EMBODIED MIND, MEANING, AND REASON

HOW OUR BODIES GIVE RISE TO UNDERSTANDING

MARK JOHNSON

MARK JOHNSON

Embodied Mind, Meaning, and Reason

HOW OUR BODIES GIVE RISE TO UNDERSTANDING

The University of Chicago Press Chicago & London

The University of Chicago Press, Chicago 60637
The University of Chicago Press, Ltd., London

Published 2017
Printed in the United States of America

26 25 24 23 22 21 20 19 18 17 1 2 3 4 5

ISBN-13: 978-0-226-50011-9 (cloth)
ISBN-13: 978-0-226-50025-6 (paper)
ISBN-13: 978-0-226-50039-3 (e-book)
DOI: 10.7208/chicago/9780226500393.001.0001

Library of Congress Cataloging-in-Publication Data
Names: Johnson, Mark, 1949– author.
Title: Embodied mind, meaning, and reason : how our bodies give rise to understanding / Mark Johnson.
Description: Chicago : The University of Chicago Press, 2017. | Includes bibliographical references and index.
Identifiers: LCCN 2017018048 | ISBN 9780226500119 (cloth : alk. paper) | ISBN 9780226500256 (pbk. : alk. paper) | ISBN 9780226500393 (e-book)
Subjects: LCSH: Mind and body. | Philosophy of mind. | Human body (Philosophy) | Meaning (Philosophy) | Consciousness. | Cognition.
Classification: LCC BF161 .J64 2017 | DDC 128—dc23
LC record available at https://lccn.loc.gov/2017018048

♾ This paper meets the requirements of ANSI/NISO Z39.48–1992 (Permanence of Paper).

CONTENTS

INTRODUCTION

Bringing the Body to Mind

This book develops an argument for the central importance of our bodies in everything we experience, mean, think, say, value, and do. It proposes an embodied conception of mind and then shows how meaning and thought are profoundly shaped and constituted by the nature of our bodily perception, action, and feeling. In short, it argues that we will not understand any of the issues that are so dear to philosophy until we have a deep and detailed understanding of how our embodiment gives rise to experience, meaning, and thought.

The view of mind, meaning, thought, and language that I elaborate here was anticipated, in part, in the writings of the American pragmatist philosopher John Dewey, and to a lesser extent in the works of William James and C. S. Peirce. However, I am not just serving up a heaping portion of warmed-over Dewey. Since Dewey's day we have had the privilege of important scientific and philosophical developments that supply crucial details about the processes of meaning and understanding that take us beyond what Dewey could provide. This research from the sciences of mind helps give flesh and blood to some of Dewey's more skeletal remarks about how organism-environment interactions generate meaningful experience. I do, nonetheless, remain a fan of Dewey's insistence on the key role of experience as the starting and ending point of any useful philosophical inquiry. Consequently, I take issue along the way with the orientation known as "linguistic" or "analytic" pragmatism, which grew mostly under the inspiration and influence of Richard Rorty, who saw philosophy as focused on language and what he called

"vocabularies," while rejecting any appeal to experience in the sense that Dewey understood that term.

My other important targets of criticism are traditional Anglo-American analytic philosophy of mind and language, along with what George Lakoff and I (1999) have called first-generation (disembodied) cognitive science. However, my focus is not primarily on the criticism of existing views, but rather on constructing a positive account of human meaning-making and understanding that draws on the cognitive science of the embodied mind. As I work up the details of that positive account, it will become clear how the cognitive science research on which I rely calls into question many key tenets of the analytic tradition in philosophy. The account of embodied mind, meaning, thought, and language developed in these essays runs directly counter to some of the fundamental assumptions in analytic philosophy and early cognitive science of the last seventy-five years. It behooves us, therefore, to begin with an explanation of why the body has mostly been ignored in mainstream analytic philosophy and its correlative conception of cognitive science.

The Invasion of the Body-Snatchers: Philosophy without the Body

When I was a graduate student in philosophy back in the mid-1970s, people did not have bodies. Well, perhaps I exaggerate a bit. What I mean is that a good deal of mainstream philosophy, both in Anglo-American and European traditions, acted as if our bodies aren't really that important for the structure of mind, and that our bodies don't play any significant role in anything that mattered to philosophers. What mattered to them, especially in so-called analytic philosophy that dominated the last three-quarters of the twentieth century in the Anglophone philosophical world, was language, concepts, logic, reason, knowledge, and truth. In all the massive literature that was generated on these topics from this analytic perspective, there is hardly any mention of the body, beyond the fact that one needs a body to secure perceptual inputs into our conceptual systems and knowledge structures, plus occasional recognition that we have feelings and emotions.

In this tradition, philosophy was defined by what Richard Rorty, borrowing a term from Gustav Bergmann, called the "linguistic turn." Bergmann described this turn as "the shared belief that the relation between language and philosophy is closer than, as well as essentially different from, that between language and any other discipline" (1967, 64–65).

He went on to emphasize the exclusively linguistic focus of philosophy when he said, "Generally, no philosophical question is ever settled by experimental or, for that matter, experiential evidence. Things are what they are. In some sense philosophy is, therefore, verbal or linguistic" (ibid., 65). In three short sentences, Bergmann has drastically restricted philosophy to linguistic analysis, and he denies any significant role for either experimental scientific research or experiential evidence! Here we have a vision of philosophy as an autonomous armchair discipline, entirely independent from science, and consisting of rational analysis of linguistic structures, terms, speech act conditions, and knowledge claims.

Rorty appropriately titles his highly influential anthology *The Linguistic Turn: Recent Essays in Philosophical Method* (1967), in which he collects many of the defining documents of what came to be known as "analytic" philosophy. In the introduction to that book, Rorty explains that "the purpose of the present volume is to provide materials for reflection on the most recent philosophical revolution, that of linguistic philosophy. I shall mean by 'linguistic philosophy' the view that philosophical problems are problems which may be solved (or dissolved) either by reforming language, or by understanding more about the language we presently use" (1967, 3).

The two methodological orientations that Rorty is describing came to be known as the "ideal language" and "ordinary language" perspectives. Those who lament the messiness, ambiguity, and incompleteness of everyday language argue that we need a clarified, precise "ideal language," if we ever hope to see how words have meanings and how genuine knowledge and truth are possible. Those who, like J. L. Austin (1970), see everyday speech as manifesting the accumulated insights and values of speech communities, argue that philosophical analysis should always start from distinctions embedded in ordinary language, even it if turns out that some of those distinctions are misleading and ought to be abandoned. In Austin's words, "Certainly, then, ordinary language is *not* the last word: in principle it can everywhere be supplemented and improved upon and superseded. Only remember, it *is* the *first* word" (1970, 185).

Consequently, linguistic philosophy went off in two different directions, one in search of a reconstructed ideal language of thought capable of expressing knowledge claims, and the other in search of an expansive mining of the conceptual resources embedded in ordinary language. Both movements, however, thought that linguistic analysis would eventually help us either to solve certain perennial questions about mind,

meaning, thought, and knowledge, or else to show them up as pseudo-problems that have needlessly perplexed us and ought to be jettisoned.

Now, the question I want to address concerning linguistic philosophy in either its "ideal language" or "ordinary language" versions is this: What is it about the character of this language-oriented philosophy that led it to almost completely ignore the body? The answer, I shall argue, is that (1) its exclusive focus on language as the object of philosophical analysis turned attention away from anything that was not linguaform, and (2) it operated with a remarkably impoverished, and scientifically unsound, view of language as entirely conceptual and propositional.

This seriously inadequate view of language resulted in large measure from the influence—on both the ideal language and ordinary language schools—of Gottlob Frege's celebrated conception of meaning and thought developed in a number of essays collected by Peter Geach and Max Black as *Translations from the Philosophical Writings of Gottlob Frege* (1966). In his classic 1892 essay, "Uber Sinn und Bedeutung" ("On Sense and Reference"), Frege hoped to validate the universal and objective stature of mathematical, logical, and scientific claims. In order to explain the alleged objectivity possible within these disciplines, Frege distinguished sharply between (1) the sign (the word or expression), (2) its reference (the object or state of affairs referred to), (3) its sense (the objective understanding, or the mode of presentation, of the reference), and (4) any subjective "associated ideas" that might be triggered in an individual's mind by a given sign. The *sense* was supposedly the public, shared meaning or understanding of the referred-to object or state of affairs, whereas the *associated idea* was merely an image or idea called up by a sign in the subjective mind of a particular individual. Frege claimed that it was the objective sense of a thought or proposition, not any associated ideas, that made shared understanding and knowledge possible. He summarized the relations between sign, sense, reference, and associated idea as follows:

> The reference and sense of a sign are to be distinguished from the associated idea. If the reference of a sign is an object perceivable by the senses, my idea of it is an internal image, arising from memories of sense impressions which I have had and acts, both internal and external, which I have performed. . . . The same sense is not always connected, even in the same man, with the same idea. The idea is subjective: one man's idea is not that of another. . . . This constitutes the essential distinction between the

> idea and the sign's sense, which may be the common property of many and therefore is not part or a mode of the individual mind. (Frege [1892] 1966, 59)

Notice that, in this famous passage, there is no mention of the body in relation to the sense of a sign. As presumably objective, senses supposedly cannot depend on the peculiarities of particular minds, let alone of particular bodies. They are universal and objective, in sharp contrast to associated "ideas," which depend on the body and experiences of those who have the ideas. Thus, Frege said, "One need have no scruples in speaking simply of *the* sense, whereas in the case of an idea one must, strictly speaking, add to whom it belongs and at what time" ([1892] 1966, 60). For example, the sense of the English word *mother* would allegedly be an abstract meaning or understanding "grasped" (to use Frege's term) by all who understand English. In addition, each of those individuals would have their own associated (and highly subjective) ideas that come to mind when he or she thinks about mothers, but none of this is held to be part of the objective sense of the term. Consequently, Frege claimed that senses are not dependent on the particulars of the bodies and brains that grasp them, so they constitute universal meanings, whereas associated ideas and images lay no claim to universality, precisely because they depend on our embodiment and experiences: "The reference of a proper name is the object itself which we designate by its means; the idea, which we have in that case, is wholly subjective; in between lies the sense, which is indeed no longer subjective like the idea, but is yet not the object itself" ([1892] 1966, 60).

Frege went on to argue that the *proposition*, not the word or concept, was the basic unit of meaning. Propositions have a subject-predicate structure. When the subject is specified and a concept is predicated of it, only then does the whole expression (i.e., the proposition) have a truth value (i.e., true or false). As a mathematician and logician, Frege was especially concerned with explaining how there could be shared, public meaning that provides a basis for objective knowledge and truth. His answer was that to understand the thought (i.e., proposition) expressed in a sentence is to grasp its public, universal sense, which is "not the subjective performance of thinking but its objective content, which is capable of being the common property of several thinkers" (Frege [1892] 1966, 62n.).

In order to explain the objectivity of the senses of terms, Frege pro-

posed what many consider to be a somewhat odd ontology consisting of three independent realms: the physical, the mental, and a third realm (to which he gave no name) that consists of abstract quasi-entities including senses, concepts, propositions, numbers, functions, and the strange objects "the True" and "the False." Because Frege believed that both physical (bodily) events and mental (psychological) processes are incapable of guaranteeing the objective and universal character of publicly shareable meaning and thought, he posited the third realm to house the objective contents of thought. Consequently, in this view, a theory of language need not pay any special attention to our embodiment, other than to notice how perception might be shaped by our bodily capacities.

With Frege, the die were fatefully cast. Few philosophers could fully embrace Frege's unusual ontological picture (especially his third realm), but the vast majority of so-called "analytic" philosophers agreed with his basic assumption that thought is propositional and relies on the objective senses of the component concepts of the proposition. They shared his view that thought is linguaform—that is, sentential, propositional, and conceptual in nature. Not surprisingly, one can find no serious account in Frege of the body's contributions to meaning and thought. This neglect of the body carried over into most of the major figures in the analytic tradition, such as Bertrand Russell, Rudolf Carnap, Carl Hempel, Gustav Bergmann, J. L. Austin, W. V. O. Quine, Donald Davidson, and a host of other philosophers, *none of whom had anything deep or extensive to say about the body's role in meaning and thought*. Even Hilary Putnam—who is much celebrated for his brain-in-a-vat thought experiments (1981), in which he emphasized that meaning requires a body interacting with a world—never supplied any detailed account of how the body shapes our thought and communicative practices. This is not to deny that there may be some insightful comments on embodiment scattered throughout their writings (especially in Wittgenstein and in Putnam); but their perspective remains mostly disembodied in its accounts of meaning, language, and thought. The overwhelming tendency in mainstream analytic philosophy of language is to begin with concepts more or less well formed, and then to analyze their relations to one another in propositions and to objects of reference in the world. This leads one to overlook the bodily origins of those concepts and patterns of thought that constitute our understanding of, and reasoning about, our world.

What is at stake here is not just analyses of the meaning of particular terms or sentences, but something much more important: the very nature of meaning and thought as grounded in and shaped by our human embodiment. Moreover, the na-

ture of philosophy itself is called into question, once we realize that it is inextricably tied to our embodiment! At issue here are the origins of meaning, language, and all our forms of symbolic expression and communication that define our world and our personal identity. Although most analytic philosophers are not strictly Fregean, Anglophone philosophy of language developed mostly in this "disembodied" Fregean mode, in the sense that a theory of meaning, thought, and language is given without any serious study of the workings of the body and brain in how we make and communicate meaning or how we think. In the seventy or so years since the emergence of the field of philosophy of language, there has been remarkably—and depressingly—little variance from these early ideas that (1) language is conceptual and propositional, and (2) other than noting that we need a body to have perceptual inputs, it is asserted that concepts, propositions, and thoughts are not profoundly shaped by the nature of our bodily capacities and modes of engagement with our material environments.

The first-generation cognitive science that developed within this linguistic framework was therefore a blending of analytic philosophy of language, Chomskyan generative linguistics, information-processing psychology, computer science, and budding artificial intelligence research—all of which were relatively disembodied perspectives. Moreover, in line with Bergmann's dismissal of empirical scientific research and experiential evidence (as supposedly being irrelevant to the primary analytic tasks of philosophy), early cognitive "science" seemed often to be driven more by armchair philosophical assumptions than by empirical research on cognition. Consequently, little of the vast scientific research on how our bodies and brains underlie cognition found its way into the philosophy of mind and language during most of the twentieth century.[1]

Although Rorty's particular version of the linguistic turn obviously does not adequately represent all the methods and perspectives that make up so-called "analytic" philosophy, I suggest that his view of language and philosophy captures several of the deepest assumptions and motivations of the larger movement of linguistic philosophy. Consequently, a brief account of his argument about the priority of language gives us a good understanding of why the body played little or no role in large parts of analytic philosophy for most of the previous century. Fifteen years after *The Linguistic Turn*, Rorty collected several of his essays into another influential volume, entitled *Consequences of Pragmatism* (1982). In this book, Rorty articulates his particular interpretation of

what some of his followers would later call linguistic (or analytic) pragmatism. He correctly praises pragmatism for its antidualism and its antifoundationalist view of knowledge and truth. However, he then goes on to claim, mistakenly, that antidualism and antifoundationalism require a concomitant rejection of any metaphysical commitments whatsoever. In particular, Rorty argues that when Dewey claimed that philosophy must start from "experience" in all its fullness, and then develop an "empirical metaphysics" that identifies recurrent structures and characteristics of all experience, Dewey was inconsistently falling right back into the very foundationalist metaphysics he had earlier so deftly criticized. Rorty sums up this critique as follows: "Dewey's mistake—and it was a trivial and unimportant mistake, even though I have devoted most of this essay to it—was the notion that criticism of culture had to take the form of a redescription of 'nature' or 'experience' or both" (ibid., 85).

Rorty liked Dewey's insightful criticism of deeply rooted epistemic and metaphysical assumptions that underlie different cultural systems, but he could not abide the idea that such a criticism might need to be based on a metaphysics of "experience" or "nature." Contrary to what he says in the previous quotation, Rorty's large corpus of later writings makes it quite clear that he did not think Dewey's "mistake" was trivial, insofar as Rorty saw such a project as leading us to an illegitimate foundationalist metaphysical program. Rorty conceives of philosophy as linguistic analysis and, where appropriate, criticism of our language games and linguistic practices. However, he also insists that we cannot carry out such analysis or criticism by claiming to "ground" it in some allegedly mind-independent "experience" or "world." According to Rorty, language communities operate with what he calls "vocabularies" that carry with them norms for what can be said and how any claim can be validated or criticized. We inhabit these vocabularies, but we cannot step outside any one of them to compare it to the world beyond language that it pretends to describe:

> This Davidsonian way of looking at language . . . lets us see language not as a *tertium quid* between Subject and Object, nor as a medium in which we try to form pictures of reality, but as part of the behavior of human beings. On this view, the activity of uttering sentences is one of the things people do in order to cope with their environment. The Deweyan notion of language as tool rather than picture is right as far as it goes. But we must be careful *not* to phrase this analogy so as to suggest that one can separate the tool, Language, from its users and inquire as to its "adequacy" to achieve

> our purposes. The latter suggestion presupposes that there is some way of breaking out of language in order to compare it with something else. But there is no way to think about either the world or our purposes except by using our language. One can use language to criticize and enlarge itself, as one can exercise one's body to develop and strengthen and enlarge it, but one cannot see language-as-a-whole in relation to something else to which it applies, or for which it is a means to an end. (Rorty 1982, xviii–xix)

From Rorty's perspective, we dwell in our vocabularies and language games, and we cannot extricate ourselves from some particular vocabulary to see how it—or any other vocabulary, for that matter—might or might not map onto a mind- and language-independent reality that we call either "experience" or "the world." Moreover, since Rorty defines metaphysics as precisely such attempts to see how language could be foundationally grounded in "reality" or "the world," he rejects all metaphysics out of hand. That is why he denigrates Dewey's "empirical metaphysics," and *any* metaphysics, for that matter. For Rorty, there can be *no* good metaphysics.

Here's where embodiment gets discarded, insofar as Rorty thinks that any talk of "embodied" or "body-based" cognition, thought, meaning, experience, or language could be nothing but one more misguided attempt to find absolute foundations for our preferred vocabularies. In other words, Rorty is led to regard reference to "the body" as nothing but one more deluded metaphysical grounding or founding device, which throws us right back into illegitimate and unachievable foundationalist metaphysical and epistemological projects. Therefore, it should not be surprising that Rorty has little or nothing to say about embodiment, the body, and experience, and that he has almost nothing significant to say about aesthetics, insofar as aesthetics focuses on the embodied patterns, images, schemas, feelings, emotions, and qualities through which we experience things and events in our world. Nor should it be surprising that Rorty eschews utilization of any privileged scientific accounts of mind, cognition, meaning, language, and values, since he thought of them as just one more vocabulary among the many we might employ to order our lives. In short, for Rorty, "the body" and "experience" are linguistic and textual notions that only have meaning in the context of some particular vocabulary, and so they have no special status for grounding philosophy.[2]

As I said earlier, Rorty does not speak for all linguistic philosophy, but he perceptively understands the deepest motives and implications of

the linguistic turn. He accepts the defining idea of linguistic philosophy as exclusively focusing on language. He regards thought as linguaform. And he thinks that any attempt to talk about "experience" or "the body" as the ground of meaning and thought is a reversion to misleading and counterproductive foundationalist metaphysical systematizing. It comes as no surprise, then, that there is little or no talk in linguistic philosophy about the body and its role in meaning and thought.

Incidentally, it was not just analytic philosophy that overlooked our embodiment. There were parallel dismissals in certain strands of European philosophy. Although phenomenology—especially the variety developed in the later Husserl, in Heidegger, and in Merleau-Ponty—did most certainly acknowledge the fundamental role of embodiment in our experience, the more deconstructivist developments in the 1960s and '70s had no place for the body as a locus of meaning and thought. Among those who recognized the importance of our bodies, Edmund Husserl made remarks on the body that are fairly general; and, in my opinion, he never overcame his reliance on a transcendent ego as the ultimate unifying agent of thought and action (Stawarska 2009). Martin Heidegger's key notions of "earth" and "world" also evoke bodily experiences, but his criticism of science and technology left most Heideggerians uninterested in scientific treatments of embodied cognition. Following Husserl and Heidegger, Maurice Merleau-Ponty uncovered the central role of the "lived body" in how our world opens up to us, but, unfortunately, Merleau-Ponty's philosophy of the body came to be eclipsed by Jacques Derrida's attack on what he called the "metaphysics of presence." Derrida argued that words have meanings, not by indicating nonlinguistic realities, but only through a network of relations and differences with other terms in one's language. There is no way to specify "the" meaning of a term by connecting it up with some underlying reality. Instead, each term takes its place in a web of other terms, all of which are mutually interdefined. Consequently, most deconstructionists (whether faithful to Derrida's original insights or not) reject any attempt to ground language in experience or bodily processes. Indeed, they rejected any notion of grounding whatsoever. "The body" then gets discarded along with every other "metaphysics of presence," just as Rorty dismisses any reference to embodiment as reversion to foundationalist metaphysical speculation. And it goes almost without saying that people disposed toward this deflationary view of meaning had no interest in, or even tolerance for, scientific claims about how bodies and brains generate meaning, thought, and language. In short, just as clas-

sical pragmatism's recognition of the body was buried by the upsurge of analytic philosophy, likewise the appreciation of the body's role in meaning and thought developed within phenomenology and hermeneutics was criticized and marginalized by the popular deconstructionist games of the late twentieth century.

Consequently, when I found myself immersed in linguistic philosophy as a graduate student in the 1970s, I did not even realize that I had been plunked down in a landscape that had been invaded by the body snatchers, who had systematically scoured the philosophical landscape to remove bodies from the scene. Although I did not then understand why, I found myself increasingly alienated from the rigorous linguistic analysis that most of my peers regarded as the very heart and soul of philosophy. I appreciated the rigor and thoroughness of those analyses, but they too often failed to engage me at an existential level that was meaningful and ethically motivating. I felt unmoored and adrift in conceptual analyses of the logic and structure of scientific laws and knowledge claims, analyses of the emotive use of value terms, inquiries into referential opacity and indeterminacy of translation within and across conceptual systems, and accounts of logic as conventional relations among arbitrary symbols. As an undergraduate in the late 1960s, I had fallen in love with philosophy because I thought it could help me understand who I was, whether (and if so, how) my life might be meaningful, and how I ought to live. But there I was, instead, in a top-tier graduate program, trying to answer W. V. O. Quine's question about whether there was any way for someone who does not speak the language of an alien group to learn what *gavagai* means in their language (Quine 1960). There I was, asking whether, in the presence of the animal I call a "rabbit" my alien interlocutor utters "*gavagai*," she *really* means "rabbit," "undetached rabbit part," "rabbit stuff," or none of the above! There I was, trying to figure out where Frege's third realm existed, since I had been assured that it was what made objective knowledge and truth possible. There I was, trying to give arguments for C. L. Stevenson's emotivist view that moral expressions like "*X* is good" really mean "I approve of *X*; do so as well." I found that I could not shake off the nagging suspicion that none of this had much to do with the lives of ordinary folks like me.

And so I quit philosophy. Not forever, but for two years. The questions about meaning, purpose, values, and knowledge that had enticed me into philosophy as an undergraduate at the University of Kansas in the late 1960s were not the questions being asked in the world of professional analytic philosophy. When I returned to graduate school, I did

so because I found teachers who were talking about things that might actually relate to the quality and direction of my life. At the time, I didn't understand what this really had to do with embodiment; but that would come later.

Retrieving the Body from the Body Snatchers

It was not until I took courses from Paul Ricoeur—one on metaphor, a second on hermeneutics, and a third on imagination—that I began to see that there was an experience of meaning and value that went deeper than language. I learned to see the entire hermeneutic (interpretive) process of understanding not merely as an intellectual and linguistic act, but rather as constituting our whole embodied way of being in, and making sense of, our world. Ricoeur taught me that understanding is not just a conceptual achievement, but rather a whole-body, visceral engagement with our world that defines who we are and how we comport ourselves. Ricoeur had little interest in the scientific study of meaning and thought, but as a brilliant phenomenologist and hermeneutic thinker, he understood the body's role in meaning, reasoning, imagining, and communicating a sense of the world.

At the same time, I came under the humane influence of Ted Cohen, who was not explicitly interested in embodiment, but whose courses on J. L. Austin and on the philosophy of art led me to probe beneath language proper into the very conditions that make it possible for us to experience meaning and to communicate with one another. These dimensions of meaning and value are manifest partly in language, but also in nonlinguistic events in painting, sculpture, music, dance, architecture, film, ritual practices, spontaneous gesture, theater performance, and so on. And so I came to see aesthetics as involving more than just a theoretical investigation into art and aesthetic judgment. I came to conceive of aesthetics more broadly as a general exploration of how humans make and experience meaning at many different levels of our engagement with our world. Both Ricoeur and Cohen, although coming from very different philosophical traditions, helped me begin to see some of the embodied and imaginative dimensions of human cognition and understanding.

Eventually, I had the good fortune to write my doctoral dissertation on metaphor with Cohen and Ricoeur on my committee. It was there that I got my first glimpse of how metaphor is not just a matter of words, not merely a linguistic device, but instead a basic irreducible imaginative

process by which humans are able to recruit body-based meaning for abstract conceptualization and reasoning. I also had a vague suspicion that the constitutive role of metaphor in thought called into question large parts of the linguistic philosophy framework that dominated the philosophical scene at that time, and so I began to realize that the assumptions and methods of analytic philosophy were not up to the task of giving an adequate account of the richness and visceral depths of human meaning.

For me, a fuller appreciation of the bodily dimensions of meaning would not come until later, after I went out to Berkeley as a young visiting assistant professor for the winter and spring of 1979 and met the linguist George Lakoff. In our intense and far-ranging conversations about metaphor, meaning, and thought, Lakoff and I saw immediately that abstract conceptualization and reasoning depend on conventionalized conceptual metaphors that could not be adequately accounted for by the reigning philosophical and linguistic traditions of the day. Those traditions were objectivist and literalist. They assumed the objectivist theory that a language is a formal system consisting of a set of meaningless symbols embedded within innately grounded syntactic frameworks and ordered by means of logical relations, which are themselves simply possible orderings of symbols. Ordinary language was modeled as a formal language system of this sort, and the meaningless symbols were supposedly given meaning by being placed in referential relations to objects, properties, and relations in the mind-independent world, usually by means of some idealized model of the world. According to this objectivist view, the meaning of any cognitively significant expression has to be literal, insofar as the sentence has to map directly onto states of affairs in the world (see Searle 1979, 117f). It followed from this objectivist orientation that a metaphor, which was taken to be merely a condensed statement of similarities between two different domains, has its meaning (if, indeed, it has any distinct meaning at all) only as a set of proper literal concepts and propositions to which it can be reduced. The other alternative theory popular at that time claimed that there is no distinctive metaphorical meaning, beyond the literal meaning of the words used in the utterance, and so metaphor is merely a pragmatic, not a semantic, device (Davidson 1978; Rorty 1989).

Those who regarded metaphor as a semantic phenomenon typically preferred the twenty-five-centuries-old comparison (or similarity) theory, according to which the meaning of a metaphor consisted merely in a set of literal similarity statements, so that a metaphor "*A* is *B*" was supposedly reducible to "*A* is like *B*, in respects *X*, *Y*, *Z*" (where *X*, *Y*, and

Z are literal similarities (either properties or relations) between objects *A* and *B*). Lakoff and I realized that if *that* was all a metaphor was, then it was no surprise that it was regarded as a linguistically and philosophically unimportant figure of speech that could, and should, ultimately be replaced by a string of statements about literal similarities existing between two domains of experience.

Once we came to see that virtually all our abstract concepts are defined by multiple metaphors that could not be reduced to literal similarity statements, Lakoff and I realized that something was rotten in the state of analytic philosophy of language. *Metaphors We Live By* (1980) was our first attempt to explore the implications of the central role of metaphor in our everyday—as well as our theoretically sophisticated—conceptualization and reasoning. Those implications were far reaching and stunning. Metaphor would need to be moved from the distant periphery to the very center of the study of thought and language. To understand how metaphors work, it was necessary to set out the cross-domain mapping structure that constituted the conceptual metaphor being studied, and then to show how each part of the mapping gives rise to polysemous terms and phrases and also to inferences within the source domain that get carried over into the target domain. For example, the UNDERSTANDING IS SEEING metaphor maps the entities, properties, and relations of the source domain (vision) onto the target domain (understanding) as follows:

- An object seen maps onto an idea or concept understood.
- Shedding light maps onto "illuminating" an idea.
- Seeing an object clearly maps onto understanding an idea.
- Visual acuity maps onto intellectual "vision" or "insight."
- An object blocking our view maps onto something that obstructs understanding.

A cross-domain mapping of this sort (here, from vision to understanding) is a conceptual metaphor that gives rise, via each of the submappings within the metaphor, to the use of terms with multiple related meanings (polysemy). In this manner, terms relating to vision (such as *see*, *light*, *illuminate*, *obscure*, *brilliant*, *blind*) have related meanings appropriate both for visual experience and also for the processes of intellectual understanding (e.g., "I *see* what you mean now"; "Could you *shed a little more light* on the last part of your theory?"; "That was a terribly *illuminating* explanation"; "What she said was very *enlightening*"; "I've been

so *blind* to what she was up to"). Moreover, beside the polysemous use of terms based on the metaphor mapping, we also *think* and *draw inferences* via the metaphor.

Many linguists and psychologists began to study how inferences from the source domain are used to make target-domain inferences (Gibbs 1994; Kovecses 2010; Dancygier and Sweetser 2014). For instance, if something is obscuring or blocking your line of sight, then you cannot see whatever is behind it fully or clearly (visual domain inference); correspondingly, if some idea or thought is dominating your understanding, you will not fully discern some competing or alternative idea (target domain inference). To sum up, the cross-domain mappings are conceptual and support both systems of meaning and patterns of inference and reasoning.

And where is the *body* in all of this? Well, as we investigated how metaphors work, we discovered that the source domains of common cross-cultural metaphor systems are typically based on our sensory, motor, affective, and interpersonal experiences and cognitive capacities, all of which involve our embodiment. In other words, metaphors are shaped by the nature of our bodies and brains as we engage our physical and social environments. Metaphors thus "recruit" sensory and motor experience and inferential patterns to perform abstract conceptualization and reasoning. It is in this sense that they are body based. We saw that, contrary to the reigning comparison theory, metaphors are not typically based on perceived literal similarities between two different domains (e.g., vision and understanding aren't significantly similar), but rather are based on *experienced correlations* between the source and target domains. A few years later, Joseph Grady (1997), a student of Lakoff who was investigating why we have the metaphors we do, developed a theory of how these cross-domain experiential correlations are learned unreflectively, simply by growing up with a body of the sort we have, interacting with environments of the sort we inhabit. Grady called these basic metaphors "primary" because they emerge naturally in our embodied experience, through the coactivation of sensory-motor experiences and "higher level" thought processes that establish reentrant neural connections between the source and target domains that make up the metaphor. These primary metaphors could then be combined to generate more complex metaphor systems. Although these body-based metaphors (e.g., Understanding Is Seeing, More Is Up, Causes Are Physical Forces, Thinking Is Moving, Temporal Change Is Relative Motion) are good candidates for metaphorical universals

tied to our shared embodiment, it is also the case that different cultures tend to elaborate the primary metaphors in different ways, giving rise to cultural variations in their meaning and use.

The fact that these conceptual metaphors are based on experiential correlations between aspects of the source and target domains revealed the crucial role of our embodied interactions with our environment in our ability to experience and make meaning. Consequently, the whole illusion of disembodied meaning went out the window! Also discarded were any dualistic and disembodied views of mind and cognition. In short, taking metaphor seriously required a massive rethinking of some of our most deeply rooted views about meaning, thought, and symbolic expression that had defined objectivist, literalist, and disembodied views. What was needed, then, was a philosophical and scientific perspective rich enough to explain these aspects of embodied meaning and thought.

The Embodied Cognitive Science of Meaning and Thought

In the emerging cognitive sciences, up through the mid-1970s, the dominant orientation, which Lakoff and I dubbed "first-generation cognitive science," was a blend of generative linguistics, information-processing psychology, analytic philosophy of language, and artificial intelligence (Lakoff and Johnson 1999). It got its conception of grammar from Chomsky's claims about innate formal structures. It got its view of natural languages as formal languages from logic and computer science. It got its conceptual and propositional focus from linguistic philosophy and generative linguistics, and it got its functionalist conception of mind from artificial intelligence and computer science. It got its view of mental operations from the cognitive psychology of the day, which parsed thought into a series of discrete mental operations on perceptual inputs, carried out sequentially, and eventually issuing in behavioral outputs. Mind was taken to be a capacity for formal operations and functions that was not dependent on any one particular form of embodiment. The Mind Is A Computational Program metaphor captured people's imaginations so thoroughly that they seemed not to notice how our bodies play a crucial role in what is meaningful to us, how we think about it, and how we communicate our insights.

Fortunately, by the mid-1970s there was a mushrooming interdisciplinary body of empirical work on cognition that began to challenge the most basic assumptions of the dominant first-generation paradigm

(Patricia Churchland 1986; Lakoff 1987; Varela, Thompson, and Rosch 1991). Studies of human conceptual systems changed our understanding of how our categories and concepts are structured. Research in cognitive neuroscience challenged mind/body dualism and revealed the importance of feeling and emotion in all thought processes (Damasio 1994, 1999, 2003, 2010). New research on how humans actually reason challenged our inherited Enlightenment faculty psychology and its conception of a pure, nonemotional rationality. The new orientation known as cognitive linguistics challenged Chomsky's innatist views about language by drawing on empirical studies of the syntax, semantics, and pragmatics of natural languages (Lakoff 1987; Langacker 1987–91; Talmy 2000). In this new view, language was seen to be mostly a development of cognitive capacities for perception, bodily motions, and action. The result was an emerging vision of embodied mind, meaning, and thought that Lakoff and I (1999) later named "second-generation (or embodied) cognitive science." This was not a completely unified and monolithic perspective, but at first a somewhat disparate collection of empirical research programs that began to provide converging evidence for the central role of our brains and bodies in everything we experience, think, and do.

One of the central tasks of this second-generation orientation was to determine how—precisely and in detail—our bodies give rise to the meaning we can experience, the reasoning we do, and the ways we communicate with others, not just through language proper, but also through all our many forms of symbolic action in the arts and associated practices. Where was one to look for evidence of embodied cognition? Since cognitive neuroscience was a new and relatively immature field in the late 1970s, it took a while for it to establish a more unified identity.

At the same time, however, in addition to the psychological research and linguistic theory mentioned above, there were philosophical resources for developing a broad theoretical explanation of embodied mind, meaning, and thought. One of those was pragmatist philosophy, to which I was introduced in the early 1980s by my colleague Tom Alexander in a seminar he was teaching on John Dewey's classic *Experience and Nature* (1925).[3] I began to see the pragmatism of C. S. Peirce, William James, and John Dewey as the most appropriate nondualistic and scientifically responsible framework for understanding human experience and cognition. I got a glimpse of the central role of our bodies in the habits of action and thought that define who we are, and in the patterns and qualities that make meaning possible for us.

The other useful philosophical perspective was phenomenology. A year or two earlier I had sat in on a seminar taught by another colleague, Glenn Erickson, on Maurice Merleau-Ponty's *Phenomenology of Perception* (1962), where I had come away with a similar aha moment about what a truly nondualistic, experientially based philosophy would look like—one that placed the *lived body* at the center of human reality. In spite of their very different philosophical styles and temperaments, I came to see that Dewey and Merleau-Ponty were exploring the same deep dimensions of embodied meaning and thought, in ways that were being mostly ignored in the mainstream philosophical traditions of the day.[4]

It was this philosophical background—along with a general familiarity with speech-act theory, phenomenology, and hermeneutics—that I brought into my ongoing discussions with George Lakoff about what a new embodied view of mind, thought, and language would involve. It was Lakoff, though, who convinced me that I had to pay attention to the burgeoning cognitive sciences and could no longer rely solely on my philosophical training, precisely because some of that training was profoundly at odds with the scientific research on mind. We couldn't keep doing philosophical business-as-usual. Though it sounds arrogant to say so, what was needed was a new philosophical perspective, and not merely some tinkering with existing methods and orientations. This new emerging, second-generation cognitive science perspective might be essentially pragmatist in character, but it would have to incorporate the half century of cognitive science that had emerged since the heyday of classical pragmatism. In doing so, it would give important new details of embodied cognition not available to Dewey in his day. What was needed was what Patricia Churchland (1986) called a "co-evolution" of philosophy and science to generate an empirically responsible philosophical theory of mind, thought, and language.

What follows are some of the key components of this emerging embodied, interdisciplinary framework that is giving rise to a new understanding of mind, meaning, thought, and language.

Meaning Arises from Organism-Environment Interactions

What we call "mind" is an emergent character of an ongoing series of interactions among certain kinds of organisms and their environments (Dewey [1925] 1981). Mind is not a metaphysical entity or fixed structure, and it cannot possibly exist independent of bodily processes, activities, and engagements with other people. Instead, mind has reality only

as an emergent process of meaning-making, acting, and communicating among creatures capable of certain kinds of complex functions and communicative interactions (Merleau-Ponty 1962; Varela, Thompson, and Rosch 1991). In order to see this, one has to start where all animals start: with a bounded, embodied organism as it engages its various environments in ways that allow it to maintain the basic conditions for life and growth. The more complex the organism is, the more ways it has by which it can meaningfully interact with the energy structures that make up its environments. Depending on the specific bodily makeup of the organism, particular situations will provide for the organism what James Gibson (1979) called "affordances"—patterns for meaningful perception and action relative to the nature of the organism, its needs, and its purposive activity in the world that it inhabits. For example, for human animals of our size, makeup, and interests, certain caves afford relations of containment (here, as space for habitation) and they may afford, for certain animals, protection from the elements and predators. Small caves do not afford access for large mammals, such as elephants, and so such enclosures do not have the same meaning to elephants as they do to humans. The world of an animal is demarcated by a large number of affordances provided by various objects, spaces, and structures within that animal's environment. What we call "objects" are affordances relative to the kinds of creatures we are.

Notice that, already at this basic level of animal-environment affordances and transactions, I have spoken of the "meaning" of specific environmental structures for a certain type of creature. I use "meaning" here for any experiences enacted or suggested by various affordances in our surroundings (M. Johnson 2007). Any aspect or quality of a situation means (for a specific type of creature) what it calls forth by way of experience. That includes past experiences, present experiences, and projected future experiences perceived to be possibilities developing out of one's current situation.

There are at least two very important consequences of this conception of meaning. (1) It acknowledges our evolutionary continuity with many other species, and therefore allows that certain nonhuman animals might be capable of various sorts of meaning-making. However, species lacking capacities for abstraction and symbolic interaction will have available to them a very attenuated range of meanings, relative to the richness of meaning available to humans; but it will be meaning nonetheless. (2) Conceiving of meaning in this embodied, experiential manner enables us to go beyond the narrower confines of language-

based meaning to embrace the full range of human meaning-making in such practices as painting, sculpture, music, architecture, dance, spontaneous gesture, and ritual practices, in a way that no merely linguistically centered account of meaning can. No traditional understanding of signs as having meaning only through some conceptual/propositional content grounded in reference to states of affairs in the world could even begin to capture the richness of body-based meaning that is experienced in all these varied forms of human meaning-making and communicative activity.

Body-Part Projections

One important way that the body undergirds languages and systems of meaning the world over is the use of *body-part projections* for understanding objects, events, and scenes (Lakoff and Johnson 1999, chap. 3; Talmy 2000). We use our own body-part relations to make sense of objects and spatial relations in our surroundings. A good example of this is the way we experience our own bodies as having *fronts* and *backs*, and so it seems natural for us to project these *front/back* relations onto other objects, such as trees, rocks, houses, and lines of people, none of which have inherent fronts or backs. We experience computer screens as *facing* us, when we sit *in front of* them. We tend to project fronts onto moving objects (cars, buses, airplanes, ships), with the front defined relative to the canonical direction of motion for the object. Cars, buses, airplanes, and ships mostly move forward, and so their "front" is specified by that direction of motion. If they reverse direction, they are then said to "back up." We extend this *front/back* orientation even onto simple physical objects like bottles, balls, and rocks. For instance, if I rotate a plastic water bottle into a horizontal orientation, and then move the bottle in a line through space, you will project a front onto the bottle based on the direction of its motion.

The relation *in front of* is defined in most languages relative to the space between a viewer and some object in their field of vision. So, a dog that is located between me and a tree is experienced by me as in front of that tree, as if the tree faces me. Some languages, such as Hausa, reverse this, projecting the front of the tree as facing away from the viewer. Therefore, in Hausa, a dog who is located between the viewer and the tree would be "behind" or "in back of" that tree; and if the tree is between me and the dog, then the dog is described as "in front of" the tree. However, despite such orientation reversals between English and Hausa,

in both cases the *in front of* relation is the result of a body-part or body-orientation projection, and so the meaning of spatial relations phrases is body relative. It is also common to experience objects such as mountains, trees, towers, poles, and people as being oriented *up* and *down*, as having *tops* and *bottoms*, and often as having *heads* and *feet* (as in the *foot of* a mountain, tree, or tower). Moreover, as will be discussed below, we use body-part terms imagistically and metaphorically when we conceive of rivers as having *arms*, and when we attribute body parts like *eyes* and *hearts* to objects and events, such as the *eye* of a needle or a storm, or the *heart* of an artichoke or a problem.

Image-Schematic Affordances

Body-part projections are meaningful because they enact aspects of our fundamental ways of relating to, and acting within, our environment. The way our perceptual and motor systems get characteristically wired up (neuronally) as we grow and develop—through ongoing relations with energy patterns in our environment—establishes a large number of recurring, intrinsically meaningful patterns that George Lakoff (1987) and I (M. Johnson 1987) dubbed "image schemas." The basic idea was that, given the nature of our bodies (how and what we perceive, how we move, what we value) and the general dimensions of our surroundings (stable structures in our environment), we will experience regular recurring patterns (such as up/down, left/right, front/back, containment, iteration, balance, loss of balance, source-path-goal, forced motion, locomotion, center/periphery, straight, curved) that afford us possibilities for meaningful interaction with our surroundings, both physical and social. For example, the fact that humans exist and operate within earth's gravitational field generates recurring experiences of *up/down* (i.e., *verticality*) relations. We understand objects as rising *up* and falling *down*, as *up*right or lying *down*, as *on top of* or *below* (or *under*), relative to our own bodily orientation and our physical surroundings. The fact that we routinely, and crucially, experience *balance* or *lack of balance* gives rise to a BALANCE schema that applies literally to balancing physical objects and metaphorically to our internal bodily states, to mathematical equations, and to notions of political fairness and justice (M. Johnson 1987; 1993). Through our numerous daily experiences with containers and contained spaces we develop a CONTAINER schema that consists of a boundary that defines an interior and an exterior (Lakoff 1987; M. Johnson 1987). Our thousands of daily encounters with moving

objects and with moving our own bodies gives rise to a LOCOMOTION schema (Dodge and Lakoff 2005). Importantly, such schemas are typically multimodal, and so are not tied to any single sensory or motor area of the brain. This multimodality is evident when we experience containment both through vision and touch, or when we see something far off and also hear it as far away.

A list of common image schemas might run into the scores or even hundreds (Cienki 1997; Hampe 2005). Cross-cultural analysis cannot yet verify any definite list of universal image schemas, but schemas such as CONTAINER, SOURCE-PATH-GOAL, VERTICALITY, and COMPELLING FORCE would appear to be excellent candidates, insofar as people the world over routinely have experiences that manifest such patterns. Ellen Dodge and George Lakoff conclude that, although all languages do not have the same spatial-relations concepts, nevertheless, they appear to build their particular spatial relations from "a limited inventory of basic primitive image schemas and frames of reference" (2005, 71).

Image schemas are meaningful to us both *before* and *beneath* linguistic meaning. They are intrinsically meaningful embodied structures. Focusing on image-schematic structure was my first attempt to figure out how the body might give rise to meaning. Lakoff and I realized that image schemas perform an important role in structuring the source domains of primary metaphors, and this was partly what it meant to say that conceptual metaphors are "embodied," "grounded in our bodies," and "experientially based." In *The Body in the Mind* (1987) I therefore gave numerous examples of how conceptual metaphors appropriate the image-schematic structure (relations and logic) of the source domain for abstract conceptualization and reasoning. Consider the SOURCE-PATH-GOAL schema that is present in all our experiences of seeing an object move along a path or moving ourselves from an initial location to a (temporary) terminal location. The SOURCE-PATH-GOAL schema manifests a recurring pattern for moving objects in our experience, and it has its own distinctive corporeal or spatial logic. So, if two objects start out at the same source location, moving along the same path, at the same speed, then they will both reach the same location at the same time. This is an inference grounded in our experience of moving objects. If one moves faster than the other, then the faster-moving object will arrive at the goal destination sooner. If I have moved halfway along a path from *A* to *B*, then I have "covered" all the points on the path up through the halfway point.

Such knowledge of source-path-goal movements may seem quite banal and mundane, but it is nonetheless the basis for spatial and temporal inferences we make about moving objects and the path on which they move. Moreover, if we later come to understand the path of motion metaphorically as the "path" or "course" of a temporal process, then we can appropriate the logic of spatial motion, plus our knowledge about moving objects, to draw appropriate inferences in some abstract domain, such as the domain of state change. For example, if we understand a causative change-of-state process metaphorically as motion along a path, from one state-location to another, then we can use the logic of moving objects to understand processes such as change of state. Thus, we speak of water on the stove as *going* from cold to hot in minutes. And if the water is getting hotter, then it is progressively getting less cold, and at some time in the process, it will cease to be cold, or even cool. Change of state is understood metaphorically as change of location. In this way, image schemas provide much of the embodied meaning—and correlative logic—that makes it possible to conceptualize and reason abstractly via metaphor.

Perceptual Concepts

Lawrence Barsalou (1999, 2003) argues that our perceptual symbols for various concrete objects (e.g., cars, glasses, houses) are grounded in the sensory and motor experiences afforded us by those objects. The key idea is that the same sensory, motor, and affective neural processes involved in our bodily engagement with such objects are activated when we conceptualize, reason, and talk about those objects. There are not two different and independent systems, one for perception and another for conception; instead, to conceive some object is a matter of engaging in a simulation process that activates selective sensory and motor aspects of that object and our typical physical and cultural interactions with it. For example, understanding a concept like *chair* involves a sensory, motor, and affective simulation of possible experiences with chairs of all sorts. Such simulations will involve multiple modalities (such as vision, touch, audition, and proprioception), insofar as our interactions with chairs are multimodal. We see chairs from various points of view as we walk around them, we know what it feels like to sit on and touch various types of chairs made from different materials, and we know the types of motor programs required for sitting in and standing up from

chairs. We also learn the different roles various types of chairs can play in different social and cultural situations. To know the meaning of *chair*, to understand what a chair is in a certain context, is to simulate experiences with chairs using all the sensory, motor, and affective modalities available to us.

The key idea here is that understanding a concept does not consist in accessing a list of abstract essential features or properties that define a thing. Rather, to have a concept of a particular object is to be able to simulate the kinds of perceptual, motor, and affective interactions you typically have with that kind of object. This simulation is not run in some abstract conceptual domain, but instead is enacted in the very bodily processes (employing the same functional neural clusters) involved in physically engaging that object.

Barsalou summarizes the six basic dimensions of his theory of body-based perceptual symbols as follows:

> Perceptual symbols are neural representations in sensory-motor areas of the brain; they represent schematic components of perceptual experience, not entire holistic experiences; they are multimodal, arising across the sensory modalities, proprioception, and introspection. Related perceptual symbols become integrated into a simulator that produces limitless simulations of a perceptual component (e.g., *red*, *lift*, *hungry*). Frames organize the perceptual symbols within a simulator, and words associated with simulators provide linguistic control over the construction of the simulation. (1999, 582)

Barsalou's use of the term *representation* might seem to support what is known as a representational theory of mind, in which thought proceeds via operations on internal mental representations that are somehow supposedly relatable to external, mind-independent realities. However, Barsalou's view could be made compatible with a nonrepresentational theory of mind, where having or entertaining a concept is merely running a neural simulation in which sensory, motor, and affective areas of the brain are activated not as representations mediating between an inner and outer world, but rather as *the very understanding of the concept*. In other words, the neural activations involved in the sensory, motor, and affective simulations within a specific context (including the social and cultural dimensions) *just are* what it is to grasp the meaning of the concept in question.