MICHAEL TOMASELLO

A NATURAL HISTORY OF HUMAN THINKING

A Natural History of Human Thinking

Michael Tomasello

Harvard University Press

Cambridge, Massachusetts
London, England
2014

Printed in the United States of America

Library of Congress Cataloging-in-Publication Data

Tomasello, Michael.
A natural history of human thinking / Michael Tomasello.
pages cm
Includes bibliographical references and index.
ISBN 978-0-674-72477-8 (hardcover : alk. paper)
1. Cognition—Social aspects. 2. Evolutionary psychology.
3. Psychology, Comparative. I. Title.
BF311.T6473 2014
153—dc23 2013020185

Contents

Preface

This book is a sequel—or, better, a prequel—to *The Cultural Origins of Human Cognition* (Harvard University Press, 1999). But it also has a slightly different focus. In the 1999 book the question was what makes human cognition unique, and the answer was culture. Individual human beings develop uniquely powerful cognitive skills because they grow to maturity in the midst of all kinds of cultural artifacts and practices, including a conventional language, and of course they have the cultural learning skills necessary to master them. Individuals internalize the artifacts and practices they encounter, and these then serve to mediate all of their cognitive interactions with the world.

In the current book, the question is similar: what makes human thinking unique? And the answer is similar as well: human thinking is fundamentally cooperative. But this slightly different question and slightly different answer lead to a very different book. The 1999 book was clean and simple because the data we had comparing apes and humans were so sparse. We could thus say things like "Only humans understand others as intentional agents, and this enables human culture." But we now know that the picture is more complex than this. Great apes appear to know much more about others as intentional agents than previously believed, and still they do not have human-like culture or cognition. Based on much research reported here, the critical difference now seems to be that humans not only understand others as intentional agents but also put their heads together with others in acts of shared intentionality, including everything from concrete acts of collaborative problem solving to complex cultural institutions. The focus now is thus less on culture as a process of transmission and more on culture as a process of social coordination—and indeed, we argue here that modern human cultures were made possible by an earlier evolutionary step in which individuals made a living by coordinating with others in relatively simple acts of collaborative foraging.

The specific focus on thinking means that this book does not simply document that humans participate in shared intentionality in a way that their nearest primate relatives do not, which has been done elsewhere. Rather, in addition, it examines the underlying thinking processes involved. To describe the nature of these thinking processes—in particular, to distinguish human thinking from that of other apes—we must characterize its component processes of cognitive representation, inference, and self-monitoring. The *shared intentionality hypothesis* claims that all three of these components were transformed in two key steps during human evolution. In both cases, the transformation was part of a larger change of social interaction and organization in which humans were forced to adopt more cooperative lifeways. In order to survive and thrive, humans were forced, twice, to find new ways to coordinate their behavior with others in collaborative (and then cultural) activities and to coordinate their intentional states with others in cooperative (and then conventional) communication. And this transformed, twice, the way that humans think.

The writing of this book, as most others, was made possible by the support of many institutions and people. I would like to thank the University of Pittsburgh Center for Philosophy of Science (John Norton, director and seminar leader extraordinaire) for hosting me for one peaceful semester of concentrated writing in the spring of 2012. I especially benefited during this stay from Bob Brandom's generosity with his time and thoughts on many topics central to the current enterprise. I thank Celia Brownell at the Pitt Department of Psychology and Andy Norman at Carnegie Mellon for many useful discussions during this semester as well. The ensuing summer I benefited greatly from presenting the themes of the book to the SIAS Summer Institute titled *The Second Person: Comparative Perspectives*, organized in Berlin by Jim Conant and Sebastian Rödl. The book is better for all of these encounters.

With regard to the manuscript itself, I would like to thank Larry Barsalou, Mattia Galloti, Henrike Moll, and Marco Schmidt for reading various chapters and providing very useful feedback. Of special importance, Richard Moore and Hannes Rakoczy each read the entire manuscript at a fairly early stage and provided a number of trenchant comments and suggestions, regarding both content and presentation. Thanks also to Elizabeth Knoll and three anonymous reviewers at Harvard University Press for a number of helpful comments and criticisms on the penultimate draft.

Last and most important, I thank my wife, Rita Svetlova, for providing constant and detailed critical commentary and suggestions throughout. Many ideas were made clearer through discussion with her, and confusing passages were made clear, or at least clearer, by her literate eye.

1

The Shared Intentionality Hypothesis

Only cooperation constitutes a process that can produce reason.

—JEAN PIAGET, *SOCIOLOGICAL STUDIES*

Thinking would seem to be a completely solitary activity. And so it is for other animal species. But for humans, thinking is like a jazz musician improvising a novel riff in the privacy of his own room. It is a solitary activity all right, but on an instrument made by others for that general purpose, after years of playing with and learning from other practitioners, in a musical genre with a rich history of legendary riffs, for an imagined audience of jazz aficionados. Human thinking is individual improvisation enmeshed in a sociocultural matrix.

How did this novel form of socially infused thinking come to be, and how does it work? One set of classic theorists has emphasized the role of culture and its artifacts in making possible certain types of individual thinking. For example, Hegel (1807) argued that the social practices, institutions, and ideologies of a particular culture at a particular historical epoch constitute a necessary conceptual framework for individual human reason (see also Collingwood, 1946). Peirce (1931–1935) claimed more specifically that virtually all of humans' most sophisticated types of thinking, including most especially mathematics and formal logic, are possible only because individuals have available to them culturally created symbolic artifacts such as Arabic numerals and logical notation. Vygotsky (1978) emphasized that human children grow up in the midst of the tools and symbols of their culture, including especially the linguistic symbols that preorganize their worlds for them, and during ontogeny they internalize the use of these artifacts, leading to the kind of internal dialogue that is one prototype of human thinking (see also Bakhtin, 1981).

The other set of classic theorists has focused on the fundamental processes of social coordination that make human culture and language possible in the first place. Mead (1934) pointed out that when humans interact with one another, especially in communication, they are able to imagine themselves in the role of the other and to take the other's perspective on themselves. Piaget (1928) argued further that these role-taking and perspective-taking abilities—along with a cooperative attitude—not only make culture and language possible but also make possible reasoning in which individuals subordinate their own point of view to the normative standards of the group. And Wittgenstein (1955) explicated several different ways in which the appropriate use of a linguistic convention or cultural rule depends on a preexisting set of shared social practices and judgments ("forms of life"), which constitute the pragmatic infrastructure from which all uses of language and rules gain their interpersonal significance. These social infrastructure theorists, as we may call them, all share the belief that language and culture are only the "icing on the cake" of humans' ultrasocial ways of relating to the world cognitively.

Insightful as they were, all of these classic theorists were operating without several new pieces of the puzzle, both empirical and theoretical, that have emerged only in recent years. Empirically, one new finding is the surprisingly sophisticated cognitive abilities of nonhuman primates, which have been discovered mostly in the last few decades (for reviews, see Tomasello and Call, 1997; Call and Tomasello, 2008). Thus, great apes, as the closest living relatives of humans, already understand in human-like ways many aspects of their physical and social worlds, including the causal and intentional relations that structure those worlds. This means that many important aspects of human thinking derive not from humans' unique forms of sociality, culture, and language but, rather, from something like the individual problem-solving abilities of great apes in general.

Another new set of findings concern prelinguistic (or just linguistic) human infants, who have yet to partake fully of the culture and language around them. These still fledgling human beings nevertheless operate with some cognitive processes that great apes do not, enabling them to engage with others socially in some ways that great apes cannot, for example, via joint attention and cooperative communication (Tomasello et al., 2005). The fact that these precultural and prelinguistic creatures are already cognitively unique provides empirical support for the social infrastructure theorists' claim that important aspects of human thinking emanate not from culture and language per se but, rather, from some deeper and more primitive forms of uniquely human social engagement.

Theoretically, recent advances in the philosophy of action have provided powerful new ways of thinking about these deeper and more primitive forms of uniquely human social engagement. A small group of philosophers of action (e.g., Bratman, 1992; Searle, 1995; Gilbert, 1989; Tuomela, 2007) have investigated how humans put their heads together with others in acts of so-called shared intentionality, or "we" intentionality. When individuals participate with others in collaborative activities, together they form joint goals and joint attention, which then create individual roles and individual perspectives that must be coordinated within them (Moll and Tomasello, 2007). Moreover, there is a deep continuity between such concrete manifestations of joint action and attention and more abstract cultural practices and products such as cultural institutions, which are structured—indeed, created—by agreed-upon social conventions and norms (Tomasello, 2009). In general, humans are able to coordinate with others, in a way that other primates seemingly are not, to form a "we" that acts as a kind of plural agent to create everything from a collaborative hunting party to a cultural institution.

Further in this theoretical direction, as a specific form of human collaborative activity and shared intentionality, human cooperative communication involves a set of special intentional and inferential processes—first identified by Grice (1957, 1975) and since elaborated and amended by Sperber and Wilson (1996), Clark (1996), Levinson (2000), and Tomasello (2008). Human communicators conceptualize situations and entities via external communicative vehicles *for* other persons; these other persons then attempt to determine why the communicator thinks that these situations and entities will be relevant for them. This dialogic process involves not only skills and motivations for shared intentionality but also a number of complex and recursive inferences about others' intentions toward my intentional states. This unique form of communication—characteristic not just of mature language use but also of the prelinguistic gestural communication of human infants—presupposes both a shared conceptual framework between communicative partners (a.k.a. common conceptual ground) and an appreciation of those partners' individual intentions and perspectives within it.

These new empirical and theoretical advances enable us to construct a much more detailed account than was previously possible of the social dimensions of human cognition in general. Our focus in this book is on the social dimensions of human thinking in particular. Although humans and other animals solve many problems and make many decisions based on evolved intuitive heuristics (so-called system 1 processes), humans and at least some other animals also solve some problems and make some decisions by thinking (system 2 processes; e.g., Kahneman, 2011). A specific focus on thinking is useful because it restricts our topic to a single cognitive process, but one that involves several key components, especially (1) the ability to cognitively represent experiences to oneself "off-line"; (2) the ability to simulate or make inferences transforming these representations causally, intentionally, and/or logically; and (3) the ability to self-monitor and evaluate how these simulated experiences might lead to specific behavioral outcomes—and so to make a thoughtful behavioral decision.

It seems obvious that, compared with other animal species, humans think in special ways. But this difference is hard to characterize using traditional theories of human thinking since they presuppose key aspects of the process that are actually evolutionary achievements. These are precisely the social aspects of human thinking that are our primary focus here. Thus, although many animal species can cognitively represent situations and entities at least somewhat abstractly, only humans can conceptualize one and the same situation or entity under differing, even conflicting, social perspectives (leading ultimately to a sense of "objectivity"). Further, although many animals also make simple causal and intentional inferences about external events, only humans make socially recursive and self-reflective inferences about others' or their own intentional states. And, finally, although many animals monitor and evaluate their own actions with respect to instrumental success, only humans self-monitor and evaluate their own thinking with respect to the normative perspectives and standards ("reasons") of others or the group. These fundamentally social differences lead to an identifiably different type of thinking, what we may call, for the sake of brevity, *objective-reflective-normative thinking*.

In this book we attempt to reconstruct the evolutionary origins of this uniquely human objective-reflective-normative thinking. The *shared intentionality hypothesis* is that what created this unique type of thinking—its processes of representation, inference, and self-monitoring—were adaptations for dealing with problems of social coordination, specifically, problems presented by individuals' attempts to collaborate and communicate with others (to *co*-operate with others). Although humans' great ape ancestors were social beings, they lived mostly individualistic and competitive lives, and so their thinking was geared toward achieving individual goals. But early humans were at some point forced by ecological circumstances into more cooperative lifeways, and so their thinking became more directed toward figuring out ways to coordinate with others to achieve joint goals or even collective group goals. And this changed everything.

There were two key evolutionary steps. The first step, reflecting the focus of social infrastructure theorists such as Mead and Wittgenstein, involved the creation of a novel type of small-scale collaboration in human foraging. Participants in this collaborative foraging created socially shared joint goals and joint attention (common ground), which created the possibility of individual roles and perspectives within that ad hoc shared world or "form of life." To coordinate these newly created roles and perspectives, individuals evolved a new type of cooperative communication based on the natural gestures of pointing and pantomiming: one partner directed the attention or imagination of the other perspectivally and/or symbolically about something "relevant" to their joint activity, and then that partner made cooperative (recursive) inferences about what was intended. To self-monitor this process the communicator had to simulate ahead of time the recipient's likely inferences. Because the collaboration and communication at this point were between ad hoc pairs of individuals in the moment—based on purely second-personal social engagement between "I" and "you"—we may refer to all of this as *joint intentionality*. When put to use in thinking, joint intentionality comprises perspectival and symbolic representations, socially recursive inferences, and second-personal self-monitoring.

The second step, reflecting the focus of culture theorists such as Vygotsky and Bakhtin, came as human populations began growing in size and competing with one another. This competition meant that group life as a whole became one big collaborative activity, creating a much larger and more permanent shared world, that is to say, a culture. The resulting group-mindedness among all members of the cultural group (including in-group strangers) was based on a new ability to construct common *cultural* ground via collectively known cultural conventions, norms, and institutions. As part of this process, cooperative communication became conventionalized linguistic communication. In the context of cooperative argumentation in group decision making, linguistic conventions could be used to justify and make explicit one's reasons for an assertion within the framework of the group's norms of rationality. This meant that individuals now could reason "objectively" from the group's agent-neutral point of view ("from nowhere"). Because the collaboration organized circularly, as feedback control systems, with built-in goal states and action possibilities. Starting from this foundation, cognition evolves not from a complexifying of stimulus-response linkages but, rather, from the individual organism gaining (1) powers of flexible decision-making and behavioral control in its various adaptive specializations, and (2) capacities for cognitively representing and making inferences from the casual and intentional relations structuring relevant events.

Adaptive specializations are organized as self-regulating systems, as are many physiological processes such as the homeostatic regulation of blood sugar and body temperature in mammals. These specializations go beyond reflexes in their capacity to produce adaptive behavior in a much wider range of circumstances, and indeed, they may be quite complex, for example, spiders spinning webs. There is no way that a spider can spin a web using only stimulus-response linkages. The process is too dynamic and dependent on local context. Instead, the spider must have goal states that it is motivated to bring about, and the ability to perceive and act so as to bring them about in a self-regulated manner. But adaptive specializations are still not cognitive (or only weakly cognitive) because they are unknowing and inflexible by definition: perceived situations and behavioral possibilities for goal attainment are mostly connected in an inflexible manner. The individual organism does not have the kind of causal or intentional understanding of the situation that would enable it to deal flexibly with "novel" situations. Natural selection has designed these adaptive specializations to work invariantly in "the same" situations as those encountered in the past, and so cleverness from the individual is not needed.

Cognition and thinking enter the picture when organisms live in less predictable worlds and natural selection crafts cognitive and decision making processes that empower the individual to recognize novel situations and to deal flexibly, on its own, with unpredictable exigencies. What enables effective handling of a novel situation is some understanding of the causal and/or intentional relations involved, which then suggests an appropriate and potentially novel behavioral response. For example, a chimpanzee might recognize that the only tool available to her in a given situation demands, based on the physical causality involved, manipulations she has never before performed toward this goal. A cognitively competent organism, then, operates as a control system with reference values or goals, capacities for attending to situations causally or intentionally "relevant" to these reference values or goals, and capacities for choosing actions that lead to the fulfillment of these reference values or goals (given the causal and/or intentional structure of the situation).

This description in control system terms is basically identical to the classic belief-desire model of rational action in philosophy: a goal or desire coupled with an epistemic connection to the world (e.g., a belief based on an understanding of the causal or intentional structure of the situation) creates an intention to act in a particular way.[1]

We will refer to this flexible, individually self-regulated, cognitive way of doing things as *individual intentionality*. Within this self-regulation model of individual intentionality, we may then say that thinking occurs when an organism attempts, on some particular occasion, to solve a problem, and so to meet its goal not by behaving overtly but, rather, by imagining what would happen if it tried different actions in a situation—or if different external forces entered the situation—before actually acting. This imagining is nothing more or less than the "off-line" simulation of potential perceptual experiences. To be able to think before acting in this way, then, the organism must possess the three prerequisites outlined above: (1) the ability to cognitively represent experiences to oneself "off-line," (2) the ability to simulate or make inferences transforming these representations causally, intentionally, and/or logically, and (3) the ability to self-monitor and evaluate how these simulated experiences might lead to specific behavioral outcomes—and so to make a thoughtful behavioral decision. The success or failure of a particular behavioral decision exposes the underlying processes of representation, simulation, and self-monitoring—indirectly, as it were—to the unrelenting sieve of natural selection.

Cognitive Representation

Cognitive representation in a self-regulating, intentional system may be characterized both in terms of its content and in terms of its format. In terms of content, the claim here is that both the organism's internal goals and its externally directed attention (NB: not just perception but attention) have as content not punctate stimuli or sense data, but rather whole *situations*. Goals, values, and other reference values (pro-attitudes) are cognitive representations of situations that the organism is motivated to bring about or maintain. Although we sometimes speak of an object or location as someone's goal, this is really only a shorthand way of speaking; the goal is the situation of *having* the object or *reaching* the location. The philosopher Davidson (2001) writes, "Wants and desires are directed to propositional contents. What one wants is . . . *that* one has the apple in hand. . . . Similarly . . . someone who intends to go to the opera intends to make it the case *that* he is at the opera" (p. 126).

In this same manner, modern decision theory often speaks of the desire or preference *that* a particular state of affairs be realized.

If goals and values are represented as desired situations, then what the organism must attend to in its perceived environment is situations relevant to those goals and values. Desired situations and attended-to environmental situations are thus perforce in the same perceptually based, fact-like representational format, which enables their cognitive comparison. Of course, complex organisms also perceive less complex things, such as objects, properties, and events—and can attend to them for specific purposes—but in the current analysis they always do so as components of situations relevant to behavioral decision making.

To illustrate the point, let us suppose that the image in Figure 2.1 is what a chimpanzee sees as she approaches a tree while foraging.

FIGURE 2.1 What a chimpanzee sees

The chimpanzee perceives the scene in the same basic way that we would; our visual systems are similar enough that we see the same basic objects and their spatial relationships. But what situations does the chimpanzee attend to? Although she could potentially focus her attention on any of the potentially infinite situations that this image presents, at the current moment she must make a foraging decision, and so she attends to the situations or "facts" relevant to this behavioral decision, to wit (as described in English):

- that many bananas are in the tree
- that the bananas are ripe
- that no competitor chimpanzees are already in the tree
- that the bananas are reachable by climbing
- that no predators are nearby
- that escaping quickly from this tree will be difficult
- etc., etc.

For a foraging chimpanzee with the goal of obtaining food, given all of its perceptual and behavioral capacities and its knowledge of the local ecology, all of these are *relevant situations* for deciding what to do—all present in a single visual image and, of course, nonverbally. (NB: Even the absence of something expected, such as food not in its usual location, may be a relevant situation.)

Relevance is one of those occasion-sensitive judgments that cannot be given a general definition. But in broad strokes, organisms attend to situations as either (1) opportunities or (2) obstacles to the pursuit and maintenance of their goals and values (or as information relevant to predicting possible future opportunities or obstacles). Different species have different ways of life, of course, which means that they perceive or attend to different situations (and components of situations). Thus, for a leopard, the situation of bananas in a tree would not represent an opportunity to eat, but the presence of a chimpanzee would. For the chimpanzee, in contrast, the leopard's presence now presents an obstacle to its value of avoiding predators, and so it should look for a situation providing opportunities for escape, such as a tree to climb without low-hanging limbs—given its knowledge that leopards cannot climb such trees and its familiarity with its own tree-climbing prowess. If we now throw into the mix a worm resting on the banana's surface, the relevant situations for the three different species—the obstacles and opportunities for their respective goals—would overlap even less, if at all. Relevant situations are thus determined jointly by the organism's goals and values, its perceptual abilities and knowledge, and its behavioral capacities, that is to say, by its overall functioning as a self-regulating system. Identifying situations relevant for a behavioral decision thus involves an organism's whole way of life (von Uexküll, 1921).[2]

In terms of representational format, the key is that to make creative inferences that go beyond particular experiences, the organism must represent its experiences as types, that is to say, in some generalized, schematized, or abstract form. One plausible hypothesis is a kind of exemplar model in which the individual in some sense "saves" the particular situations and components to which it has attended (many models of knowledge representation have attention as the gateway). There is then generalization or abstraction across these in a process that we might call *schematization*. (Langacker's [1987] metaphor is of a stack of transparencies, each depicting a single situation or entity, and schematization is the process of looking down through them for overlap.) We might think of the result of this process of schematization as cognitive models of various types of situations and entities, for example, categories of objects, schemas of events, and models of situations. Recognizing a situation or entity as a token of a known type—as an exemplar of a cognitive category, schema, or model—enables novel inferences about the token appropriate to the type.

Categories, schemas, and models as cognitive types are nothing more or less than imagistic or iconic schematizations of the organism's (or, in some cases, its species') previous experience (Barsalou, 1999, 2008). As such, they do not suffer from the indeterminacy of interpretation that some theorists attribute to iconic representations considered as mental pictures, that is, the indeterminacy of whether this image is of a banana, a fruit, an object, and so forth (Crane, 2003). They do not because they are composed of individual experiences in which the organism was attending to a relevant (already "interpreted") situation. Thus, the organism "interprets," or understands, particular situations and entities in the context of its goals as it assimilates them to known (cognitively represented) types: "This is another one of those."

Simulation and Inference

Thinking in an organism with individual intentionality involves simulations or inferences that connect cognitive representations of situations and their components in various ways. First are those instrumental inferences that occur

Thinking like an Ape

We begin our natural history of the evolutionary emergence of uniquely human thinking with a focus on the last common ancestor of humans and other extant primates. Our best living models for this creature are humans' closest primate relatives, the nonhuman great apes (hereafter, great apes), comprising chimpanzees, bonobos, gorillas, and orangutans—especially chimpanzees and bonobos, who diverged from humans most recently, around 6 million years ago. When cognitive abilities are similar among the four species of great ape but different in humans, we presume that the apes have conserved their skills from the last common ancestor (or before) whereas humans have evolved something new.

Our characterizations of the cognitive skills of this last common ancestor will derive from empirical research with great apes, cast in the theoretical framework of individual intentionality just elaborated: behavioral self-regulation involving cognitive models and instrumental inferences, along with some form of behavioral self-monitoring. Because humans share with other apes such a recent evolutionary history—along with the same basic bodies, sense organs, emotions, and brain organization—in the absence of evidence our default assumption will be evolutionary continuity (de Waal, 1999). That is to say, when great apes behave identically with humans, especially in carefully controlled experiments, we will assume continuity in the underlying cognitive processes involved. The onus of explanation is thus on those who posit evolutionary discontinuities, a challenge we embrace in later chapters.

Great Apes Think about the Physical World

Processes of great ape cognition and thinking may be usefully divided into those concerning the physical world, structured by an understanding of physical causality, and those concerning the social world, structured by an understanding of agentive causality, or intentionality. Primate cognition of the physical world evolved mainly in the context of foraging for food (see Tomasello and Call, 1997, for this theoretical claim and supporting evidence); this is thus its "proper function" (in Millikan's [1987] sense). In order to procure their daily sustenance, primates (as mammals in general) evolved the proximate goals, representations, and inferences for (1) finding food (requiring skills of spatial navigation and object tracking), (2) recognizing and categorizing food (requiring skills of feature recognition and categorization), (3) quantifying food (requiring skills of quantification), and (4) procuring or extracting food (requiring skills of causal understanding). In these most basic skills of physical cognition, all nonhuman primates would seem to be generally similar (Tomasello and Call, 1997; Schmitt et al., 2012).

What great apes are especially skillful at, compared with other primates, is tool use—which one might characterize as not just understanding causes but actually manipulating them. Other primates are mostly not skilled tool users at all, and when they are it is typically in only one fairly narrow context (e.g., Fragaszy et al., 2004). In contrast, all four species of great ape are highly skilled at using a variety of tools quite flexibly, including using two tools in succession in a task, using one tool to rake in another (which is then needed to procure food), and so forth (Herrmann et al., 2008). Classically, tool use is thought to require the individual to assess the causal effect of its tool manipulations on the goal object or event (Piaget, 1952), and so the flexibility and alacrity with which great apes succeed in using novel tools suggest that they have one or more general cognitive models of causality guiding their use of these novel tools.

Great apes' skills with manipulating causal relations via tools may be combined in interesting ways with processes of cognitive representation and inference. For example, Marín Manrique et al. (2010) presented chimpanzees with a food extraction problem that they had never before seen. Its solution required a tool with particular properties (e.g., it had to be rigid and of a certain length). The trick was that the potential tools they could use were in a different room, out of sight of the problem. To solve this task, individuals had to first comprehend the causal structure of the novel problem, and then keep that structure cognitively represented while approaching and choosing a tool in the other room. Many individuals did this, often from the first trial onward, suggesting that they assimilated the novel problem to a known cognitive model having a certain causal structure, which they then kept with them as they entered the adjoining room. They then simulated the use of at least some of the available tools and the likely outcome in each case through the medium of this cognitive model—before actually choosing a tool overtly. In the study of Mulcahy and Call (2006), bonobos even saved a tool for future use, presumably imagining the future situation in which they would need it.

The simulations or inferences involved here have logical structure. This is not the structure of formal logic but, rather, a structure based on causal inferences. The idea is that causal inferences have a basic if-then logic and so lead to "necessary" conclusions: if A happens, then B happens (because A caused B). Bermudez (2003) calls inferences of this type protoconditional because the necessity is not a formal one but a causal one. In the experiment of Marín Manrique et al. (2010), as an ape simulates using the different tools, she infers "if a tool with property A is used, then B must happen." One thus gets a kind of proto-*modus ponens* by then actually using the tool with property A in the expectation that B will indeed happen as a causal result (if A happens, then B happens; A happens; therefore B will happen). This is basically a forward-facing inference, from premise or cause to conclusion or effect.

In another set of recent experiments, we can see backward-facing inferences, that is, from effect to cause. Call (2004) showed chimpanzees a piece of food, which was then hidden in one of two cups (they did not know which). Then, depending on condition, the experimenter shook one of the cups. The relevant background knowledge for success in this experiment is as follows: (1) the food is in one of the two cups (learned in pretraining), and (2) shaking the cup with food will result in noise, whereas shaking the cup without food will result in silence (causal knowledge brought to the experiment). The two conditions are shown in Figure 2.2, using iconic representations to depict something of the way the apes understand the situation.

(The iconic diagrams modeling great ape cognitive representations in Figure 2.2 are not uninterpreted pictures but symbols in a theoretical metalanguage that mean what we agree that they mean. So they are meant to depict the ape's interpreted experience when she has seen the cup as a cup and the noise as coming from the cup, and so on. Importantly, these diagrams are created within the confines of a restrictive theory of the possibilities of great ape cognition. Following Tomasello's (1992) depictions for one-year-old human children, we make the diagrams out of concrete spatial-temporal-causal elements that may be posited to be a part of the apes' cognitive abilities based on empirical research. Then, the logical structure—based on the protoconditional and protonegation—is posited to be necessary to explain apes' actions in specific experimental situations. The logical operations are depicted in English words, since the ape does not have perception-based representations of them, but only procedural competence with them.)

In condition 1, an experimenter shook the cup with food. In this case the chimpanzee observed a noise being made and had to infer backward in the causal chain to what might have caused it, specifically, the food hitting the inside of the cup. This is a kind of abduction (not logically valid, but an "inference to best explanation"). That is, (1) the shaking cup is making noise;

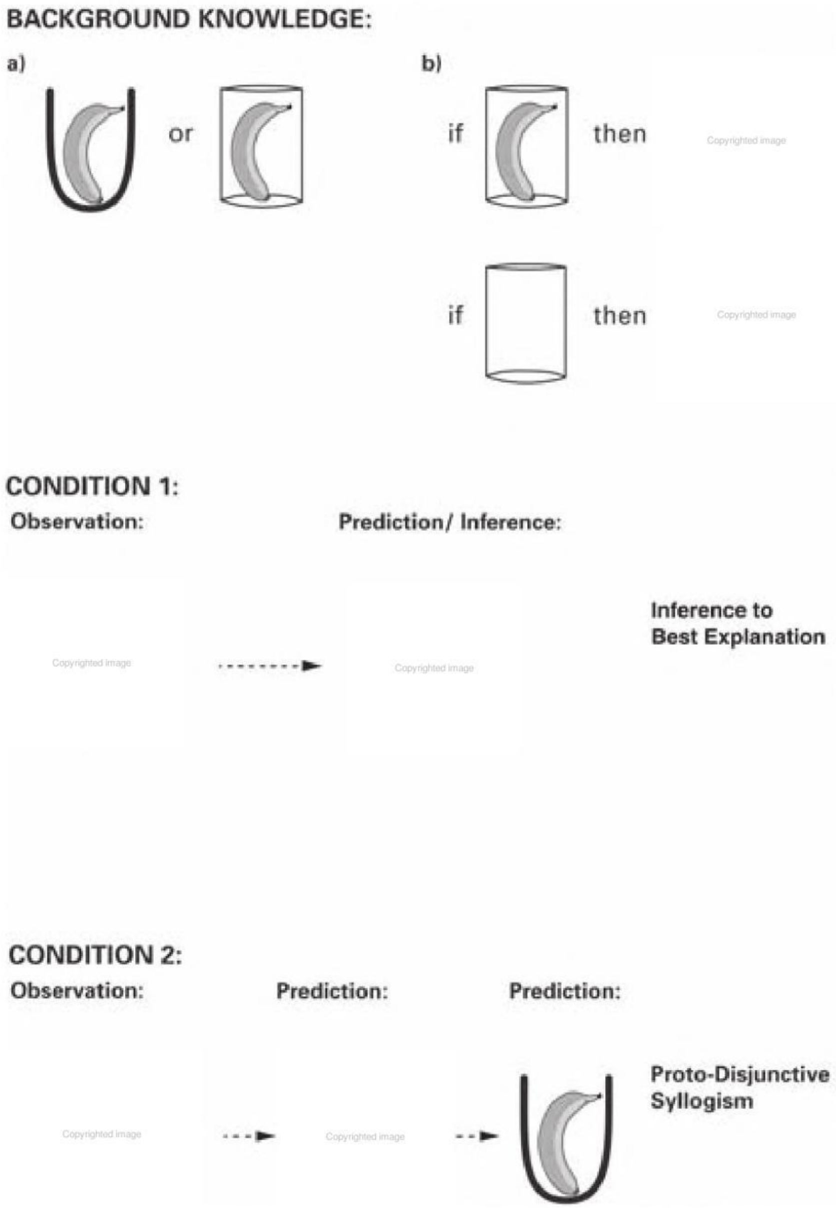

FIGURE 2.2 Ape inferences in finding hidden food (Call, 2004)

(2) if the food were inside the shaking cup, then it would make noise; (3) therefore, the food is inside the cup. In condition 2, the experimenter shook the empty cup. In this case the chimpanzee observed only silence and had to infer backward in the causal chain to why that might be, specifically, that there was no food in the cup. This is a kind of proto-*modus tollens*: (1) the shaking cup is silent; (2) if the food were inside the shaking cup, then it would make noise; (3) therefore, the food must not be in the cup (the shaken cup must be empty). The chimpanzees made this inference, but they also made an additional one. They combined their understanding of the causality of noise making in this context with their preexisting knowledge that the food was in one of the two cups to locate the food in the *other*, nonshaken cup (if the food is not in this one, then it must be in that one; see bottom row in Figure 2.2). This inferential paradigm thus involves the kind of exclusion inference characteristic of a disjunctive syllogism.

Negation is a very complex cognitive operation, and one could easily object to the use of negation in these proposed accounts of great ape logical inferences. But Bermudez (2003) makes a novel theoretical proposal about some possible evolutionary precursors to formal negation that make these accounts much more plausible. The proposal is to think of a kind of protonegation as simply comprising exclusionary opposites on a scale (contraries), such as presence-absence, noise-silence, safety-danger, success-failure, and available-not available. If we assume that great apes understand polar opposites such as these as indeed mutually exclusive—for example, if something is absent, it cannot be present, or if it makes noise it cannot be silent—then this could be a much simpler basis for the negation operation. All of the current descriptions assume protonegation of this type.

When taken together, the conditional (if-then) and negation operations structure all of the most basic paradigms of human logical reasoning. The claim is thus that great apes can solve complex and novel physical problems by assimilating key aspects of the problem situation to already known cognitive models with causal structure and then use those models to simulate or make inferences about what has happened previously or what might happen next—employing both a kind of protoconditional and a kind of protonegation in both forward-facing and backward-facing paradigms. Our general conclusion is thus that since the great apes in these studies are using cognitive models containing general principles of causality, and they are also simulating or making inferences in various kinds of protological paradigms, with various kinds of