the organization of intelligent action (Damasio, 1994). The sensorimotor organization is also replete with representations that are cognitive (Jeannerod, 1997; 1999; Rizzolatti and Arbib, 1998). These are not meaningless bodily events. The class of representations is larger than simply propositional ones; cognitive structure is pervasive, of course, amidst common human practices that we share with others (Jackson and Decety, 2004).

The anatomical/functional relations have relevance to considerations about regions of the brain and social knowledge. It may be the case that my acting in an intentional way, or imagining that I am, and my attributing intentional action to you activate the same brain regions, and that these brain regions, when decoded, may contribute to the organization of actions (e.g., Gallese, 2007; Gallese and Goldman, 1998).  
AQ88

## Conclusion

The human condition is determined by our evolutionary history and by pervading historical and social factors. We are thrust out in the world. The process of evolution engendered in humans a desire to compete and to cooperate, to form bonds of intimacy, and to deceive.

To placate the uncertainty and insecurity of existence, science has in some ways replaced the traditional theologies that attempted to satisfy our quest for certainty (Dewey, 1929/1960). Human desire for certain knowledge and boundless security is infinite in scope. Our aspirations rise higher than any barometer we can construct. Whatever science is, it is not about certainty, but meaning and knowledge. And it does represent one of the glorious treasures of humankind.  
AQ89

A sense of history matters. Science, like everything else human, is historically contingent (Dear, 1995; 2006; Hull, 1988; Shapin, 1995; 1996; Todes, 1989; 1997) and always needs to be linked to what matters to the human condition (Polanyi, 1946/1964). Perhaps a phrase from Kuhn is applicable: "post Darwinian/Kantianism" (Kuhn, 2000, p. 164), enriched by lexical entries in the common vocabulary of understanding (Levinson and Jaisson, 2006) and linked to diverse social practices that are quite easily available.  
AQ90

It is the meaning of the lifeworld with others that, in our transactions, the theory and practice of the sciences are to be uncovered in everyday life (e.g., Dewey, 1925/1989; Heelan and Schulkin, 1998). These events are meaningful both existentially and historically, and should be explored by self-corrective methods of inquiry and hypothesis testing as broadly conceived across the spectrum of human experience,  
AQ91



# PROOF

246 *Jay Schulkin and Patrick Heelan*

from language to art and from mathematics to religion. Central metaphors define our approaches to inquiry (e.g., Galison, 1988); our cognitive toolbox provides the orientation and the set of skills [Carey, 2004; Lakoff and Johnson, 1999].

The emergence of rich linked semantic networks and external source codes that expand memory for action and perception contributed to the evolving cephalic capabilities of humans [Clarke, 2008; Donald, 2004; Wheeler, 2005]. These diverse cognitive abilities constitute what it takes to "read"—that is, to decode—symbolic expressions (e.g., Atran et al., 2005; Deacon, 1997; Levinson, 2003; 2006; Sperber, 1985). Symbolic expression permeates our human understanding (Cassirer, 1944/1978). The ability to use symbols in this way is an adaptation, for, in manipulating symbols, humans learned to give meaning to—to "read"—symbolic objects in the environment; from this came languages, politics, technologies, and the making of historical social cultures.

Symbolic representation probably emerged from a rudimentary stage in which early humans represented the objects they used and encountered to comprehend them. The use of symbols is "the distinctive mark of human life" [Cassirer, 1946; Langer, 1937; Sperber, 1975]. Symbolic expressions line the vast array of human social contact and meaning. These are adaptive responses to attempt to tame the perplexities of uncertainty and insecurity. Human understanding lies in the ability to secure stability and security in representing events amidst social solidarity (Dewey, 1925/1989; Gigerenzer, 2000; 2007; Moreno, 1995; 1999/2003).

Natural science uses hermeneutic methods in its growth and development, in its use of models, and in establishing continuity with its archaic traditions. In the existential analysis of acts of scientific knowledge about the world, however, the central role of hermeneutics turns out to be exercised where it is to be least expected according to the standard accounts of natural science, namely, in the process of scientific observation.

The Book of Nature which science "reads" was not, as Galileo thought, "written" in final form from time immemorial, but its "text" is generated in response to a human form of inquiry, the purpose of which is to reveal universal antecedent conditions of possibility of all cultural activity. What is meant by the question, however, and what counts as a satisfactory answer depends on the time and place, and the cultural interests of the scientific community.

Embodied cognitive systems are endemic to central nervous system function; the brain is a decoding information-processing system, a

# PROOF



Action and Cephalic Expression 247

diverse collection of systems that reflect the meaning-making and cognitive machinations of the brain (Schulkin, 2004; 2009). The idea of embodied cognitive systems requires two concepts – a sense of animacy and a sense of agency – key cognitive categories that underlie our sense of one another and of ourselves, and that figure in our understanding of human history.

The classificatory distinction of the living from the nonliving is a fundamental cognitive adaptation much expanded and developed. Categorical attributions of animacy and agency are dominant early, but not totally unconstrained (Gelman, 2003), and they are intertwined. Both categories matter in determining the world around us (e.g., Atran, 1990/1996; Boyer, 2002) amidst a toolbox of cephalic capabilities and an all-pervasive social world of ready-to-hand social practices that bind us together.

AQ96

Diverse senses of embodied cognitive systems are inherent in the organization of action (Barsalou, 2003; Wilson, 2002). Our cognitive evolution, with its extra premium on memory and language, implodes our cognitive capabilities; our cultural evolution draws on all resources, as our memory and communicative systems are both internal and external to us (Clark, 1997; Donald, 1991). With the onset of symbolic extended and external memory in human evolution, our universe expanded exponentially.

AQ97

Although many regions of the brain underlie the animate–inanimate distinction, a word's depiction of an action (e.g., viewing a hammer) is more likely to activate motor and premotor areas of the cortex (see, e.g., Caramazza and Mahon, 2005; Martin and Caramazza, 2003). This activation is a representation of – that is, symbolic of – the meaning attached to it; the imagery of action, [watching others], getting anchored by a symbolic representation of a world in which the representation is decoded in perception and action (Jackson and Decety, 2004). Importantly, looking at action words (Hauk et al., 2004; Pulvermuller et al., 2005) and the performance of an action activate many of the same symbolic, or representational, regions of the brain (e.g., Martin et al., [2000]; Perani et al., 1995). The coherence in the organization of the brain and its linked “decoders” is the tight link between cognitive systems and action and function.

AQ98

AQ99

Imitating others is represented symbolically in many different regions of the brain that reflect coded representations of action, of motion, of agency (Chaminade et al., 2005). Coded representations of action words and regions of the brain that underlie action, and category-specific prepared knowledge about objects linked to action, show that cognitive

AQ100



# PROOF

248 Jay Schulkin and Patrick Heelan

systems traverse most, if not all, regions of the brain (Jackson and Decety, 2004; Schulkin, 2004). Importantly, these areas of the brain underlie – code for – the adaptation of shared experience, a sense of community (Dewey, 1934; Flanagan, 2007) and meaning (Jaspers, 1913/1997).

Indeed, we know that our success is linked to social cooperative behaviors (Levinson, 2003; 2006). One pragmatist who continually wrote about cooperative behaviors, an essential feature of our evolution and our intelligence, was John Dewey. He believed that social behavior is rooted in our psychobiology, anchored to an individual that is social in nature, for which pervasive moral experience and an ethics that sets the condition for a participatory democracy emerges, and our capacity to work through endless conflicts.

Dewey anchored a rich sense of human experience to the social context of civilized action, and [also to cephalic] propensity embedded in a sense of objects; we are rooted in objects, as he often put it (Dewey, 1925/1989), forged in communicative social contexts of adaptation while coping with diverse forms of precarious experience. One key is to anchor the prosocial sensibility into adaptive, culturally and socially bonded individuals, diluting differences and encouraging engagement, and, as Dewey often noted, “all human experience” has a social component” (Dewey, 1938/1973).

Embodied cognitive systems are decoded and expressed in our human experience (e.g., Clark, 1999; Gallagher, 2005; Lakoff and Johnson, 1999; Stewart et al., 2011; Varela et al., 1991; Wheeler, 2005), and they underlie as codes for a multitude of forms of practice across human symbolic expression (Bourdieu, 1980/1990). What underlies those, in turn, are rich decoding information-processing systems in the brain, now being studied in some depth (Gazzaniga, 1995/2000) and always embodied in human knowing (Changeux, 2002) amidst shared human experiences (Dewey, 1896; 1925/1989; 1938/1973; Heidegger, 1927/1962).

Still, with an eye to the past and the organization of the mind/brain (Gazzaniga, 1998) and our cultural expressions, advances in inquiry need to be linked to our existential fate (Moreno, 1999/2003); we are great decoding information-processing systems who have now uncovered great depths of scientific understanding. But we are obviously more than decoding information-processing systems; we are more than the methods of our age. A variety of forms of human reasoning mark us, but two stand out: our orientation toward method and our orientation toward synthesis of meanings (or ideas), what Whitehead (1929/1958) called the way of Ulysses versus Plato. Both are essential; one has been traditionally linked to the sciences and the other

to the humanities. In fact, this is misleading, for both the sciences and the humanities embody both of them. The distinction is real, but cuts across disciplines, just as perception, readiness, and traditions of practice are inherent in all disciplines of inquiry. A self-corrective conception of investigation and invention cuts across the broad range of human inquiry but needs to be anchored in historical and existential recognition of the frail sense of our human achievements with regard to the treatment of one another.

## References

- Aristotle (1962). *The Nicomachean Ethics*. New York: Macmillan.
- Arnheim, R. (1974). *Art and Visual Perception* (Berkeley and Los Angeles: University of California Press).
- Atran, A., Medin, D.L., and Ross, N.O. (2005). The cultural mind. *Psychological Review*, 112, 744–66.
- Austin, J. L. (1970). *Philosophical Papers*. Oxford: Oxford University Press.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. San Francisco: Freeman.
- Baron-Cohen, S., Ring, H. A., Wheelwright, S., Bullmore, E. T., Brammer, M. J., Simmons, A., and Williams, S. C. R. (1999). Social intelligence in the normal and autistic brain: An fMRI study. *European Journal of Neuroscience*, 11, 1–8.
- Barsalou, L.W. (2003). Abstraction in perceptual symbol systems. *Phil. Trans. Soc. London*, 358, 1,177–87.
- Barton, R. A. (2004). Binocularity and brain evolution in primates. *PNAS*, 101, 10,113–5.
- Barton, R.A. (2006). Primate brain evolution: integrating comparative neurophysiological and ethological data. *Evolutionary Anthropology*, 15, 224–36.
- Berridge, K. C. (2004). Motivation concepts in behavioral neuroscience. *Physiology and Behavior*, 81, 179–209.
- Berthoz, A. (2002). *The Brain's Sense of Movement*. Cambridge, MA: Harvard University Press.
- Berthoz, A., and Petite, J. L. (2006). *The Physiology and Phenomenology of Action*. Oxford: Oxford University Press.
- Bourdieu, P. (1980/1990). *The Logic of Practice*. Palo Alto: Stanford University Press.
- Boyer, P. (1990). *Tradition as Truth and Communication*. Cambridge: Cambridge University Press.
- Brook, A. and Akins, K. (2005) *Cognition and the Brain: The Philosophy and Neuroscience Movement* (Cambridge: Cambridge University Press).
- Brothers, L. (1997). *Friday's Footprint*. Oxford: Oxford University Press.
- Byrne, R. W., and Corp, N. (2004). Neocortex size predicts deception rate in primates. *Proc. R.Soc.* 271, 1,693–9.
- Calvert, G. A., Bullmore, E. T., Brammer, M. J., Campbell, R., Williams, S. C. R., McQuire, P. K., Woodruff, P. W. R., Iverson, S.D., and David, A. S. (1997). Activation of auditory cortex during silent lipreading. *Science*, 276, 593–6.
- Campbell, J. (1995). *Understanding John Dewey: Nature and Cooperative Intelligence*. Chicago: Open Court.

# PROOF

250 Jay Schulkin and Patrick Heelan

- AQ106 Caramazza, and Mahon, B.Z. (2005). The organization of conceptual knowledge in the brain: the future's past and some future directions. *Cognitive Neuropsychology*, 22, 1–25.
- Carey, S. (1985/1987). *Conceptual Change in Childhood*. Cambridge: MIT Press.
- Carey, S. (2009). *The Origin of Concepts*. Cambridge: Harvard University Press.
- Cassirer, E. (1944/1978). *An Essay on Man*. New Haven: Yale University Press.
- Changeux, J. P. (2002). *The Physiology of Truth*. Cambridge: Harvard University Press.
- Cheney, D.L., and Seyfarth, R.M. (1990). *How Monkeys See the World*. Chicago: University of Chicago Press.
- Cheney, D. L., and Seyfarth, R.M. (2007). *Baboon Metaphysics*. University of Chicago Press.
- Clark, A. (1998). *Being There: Bringing Brain, Body, and World Together*. Cambridge, MS: Bradford Books/MIT Press.
- Clark, A. (1999). An embodied cognitive science? *Trends in Cognitive Sciences*, 3, 345–51.
- Clark, R.W. (1983). *Benjamin Franklin*. New York: DaCapo Press.
- Cohen, M. S., Kooslyn, S. M., Breitner, H. C., DiGirolamo, G. J., Thompson, W. L., Anderson, A. K., Bookheimer, S. Y., Rosen, B. R., and Belliveau, J. W. (1996). Changes in cortical activity during mental rotation: A mapping study using functional MRI. *Brain*, 119, 89–100.
- Damasio, A. R. (1994). *Descartes' Error: Emotion, Reason, and the Human Brain* (New York: Grosset/Putnam).
- Darwin, C. (1872/1965). *The Expression of Emotions in Man and Animals*. Chicago: University of Chicago Press.
- Deacon, T.W. (1997). *The Symbolic Species*. New York: Norton.
- AQ108 Deaner, R. O., Isler, K., Burkart, J., Burkart, J., and van Schaik, C. (2007). Overall brain size, and not encephalization quotient, best predicts cognitive ability across non-human primates. *Brain Behavior and Evolution*, 70, 115–24.
- Dear, P. (1995). *Discipline and Experience*. Chicago: University of Chicago Press.
- Dear, P. (2006). *The Intelligibility of Nature*. Chicago: University of Chicago Press.
- Decety, J., Perani, D., and Jeannerod, M. (1994). Mapping motor representations with positron emission tomography. *Nature*, 371, 600–2.
- Dennett, D. (1987). *The Intentional Stance*. Cambridge: MIT Press.
- Desimone, R. (1991). Face-selective cells in the temporal cortex of monkeys. *Journal of Cognitive Neuroscience*, 3, 1–8.
- Desimone, R., Albright, T.D., Gross, C.G., and Bruce, C. (1984). Stimulus-selective properties of inferior temporal neurons in the macaque. *Journal of Neuroscience*, 8, 2,051–62.
- deWaal, F. B. M. (2000). Primates—A natural heritage of conflict resolution. *Science*, 289, 586–90.
- Dewey, J. (1896). The reflex arc concept in psychology. *Psychological Review*, 3, 357–70.
- Dewey, J. (1925/1989). *Experience and Nature*. LaSalle, IL: Open Court Press.
- Dewey, J. (1938). *Logic: The Theory of Inquiry*. New York: Holt, Rinehart.
- Donald, M. (1991). *Origins of Modern Man*. Cambridge: Harvard University Press.
- Donald, M. (2001). *A Mind So Rare: The Evolution of Human Consciousness*. New York: Norton.

# PROOF

Action and Cephalic Expression 251

- AQ109 Donald, M. (2004). Hominid enculturation and cognitive evolution. In J.M. Luraciello et al. (Eds), *The Development of the Mediated Mind*. Mawash, NJ: Erlbaum Press.
- Dunbar, R. I. M. (2003). The social brain. *Annual Review of Anthropology*, 32, 163–81.
- Dunbar, R. I. M., and Shultz, S. (2007). Evolution in the social brain. *Science*, 317, 1,344–7.
- Ekman, P. (1973). Cross-cultural studies of facial expression. In P. Ekman (Ed.), *Darwin and Facial Expression*. New York: Academic.
- AQ110 Falk, D. (1983). Cerebral cortices of East African early hominids. *Science*, 221, 1,072–4.
- Farah, M. J. (1984). The neurobiological basis of visual imagery: A componential analysis. *Cognition*, 18, 245–72.
- Fodor, J. (1983). *The Modularity of Mind*. Cambridge: MIT Press.
- Frith, C., and Wolpert, D. (2003). *The Neuroscience of Social Interaction*. Oxford: Oxford University Press.
- Galison, P. (1988). History, philosophy and the central metaphor. *Science in Context*, 3, 197–212.
- Gallagher, S. (2005). *How the Body Shapes the Mind*. Oxford: Oxford University Press.
- Gallese, V. (2007). Before and below “theory of mind”: Embodied simulation and the neural correlates of social cognition. *Phil. Trans. R. Soc. B*, 362, 659–99.
- Gallistel, C. R. (1980). *The Organization of Action: A New Synthesis*. Hillsdale, NJ: Lawrence Erlbaum.
- Gallistel, C. R. (1992). *The Organization of Learning*. Cambridge, MA: MIT Press.
- Garcia, J., Hankins, W. G., and Rusiniak, K. W. (1974). Behavioral regulation of the milieu interne in man and rat. *Science*, 185, 824–31.
- Gazzaniga, M. S. (1995/2000). *The New Cognitive Neurosciences*. Cambridge, MA: MIT Press.
- Gazzaniga, M. S. (1998). *The Mind’s Past*. Berkeley: University of California Press.
- Gelman, S. A. (2003). *The Essential Child*. Oxford: Oxford University Press.
- Gibbs, R. W. (2006). *Embodiment and Cognitive Science*. Cambridge: Cambridge University Press.
- Gibson, J. J. (1966). *The senses considered as perceptual systems*. New York: Houghton-Mifflin.
- Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. New York: Houghton-Mifflin.
- Gigerenzer, G. (2000). *Adaptive thinking, rationality in the real world*. New York: Oxford University Press.
- Gigerenzer, G. (2007). *Gut Feelings*. New York: Viking Press.
- Gigerenzer, G., and Brighton, H. (2009). *Homo heuristicus*: Why biased minds make better inferences. *Topics in Cognitive Science*, 1, 107–43.
- Gigerenzer, G., and Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, 103, 650–69.
- Gigerenzer, G., and Selten, R. (2001). *Bounded Rationality*. Cambridge: MIT Press.
- Gigerenzer, G., Todd, P. M., and the ABC Research Group (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.
- Godfrey-Smith, P. (2002). Dewey on naturalism, realism and science. *Philosophy of Science*, 69, S1–S11.

# PROOF

252 Jay Schulkin and Patrick Heelan

AQ111 Godfrey-Smith, P. (2008). Induction, Samples, and Kinds, in *Carving Nature at its Joints: Topics in Contemporary Philosophy, Volume 8*. MIT Press.

Goffman, E. (1959). *The Presentation of Self in Everyday Life*. New York: Doubleday Anchor.

Goffman, E. (1971). *Relations in Public*. New York: Harper and Row.

Goldman, A. I. (1999). *Knowledge in a Social World*. Oxford: Clarendon Press.

Goldstein, D. G., and Gigerenzer, G. (2002). Models of Ecological Rationality: The Recognition Heuristic, *Psychological Review*, 109, 75–90.

AQ112 Grice, P. (1957). Meaning. *Philosophical Review*.

Gross, C. G. (1999). *Brain, Vision and Memory*. Cambridge: MIT Press.

Hanson, N. R. (1971). *Observation and Explanation*. New York: Harper Press.

Hanson, N.R. (1958/1972). *Patterns of Discovery*. Cambridge: Cambridge University Press.

Hari, R., Forss, N., Avikainen, S., Kirverskari, E., Salenius, S., and Rizzolatti, G. (1998). Activation of human primary cortex during action observation: A neuromagnetic study. *Proceedings of the National Academy of Sciences*, 95, 15,061–5.

Hauk, O., Johnsrude, I., and Pulvermuller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron*, 41, 301–7.

Heelan, P. (1983/1988). *Space-Perception and the Philosophy of Science*. Berkeley: University of California Press.

Heelan, P. A., and Schulkin, J. (1998). Hermeneutical philosophy and pragmatism: A philosophy of the science. *Synthese*, 115, 269–302.

Heidegger, M. (1927/1962). *Being and Time*. (J. Macquarrie and E. Robinson, Trans.) London: SCM Press.

Helmholtz, H. (1873). *Popular Lectures in Scientific Subjects*. (E. Atkinson, Trans.; Introduction by J. Tyndale). London: Longmans Green.

Hempel, C. G. (1965). *Aspects of Scientific Explanation*. New York: Free Press.

Hinde, R. A. (1970). *Animal behavior (2nded)*. New York: McGraw-Hill.

Hobart, M. E., and Schiffman, Z.S. (1998). *Information Ages*. Baltimore: Johns Hopkins University Press.

AQ113 Houser, N., Roberts, D. D., and Van Evra (Eds) (1997). *Studies in the Logic of Charles Sanders Peirce*. Indianapolis: University of Indiana Press, p. 107.

Houser, N., Eller, J. R., Lewis, A. C., De Tienne, A., Clark, C. L., and Bront Davis, D. (1998). *The Essential Peirce Volume 2 (1893-1913)*. Bloomington: Indiana University Press.

Hubel, D. H., Wiesel, T. N., and LeVay, S. (1977). Plasticity of ocular dominance columns to monkey striate cortex. *Phil. Trans. Lond. Soc. B.*, 278, 377–409.

Hubel, D. H., and Wiesel, T. N. (2005). *Brain and visual perception: The story of a 25 year collaboration*. Oxford: Oxford University Press.

Hull, D.L. (1988). *Science as a Process*. Chicago: University of Chicago Press.

Humphrey, N. (1976). The social function of intellect. In P.P.G. Bateson, and R.A. Hinde (Eds), *Growing Points in Ethology*. Cambridge: Cambridge University Press, pp.307–17.

Husserl, E. (1969). *Formal and Transcendental Logic*. (J. Cairns, Trans.) The Hague: Nijhoff.

Husserl, E. (1973). *Experience and Judgment: Investigations in a Genealogy of Logic*. (J. Churchill and K. Ameriks, Trans.). Evanston, IL: Northwestern University Press.

# PROOF

Action and Cephalic Expression 253

- Ishai, A., Ungerleider, L. G., Martin, A., Schouten, J. L., and Haxby, J. V. (1999). Distributed representations of objects in the human visual pathway. *Proceedings of the National Academy of Sciences*, 96, 9,379–84.
- Jackson, P. L., and Decety, J. (2004). Motor cognition: A new paradigm to self and other interactions. *Current Opinion in Neurobiology*, 14, 259–63.
- Jacob, P., and Jeannerod, M. (2004). *Ways of Seeing: The Scope and Limits of Visual Cognition*. Oxford: Oxford University Press.
- Jacob, P., and Jeannerod, M. (2005). The motor theory of social cognition: A critique. *Trends in Cognitive Science*, 9, 21–5.
- James, W. (1890/1952). *Principles of Psychology*, Vol. 1. New York: Henry Holt, p. 79.
- Jaspers, K. (1913/1997). *General Psychopathology*, Vol. I and II. (J. Hoenig and M. W. Hamilton, Trans.) Baltimore: The Johns Hopkins University Press.
- Jeannerod, M. (1997). *The Cognitive Neuroscience of Action*. Oxford: Blackwell.
- Jeannerod, M. (1999). To act or not to act: perspectives on the representation of action. *Quarterly Journal of Experimental Psychology*, 52, 1–29.
- Johnson, M. (1987/1990). *The Body in the Mind*. Chicago: University of Chicago Press.
- Johnson, M. (2007). *The Meaning of the Body*. Chicago: University of Chicago Press.
- Johnson, M., and Rohrer, T. (2007). We are live creatures: embodiment, American pragmatism and the cognitive organism. In *Body, Language and Mind* Vol. 1, Berlin: Mouton de Gruyter, in press.
- Kagan, J. (1984). *The nature of the child*. New York: Basic Books.
- AQ114  
Kagan, J. (2002). *Surprise, uncertainty and mental structure*. Cambridge, MA: Harvard University Press.
- Knoblich, G., and Sebanz, N. (2006). The social nature of perception. *Current Directions in Psychological Science*, 15, 99–104.
- Kornblith, H. (1993). *Inductive Inference and its Natural Ground*. Cambridge: MIT Press.
- Kosslyn, S. M. (1980). *Image and Mind*. Cambridge, MA: Harvard University Press.
- Kosslyn, S. (1994). *Image and Brain*. Cambridge: MIT Press.
- Kosslyn, S. M., Alpert, N. M., and Thompson, W. L. (1993). Visual mental imagery and visual perception: PET studies. In Society for Magnetic Resonance Imaging. *Functional MRI of the Brain*. Virginia: Arlington.
- Kuhn, T. S. (1962). *The Structure of Scientific Revolution*. Chicago: University of Chicago Press.
- Kuhn, T. S. (2000). *The Road Since Structure*. Chicago: University of Chicago Press.
- Lakoff, G., and Johnson, M. (1999). *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*. New York: Basic Books.
- Lakoff, G., and Nunez, R. (2000). *Where Mathematics Comes From*. New York: Basic Books.
- Langer, S. (1937). *Philosophy in a New Key*. Cambridge, MA: Harvard University Press.
- Langer, S. K. (1962). *Philosophical Sketches*. New York: Mentor Books.
- Lashley, K. S. (1951). The problem of serial order in behavior. In L.A. Jeffres (Ed.), *Cerebral mechanisms in behavior*. New York: Wiley, pp.110–33.

# PROOF

254 Jay Schulkin and Patrick Heelan

- Levi, I. (1967). *Gambling with Truth*. Cambridge: MIT Press.
- Levi, I. (2004). Beware of syllogism: reasoning and conjecture according to Peirce. In *The Cambridge Companion to Peirce*. Cambridge: Cambridge University Press.
- Levinson, S. (2003). *Space in language and cognition*. Cambridge: Cambridge University Press.
- Levinson, S. (2006). Cognition at the heart of human interaction. *Discourse Studies*, 8, 85–93.
- Levinson, S., and Jaisson P. (Eds) (2006). *Evolution and Culture*. Cambridge, MA: MIT Press, pp. 105–32.
- Loewenstein, G. (1994). The psychology of curiosity. *Psychological Bulletin*, 116, 75–98.
- Loewenstein, G. F. (1996). Out of control: Visceral influences on behavior. *Organizational Behavior and Human Decision Processes*, 35, 272–92.
- Martin, A. (2007). The representation of object concepts in the brain. *Annual Review of Psychology*, 58, 25–45.
- Martin, A. (2011). Circuits in Mind: The neural foundations for object knowledge. In M. Gazzamiga (Ed.), *The Cognitive Neurosciences*. Cambridge: MIT Press.
- Martin, A., and Caramazza, A. (2003). Neuropsychological and neurimaging perspectives on conceptual knowledge. *Cognitive Neuroscience*, 30, 195–212.
- Mead, G. H. (1928/1972). *The Philosophy of the Act*. Chicago: University of Chicago Press.
- Mead, G. H. (1934/1972). *Mind Self and Society*. Chicago: University of Chicago Press.
- Menary, R. (2010). *The Extended Mind*. Cambridge: MIT Press.
- Merleau-Ponty, M. (1942/1967). *The Structure of Behavior*. Boston: Beacon Press.
- Merleau-Ponty, M. (1962). *The Phenomenology of Perception*. New York: Routledge and Kegan Paul.
- Miller, A.D., and Goodale, M. A. (1995). *The Visual Brain in Action*. Oxford: Oxford University Press.
- Misak, C. J. (1991). *Truth and the End of Inquiry, A Peircean Account of Truth*. Oxford: Clarendon Press.
- Misak, C. (2004). *The Cambridge Companion to Peirce*. Cambridge: Cambridge University Press.
- Mithen, S. (1996). *The Prehistory of the Mind*. London: Thames and Hudson.
- Moreno, J. D. (1995). *Deciding Together*. Oxford: Oxford University Press.
- Moreno, J. D. (1999/2003). Bioethics is a Naturalist. In G. McGee (Ed.), *Pragmatic Bioethics*. Cambridge: MIT Press.
- Moscovitch, M., Behrmann, M., and Winocur, G. (1994). Do PETs have long or short ears? Mental imagery and neuroimaging. *Trends in Neuroscience*, 17, 292–4.
- Neville, R.C. (1974). *The Cosmology of Freedom*. New Haven, CT: Yale University Press.
- Noe, A. (2004). *Action in Perception*. Cambridge, MA: MIT Press.
- Oram, M. W., and Perrett, D. I. (1996). Integration of form and motion in the anterior superior temporal polysensory area (STPa) of the macaque monkey. *Journal of Neurophysiology*, 76, 109–29.
- Parrott, W. G., and Schulkin, J. (1993). Neuropsychology and the cognitive nature of emotions. *Cognition and Emotion*, 7, 43–59.

# PROOF

- Pashler, H. E. (1998). *The Psychology of Attention*. Cambridge, MA: MIT Press.  
[AQ115] Peirce, C. S. (1877). In *The Essential Peirce: Vol. 1*, p. 115.
- [AQ116] Peirce, C. S. (1878). Deduction, induction and hypothesis. *Popular Science and Monthly*, 13, 470–82.
- Peirce, C. S. (1883). *Studies in Logic, by Members of the Johns Hopkins University*. Boston: Little Brown.
- Peirce, C. S. (1889/1992). *Reasoning and the logic of things*. (K. L. Ketner, and H. Putnam, Eds). Cambridge: Harvard University Press.
- Peirce, C.S. (1892). The architecture of theories. *The Monist*, 1, 61–76.
- [AQ117] Perani, D., Cappa, S.F., Bettinardi, V., Bressi, S., Gorno-Tempini, Matarrese, M., and Fazio, F. (1995). Different neural systems for the recognition of animals and hand-made tools. *Neuroreport*, 6, 1,636–41.
- Perrett, D., Rolls, E. T., and Caan, W. (1982). Visual neurones responsive to faces in the monkey temporal cortex. *Exp. Brain Res.*, 47, 329–42.
- Perrett, D. I., Hietanen, J. K., Oram, M. W., and Benson, P. J. (1992). Organization and functions of cells responsive to faces in the temporal cortex. *Philosophical Transactions of the Royal Society of London, B* 335, 23–30.
- Pinker, S. (1994). *The Language Instinct*. New York: William Morrow and Co.
- Plato (1985). *Meno*. Oxford: Arisand Phillips.
- Premack, D., and Premack, A. J. (1995). Origins of human social competence. In M. S. Gazzaniga (Ed.), *The Cognitive Neurosciences*. Cambridge, MA: MIT Press, pp. 205–18.
- [AQ118] Pribram, K. (1991). *Brain and Perception: Holonomy and Structure in Figural Processing* (Hillsdale, NJ: Erlbaum).
- Prinz, J. L., and Barsalou, L.W. (2000). Steering a course for embodied representation. In Dietrich and E. Markman(Eds), *Cognitive Dynamics: Conceptual Change in Humans and Machines*. Cambridge, MA: MIT Press.
- [AQ119] [AQ120] Prinz, W. (2003). Experimental approached to action. In J. Rössler and N. Eilan (Eds), *Agency and Self-awareness*. Oxford: Oxford University Press.
- Pulvermüller, F., Shtyrov, Y., and Ilmoniemi, R. (2005). Brain signatures of meaning in action word recognition. *J. of Cognitive Neuroscience*, 6, 884–92.
- Quine, W.V.O. (1953/1961). *From a Logical Point of View*. New York: Harper Torchbooks.
- Quine, W.V.O. (1974). *The Roots of Reference*. La Salle, IL: Open Court.
- Rizzolatti, G., and Arbib, M. A. (1998). Language within our grasp. *Trends in Neuroscience*, 21, 188–94.
- Rizzolatti, G., and Luppino, G. (2001). The cortical motor system. *Neuron*, 31, 889–901.
- Roland, P. E., and Gulyas, B. (1994). Visual imagery and visual representation. *Trends in Neurosciences*, 17, 281–7.
- Rolls, E.T., and Deco, G. (2002). *Computational Neuroscience of Vision*. Oxford: Oxford University Press.
- Rolls, E.T., and Treves, A. (1998). *Neural Networks and Brain Function*. New York: Oxford University Press.
- Rolls, E. T., Treves, A., and Tovee, M. T. (1997). The representational capacity of the distributed encoding of information provided by populations of neurons in primate visual cortex. *Experimental Brain Research*, 114, 149–62.
- [AQ121] Ronchi, V. (1991). *Optics: A Science of Vision*(E. Rosen, Trans.) New York: Dover, orig. 1957.

# PROOF

256 Jay Schulkin and Patrick Heelan

- Rowlands, M. (2010). *The New Science of the Mind*. Cambridge: MIT Press.
- Rozin, P. (1976). The evolution of intelligence and access to the cognitive unconscious. In J. Sprague and A. N. Epstein (Eds), *Progress in psychobiology and physiological psychology*. New York: Academic Press.
- Rozin, P. (1998). Evolution and development of brains and cultures: some basic principles and interactions. In M. S. Gazzaniga and J. S. Altman (Eds), *Brain and mind: Evolutionary perspectives*. Strasbourg: Human Frontiers Science Program.
- Sabini, J., and Schulkin, J. (1994). Biological realism and social constructivism. *J. for the Theory of Social Behavior*, 224, 207–17.
- Schulkin, J. (2000). *Roots of social sensibility and neural function*. Cambridge, MA: MIT Press.
- Schulkin, J. (2004). *Bodily sensibility: Intelligent action*. Oxford: Oxford University Press.
- Schulkin, J. (2007). Effort and Will: A Cognitive Neuroscience Perspective. *Mind and Matter*, 5, 111–26.
- Schulkin, J. (2009). *Cognitive Adaptation: A Pragmatist Perspective*. Cambridge: Cambridge University Press.
- Schull, J. (1992). Selection: James's Principal Principle. In M. E. Donnelly (Ed.), *Reinterpreting the Legacy of William James*. Washington, DC: American Psychological Association, pp. 139–52.
- Schutz, A., and Luckman, A. (1973). *The Structures of the Life-World*. Evanston, IL: Northwestern University.
- Searle, J.R. (1983). *Intentionality*. Cambridge: Cambridge University Press.
- Sellars, W. (1962). *Science, Perception, and Reality*. New York: Routledge and Kegan Paul.
- Shapin, S. (1995). *A Social History of Truth*. Chicago, IL: University of Chicago Press.
- Shapin, S. (1996). *The Scientific Revolution*. Chicago, IL: University of Chicago Press.
- Shapiro, L.A. (2004). *The Mind Incarnate*. Cambridge, MA: MIT Press.
- Shepard, R. N., and Cooper, L. A. (1982). *Mental Images and Their Transformations*. Cambridge, MA: MIT Press.
- Shepard, R. N., and Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, 171, 701–803.
- Simon, H. (1967). *The Architecture of Complexity*. Cambridge, MA: MIT Press, p. 14.
- Skidelsky, E. (2008). *Ernst Cassirer: The Last Philosopher of Culture*. Princeton, NJ: Princeton University Press.
- Sperber, D. (1985). *On Anthropological Knowledge*. Cambridge: Cambridge University Press.
- Squire, L. R., and Zola, S. M. (1996). Structure and function of declarative and nondeclarative memory systems. *Proceedings of the National Academy of Sciences*, 93, 13,515–22.
- Stephan, K. M., Fink, G. R., and Passingham, R. E. (1995). Functional anatomy of the mental representation of upper extremity movements in healthy subjects. *Journal of Neurophysiology*, 73, 921–4.
- Stewart, J., Gepenne, O., and di Paolo, E.A. (2011). *Enaction*. Cambridge, MA: MIT Press.

# PROOF

Action and Cephalic Expression 257

- Tenenbaum, J.B., Kemp, C., Griffiths, T.L., and Goodman, N.D. (2011). How to grow a mind: statistics, structure and abstraction. *Science*, 331, 1,279–85.
- Thomas, E. (2001). Empathy and consciousness. *J. of Consciousness Studies*, 8, 1–35.
- Todes, D. P. (1989). *Darwin Without Malthus*. Oxford: Oxford University Press.
- Todes, D. P. (1997). Pavlov's physiology factory. *History of Science Society*, 88, 205–46.
- Tomasello, M. (2003). *Constructing a Language: A Usage Based Theory of Language Acquisition*. Cambridge, MA: Harvard University Press.
- Tulving, E. (1983/1993). *Elements of episodic memory*. Oxford: Clarendon Press.
- Tulving, E. (2002). Episodic memory: From mind to brain. *Ann. Rev. of Psychol.*, 53, 1–25.
- Ungerleider, L. G., and Haxby, J. V. (1994). "What" and "where" in the human brain. *Current Opinion in Neurobiology*, 4, 157–65.
- Ungerleider, L. G., and Mishkin, M. (1982). Two cortical visual systems. In D. Ingle, M. Goodale, and R. Mansfield (Eds), *Analysis of Visual Behavior*. Cambridge: MIT Press.
- Van Essen, D. C. (2005). Corticocortical and thalamocortical information flow in the primate visual system. *Prog Brain Res*, 149, 173–85.
- Van Essen, D. C., Anderson, C. H., and Felleman, D. J. (1992). Information processing in the primate visual system: An integrated systems perspective. *Science*, 255, 419–22.
- Varela, F.J., Thompson, E., and Rosch, E. (1991). *The embodied mind*. Cognitive Science and Human Experience Series. Cambridge, MA: MIT Press.
- Weissman, D. (2000). *A Social Ontology*. Yale University Press, p. 14.
- Wheeler, M. (1995). *Reconstructing the Cognitive World*. Cambridge: MIT Press.
- Whitehead, A. N. (1929/1958). *The function of reason*. Boston: Beacon Press.
- Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin and Review*, 9, 625–36.
- Wittgenstein, L. (1953/1958). *Philosophical Investigations*. (2nd ed.). (G. E. M. Anscombe, Trans.) Oxford: Basil Blackwell.
- Young, A.W. (1998). *Face and Mind*. Oxford: Oxford University Press.
- Zatorre, R. J. and Halpern, A. R. (2005). Mental concerts: Musical imagery and auditory cortex. *Neuron*, 47, 9–12.
- Zatorre, R. J., Belin, P., and Penhune, V. B. (2002). Structure and function of auditory cortex: music and speech. *Trends in Cognitive Science*, 6, 37–46.
- Zeki, S., Watson, J.D., and Finkenwirth, R.S. (1993). Going beyond the information given: The relation of illusory visual motion to brain activity. *Proceedings of the Royal Society of London – Series B, Biological Sciences*, 252, 215–22.

# PROOF

