2

Three Puzzles of Mindreading

BERTRAM F. MALLE

Mindreading is the human activity of inferring other people's mental states. Literatures from multiple disciplines have advanced our knowledge on this phenomenon, and we now know quite a bit about the development and functional use of mindreading in the human species (e.g., Astington, 1993; Baron-Cohen, Tager-Flusberg, & Cohen, 2000; Malle, Moses, & Baldwin, 2001; Perner, 1991; Wellman, 1990) and are even beginning to sketch a picture of its evolutionary origins (Baron-Cohen, 1999; Bogdan, 2000; Malle, 2002; Povinelli, 2001; Whiten, 1999). But as we try to integrate this growing knowledge and perhaps work toward a unified theory of mental state inference, many problems and puzzles emerge. In this chapter I focus on three such puzzles. Each of them, I will argue, has a credible solution, but what is perhaps more important is that each of these solutions points to the same general conclusion about mindreading.

THE FIRST PUZZLE: BEHAVIOR AS INPUTS AND OUTPUTS OF MINDREADING

The first puzzle concerns the relationship of behavior to inferences of mental states. On the one hand, most researchers assume (or explicitly theorize) that people typically infer mental states by observing the actor's behavior, making perceptions of behavior an important input to

mindreading. Watching facial expressions, head and eye direction, body posture, and movements allows the perceiver to figure out a great deal about what others believe, want, feel, or intend. On the other hand, perceptions of behavior can also be an output of mindreading. Plenty of research shows that people explain behavior to a considerable extent by ascribing mental states to the agent (e.g., Buss, 1978; Heider, 1958; Malle, 1999; McClure, 2002; Read, 1987; Wellman, Hickling, & Schult, 1997). In the case of intentional behavior, it is primarily beliefs and desires that are seen as reasons that make the agent's action intelligible. But even in the case of unintentional behavior, mental states such as emotions, perceptions, and thoughts serve to explain why the person behaved in the observed way. Thus, mental state inferences are inputs to explaining behavior.

But if this is all true, we are caught in circularity. How can the human perceiver use an agent's observed behavior to infer his or her mental states but gather the meaning of that very behavior from inferences of mental states? We know there has got to be a solution to this puzzle; but how do people get it done?

There would seem to be two paths out of this circularity. One path requires at least some behaviors whose meaning can be assessed without reference to mental states. Those behaviors could then be noncircular inputs to certain mental state inferences. The other path requires that at least some mental states can be grasped without behavior observation, and those mental states could then be noncircular inputs to the interpretation of certain behaviors.

Along the first path, I will first discuss two less promising approaches, followed by two more promising ones.

Raw Behaviors

One possibility is that perceivers use "raw" behavior observations as input to mindreading—behaviors that don't require any further interpretation. But raw behavior observations are hard to come by. Purely physical descriptions of an agent's body (without making any reference to meaning) are difficult to produce unless one is an exercise and movement scientist. Neither young children (past the age of 4, when they make a variety of mental state inferences) nor untrained adults have such technical knowledge. Moreover, minute variations in physical behavior would demand distinct physical descriptions. These descriptions would somehow have to be translated into *types* of indicator behavior, because only then would the perceiver be able to narrow down the possible mental states to infer. Without at least a rudimentary meaning analysis, however, such type identification would seem extremely difficult.

In fact, a purely physical analysis of behavior resembles what we see in autistic individuals—who find both social behavior and mental states incomprehensible. Autistic perceivers apparently notice "raw" behaviors but do so without recognizing their meaning. One autistic person reports:

> I know people's faces down to the acne scars on the left corners of their chins and what their eyes do when they speak, and how the hairs of their eyebrows curl, and how their hairlines curve around the tops of their foreheads. . . . The best I can do is start picking up bits of data during my encounter with them because there's not much else I can do. It's pretty tiring, though, and explains something of why social situations are so draining for me. . . . That said, I'm not sure what kind of information about them I'm attempting to process. (Blackburn, Gottschewski, George, & L—, 2000)

Some other information sources have to supplement the raw perception of behavior for those behaviors to appear meaningful. Perhaps perceivers can use context information to make a type identification. Tears flowing down the agent's face during a funeral indicate a different behavior than tears flowing down the agent's face after listening to a joke. That would distinguish crying from laughing, and once these types are distinguished, they could serve as inputs to mental state inferences of sadness and joy, respectively. But what exactly is "context"? It isn't just the physical arrangement of bodies and props; it is the interpretation of those physical elements as *standing for* something, *counting as* a funeral. It would be difficult to classify most contexts without registering what the participants take the situation to be—requiring inferences about their beliefs, assumptions, and interpretations (Givón, 2005), and throwing us back into circularity.

Intentionality Concept

A more promising approach is to consider the concept of intentionality— a conceptual frame within which behavior is interpreted but, at least among infants, without the inference of mental states. Infants learn early to distinguish intentional from unintentional behavior, with estimates ranging from 9 months to 18 months (Carpenter, Akhtar, & Tomasello, 1998; Gergely, Nádasdy, Csibra, & Bíró, 1995; Woodward, 1999). At this tender age, infants must respond to certain cues: the degree of movement "smoothness" as a symptom of the agent's control (compare walking down stairs versus falling down stairs); characteristic accompanying behaviors (e.g., head turning, eye gaze); the connection and manipula-

tion of objects in the world; equifinality (the principle that intentional agents pursue their goals along multiple paths, trying a different path when the first one failed); and characteristic agent responses at the end of the behavior (e.g., "there!" or a happy face with an intentional action; "oops" or an unhappy face with an unintentional action). As far as we know, intentionality is not "mentalized" at this early age; rather, the concept serves as a means of distinguishing observable behaviors into two classes, with room for further differentiation within each class.

With increasing age (probably between 2 and 4 years), children learn to judge intentionality with more sophistication and thus begin to appreciate the involvement of mental states in action—states such as desires, beliefs, and intentions. More specifically, children slowly learn the adult concept of intentionality, which incorporates four mental states. An action is considered intentional only when the agent has desire for an outcome, beliefs about the action that leads to the outcome, an intention to perform the action, and awareness of fulfilling this intention while acting (Malle & Knobe, 1997a). This frame allows the powerful inference that whenever an action is intentional (presumably judged, like the infant does, on the basis of cues other than mental states), there must be the involvement of various mental states, most notably beliefs and desires that provide the reasons for acting (Malle, 1999). Obviously, the conceptual frame does not provide context-specific beliefs and desires— the perceiver knows only that *some* desire and *some* beliefs were involved, not which ones (Malle et al., 2001).

Intentionality thus provides an interpretational frame that even toddlers can acquire before they infer any mental states. Once this frame is in place, expectations for and responses to intentional behavior can be better coordinated, and the significant elements of a social interaction appear in relief. Moreover, with time, the intentionality concept facilitates the inference of mental states in a noncircular way, using the "trick" of postulating certain *kinds* of mental state upon encountering intentional behavior—a postulate that holds without any further analysis of the behavior's specific meaning.

Transparent Behaviors

Another promising approach along the first path is the identification of behaviors that may be sufficiently transparent so that the perceiver who uncovers their meaning gets by without any mental state inferences. The first such class is expressive behaviors, such as screaming in pain, laughing with joy, or growling with anger. Their meaning could initially be purely functional in that they are associated with certain antecedents (e.g., in the case of screaming in pain, sharp or heavy objects intruding

on the body) or certain consequences (e.g., in the case of growling with anger, destructive movements). These antecedents and consequences are part of the *physical* context and therefore require no mental state interpretation. The second class of transparent behavior comprises basic human movements such as reaching, grasping, walking, standing up, and lying down. The meaning of such actions can initially be purely functional as well—defined by the role they play in interactions with objects or other beings. For example, reaching and grasping connect the agent with objects and make those objects manipulable and consumable.

In addition to their transparency, expressive behaviors and basic movements have another important feature: They are performed early by human infants themselves. This feature has two associated characteristics. First, transparent behaviors are also the kinds of behaviors that have a high probability of being imitated or becoming "contagious" (see next section), and at least some of them have been linked to brain structures that translate perceptions of another person's behavior into motor programs of performing that same behavior (Decety, Chapter 9, this volume; Jeannerod, 1994; Rizzolatti, Fadiga, Gallese, & Fogassi, 1996). Second, behaviors that the infant performs will, from an early age on, be associated with an "inside" perspective—the experience of performing those behaviors in context (Russell, 1996). This experiential perspective may well be one starting point for inferring experiences in the *other* person who is observed performing those behaviors. First a link is created between representations of behaviors (e.g., seeing a growling face; observing a reach-and-grasp movement) with one's own characteristic experiences when performing these behaviors (e.g., feeling angry and agitated; succeeding in pursuit of a goal); later, this associative link between behavior representation and first-person experience may be "transferred" such that observing another's behavior does not just help simulate first-person experiences in the perceiver but triggers the third-person inference that the other has those characteristic experiences as well (Goldman, 2001; Meltzoff & Brooks, 2001).

In sum, expressive behaviors and fundamental movements are embedded functionally in physical contexts, which makes them transparent and thus excellent candidates for being comprehended without any mental state inference. Moreover, transparent behaviors allow for a link between representation of the other and experience of oneself, making them excellent candidates for facilitating mental state inferences.

Effect States

I now turn to the second path of resolving our puzzle—this one requiring mental state inferences that do not rely on behavior input. Such in-

ferences target one of two very different classes of mental states. Let us roughly call them mental states that are often *effects* and mental states that are often *causes* of behavior. Those that are effects include perceptions and emotions. They can be predicted fairly well from the context, general knowledge of the actor, one's own reactions (simulated or shared), and a set of lawlike generalizations about what people want, need, see, think, and feel under a variety of conditions. These mental states are relatively easy to infer because they constitute relatively reliable responses to the world. Now, when such states are associated with expressive behaviors (e.g., eye movements, facial expressions), we return to the first path of solving the puzzle. But these "effect" states may well be inferred even before they are expressed, and even in situations that do not allow for observation of the agent's expressive behavior (e.g., when the perceiver is not physically copresent with the agent).

By contrast, there are mental states that exist primarily as causes of behaviors (or of other mental states), such as desires, beliefs, and intentions. These states are far more difficult to infer, because they are context-specific and can be quite idiosyncratic. Their inference will rely on transparent behavior input (expressive behaviors, basic movements) and additional knowledge about the agent's past actions and general dispositions. For example, when an agent revealed such a mental state in past circumstances (e.g., when explaining his or her action), the ascribed desires and beliefs can be stored and held ready for the next time that certain identical or parallel triggers in the environment are observed to be present. Some form of intentional perspective taking may also be recruited for these purposes ("In his position, what would I want, like, or think?"). So, whereas *effect* states may be inferable without behavior input, *cause* states require behavior input as well as several other sources of information.

We can now summarize the likely ways in which human perceivers solve the inference puzzle that emerges between behavior and mental states. Along one path, they are sensitive to certain "transparent" behaviors that do not require prior inferences of mental states. In addition, they are sensitive to the intentionality of behaviors (judged from cues other than mental states), which in turn guides the selective search for mental states and thus kick-starts the social-cognitive system without circularity. Along a second path, human perceivers are sensitive to the difference between *effect* mental states and *cause* mental states, with the former being to some degree predictable or inferable from physical contexts or transparent behaviors, none of which require a meaning analysis that itself presupposes mental states. The human perceiver is thus not operating circularly but rather by using a parallel strategy: identifying behaviors whose meaning can be understood without prior analysis of

mental states and identifying mental states whose existence and meaning can be established without prior meaning analysis of behavior. These kinds of behaviors and these kinds of mental states provide starting material for the social-cognitive system, and once it gets started, it can easily ratchet up its inferences, because starting behaviors can then be used for inferring more difficult mental states, and starting mental states can be used to explain the meaning of more difficult behaviors.

THE SECOND PUZZLE:
HOW DO WE SPEAK ABOUT THE MIND?

The second puzzle concerns the ways human perceivers represent and identify mental states in language. It seems clear that humans talk about the mind. In doing so, they are not deterred by the fact that mental states are unobservable; after all, there are plenty of unobservable states and objects that we comfortably discuss (e.g., justice, electrons, the Big Bang). But the exact relationship between mental states and the words that refer to them is quite unclear. On the one hand, mental state terms are notoriously difficult to define, certainly vague, sometimes equivocal, as Uleman (Chapter 16, this volume) argues. Many emotion words, for example, refer to behaviors (a sad look, an angry face) as much as to internal states (I feel sad, she is very angry); it would take us a while to explain what it means to *trust* someone; *believing* can refer to a spiritual, intellectual, perceptual, or emotional internal state; waking up from our dreams, we struggle to find words that describe the feelings, thoughts, and images we just inhabited. One way to summarize these mind–language relations is to propose that there are a limited number of language terms that are used for a large variety of mental states, acts, and experiences, changing from context to context, from agent to agent, thereby blurring the boundaries of meaning.

But others diagnose the exact opposite problem. Recently, Sabini and Silver (2005) argued that the language of mental states is far richer than the mental world they describe, that is, "There are fewer unique mental states than one might have thought" (p. 9). This is a rather counterintuitive position, so let's look closely at some of the evidence that Sabini and Silver provide for their claim.

In a vignette study (Sabini, Garvey, & Hall, 2001), participants imagined being the protagonist in a story in which someone was helping them move from one office to another. In the process, the helper stumbles on a piece of pornography that either (1) truly belongs to the protagonist or (2) actually belongs to a former office occupant—though the helper doesn't know that. In case (1) participants described their emo-

tion as equal parts of shame and embarrassment, whereas in case (2) they described it more as embarrassment than shame. The authors took this to be evidence that "shame and embarrassment are different interpretations of the same raw feeling and these interpretations take into account the different conditions that surround them" (p. 6). But what is this "same raw feeling"? Who makes the judgment of sameness?

The protagonist's psychological world is, in each case, an interwoven complex of realizing, fearing, sensing, intending, and much more; and there is clearly some overlap of these complexes across the two cases. But how do we pick out exactly one "raw feeling"? Sabini and Silver (2005) treat the mental world as segmentable and its segments as uniquely identifiable, but that is a treatment best applied to physical objects and not psychological states. Moreover, the complex psychological state each protagonist experiences goes far beyond what the words *embarrassed* or *ashamed* can capture. The vignette in fact may illustrate the limits of language to reference complex mental states rather than an abundance of language for the "same" mental states.

In another study (Silver & Sabini, 1978), the researchers created a video in which a (male) student tells a (female) friend that he didn't get into medical school. Soon thereafter, a mutual friend enters and excitedly reports that he got into Harvard Medical School. On further questions by the female friend, he also mentions that he received a full scholarship. After the Harvard admittee leaves, the unsuccessful student complains to the female that their friend was bragging. Participants who saw this video are asked to describe the complaining student's emotion, and they overwhelmingly ascribe envy to him. Arguably, however, the emotion that the complaining student himself felt was righteous indignation or anger. From this Sabini and Silver (2005) conclude that "the experiences of an envious person and of a righteously indignant person can be the same" (p. 15), that "righteous indignation and envy . . . are the same experiential state" (p. 15). Thus, Sabini and Silver take this to be another illustrative case of two language labels referring to one and the same emotional state.

But do spectators and the actor himself refer to the *same* emotional states? Again, who would be the judge of this sameness? The actor would certainly deny that he feels envy—and, if we let him be perfectly honest and truthful, he will not *feel* envy (at least in the moment). So how can all spectators agree that his emotion was one of envy, not of anger? Are they right in their third-person perspective, and is the actor wrong in his first-person perspective?

I want to suggest that in both stories (the pornography discovery and the medical school announcement), we are dealing with vaguely bounded *emotion–action complexes*, and different descriptors pick out

different aspects of those complexes. Take the medical school story. By stipulation, the actor describes what he is consciously feeling at the moment; but the spectators describe the whole emotion-plus-action complex with which they are presented (reflecting a rather general actor–observer asymmetry in people's attention to experiences vs. actions; Malle & Knobe, 1997b; Malle & Pearce, 2001). No spectator would censure the actor by saying "No, no, you *feel envious*!" At best, they might say "You *are* envious," describing not a conscious experiential state but a combination of antecedents, behaviors, and unconscious emotions. Consider this parallel case: If a spectator said "He is envious" and the actor said (honestly) "I am not envious," we would not conclude that *envious* and *not envious* are two terms for the same state. Nor should we conclude, as Sabini and Silver (2005) suggest, that the terms *envy* and *righteous indignation* refer to the same single state. All that we should conclude in each case is that the parties are talking about two different things.

Likewise in the discovery story: The very fact that participants described the imagined feeling using both terms of shame and embarrassment, but in slightly different blending, suggests that we are dealing with complex and subtly different experiences. There are of course some similarities in the two variants of the story. In each, the protagonist realizes that the helper just found a piece of pornography and that the helper will assume it is the protagonist's. But then the differences begin. The one who actually owns the piece may feel "caught" and as a result be mortified; the nonowner may be surprised, perhaps shocked, fearing that the helper will have a (false) bad impression and quickly searching for ways to change the helper's impression. I have already used more than half a dozen mental state descriptors to describe these complex experiences, and I have not even begun to capture their nuances.

Emotions and experiences normally don't have boundaries that would allow us to reliably identify or count them like pebbles on the beach or words in conversation. Any given emotion (or, more generally, every state of mind) is a complex of combined and recombined mental and physical states, often tied up with intentions and actions. The corresponding language of those states is similarly complex in that it consists of terms that individually have context-sensitive meaning and are flexibly combined and recombined to represent the complexity of the mental world. And that is why, as mentioned earlier, mental language is vague and hard to define. Vagueness is the very feature that allows a limited number of terms to begin to describe a far larger (arguably infinite) number of states.

But how, one might object, could we ever use such a vague language of mind in a consistent, publicly shared way? In particular, how can ob-

servers even approximate an actor's complex experience with a handful of vague terms and limited behavioral evidence? The solution lies in a diversity of criteria that allow people—actors, interaction partners, or spectators—to assess the appropriateness of a mental state term used in a particular context. These criteria include (but are probably not limited to) introspection, memory, observation, joint attention, logic, and negotiation. Any given mental state ascription may not draw on evidence from all criteria, but as long as the available evidence reasonably converges, there will be stability in talk about the mind.

Scholars are often nervous about the role of introspection as a criterion of mental state ascriptions. Some (typically dismissed as "Cartesians") see it as the fundamental criterion against which others have to be measured; others see it as an illusion (e.g., Ryle, 1949). Wittgenstein (1953) argued that actors cannot use introspection as the criterion of their mental state talk, because it would mean speaking a "private language," which is no language at all. But there is no denying that actors sometimes do use their conscious experience as the guide to choose words of mind—for example, in response to legitimate questions such as "What are you thinking right now?" or "What are you feeling right now?" Actors just cannot always and solely rely on introspection to describe their psychological states; considerations of what others know, observe, remember from the past, and assumptions of logic and plausibility will often figure prominently as well. Similarly, observers will sometimes rely entirely on the actor's self-report to learn about another's mind—which is precisely why they ask questions such as "What are you thinking right now?" or "What are you feeling right now?" But observers do not always confine themselves to that kind of evidence; they may have reason to doubt the actor's self-report, or they may be interested in a more complex psychological state that is not solely constituted by conscious experience (as in the ascription of *envy* to the unsuccessful medical school applicant).

The meaning of mental state terms is thus not restricted to a kind of "private pointing" to inner states. Once we give up the idea that mental state terms rigidly refer to precisely bounded states, the use of any mental state term becomes a social act, and the appropriateness of this use is subject to the full variety of criteria available. To make a justifiable mental state ascription observers need to take into account what they see and know about the agent; what they generally know about people and the type of context the agent is in; what the agent reports, and how much he or she can be trusted; what others say or would say; and what the goals and stakes are of everyone involved. Precisely because observers make use of all this evidence, actors, too, must take it into account if they want to be credible to observers or converge with their judgments.

There will be cases of disagreement between actors and observers (and the medical school discovery story illustrates one of them). Facing such disagreement, people will clarify and compare their evidence, negotiate the relevance of this evidence for the claim at issue, and occasionally conclude that they were talking about different things. Once more, the vagueness of mental state terms is not a nuisance here but a blessing—it is the feature that helps bridge the actor–observer gap. Only because these terms are vague and often refer to an unspecified mix of internal states, behavioral indicators, and contextual constraints can actors and observers find sufficient agreement to render the language of mind intersubjective and meaningful. It may, in fact, be a necessary requirement of a language of mind to have unsharp boundaries and meanings that vary with context, neighboring terms, and the interlocutors' goals.

THE THIRD PUZZLE: IS MINDREADING A HIGH-LEVEL OR LOW-LEVEL ACTIVITY?

The final puzzle is that, even though mindreading appears to be a sophisticated and challenging activity of higher cognition, much of the mindreading that goes on in everyday interaction is not conscious, and some may not even be cognitive. I consider two cases here, one involving physiology in interaction, the other involving tacit inferences in real-time conversation.

When two people interact, their bodies often begin to synchronize in a number of ways: in posture, gesture, facial expression, timing and structure of speech, heart rate, and more (for reviews, see Chartrand, Maddux, & Lakin, 2005; Levenson & Ruef, 1997). Such synchronization will be imperfect and can at times be entirely absent; but often it is remarkable and has led to compelling demonstrations of behavioral mimicry (e.g., Bavelas, Black, Lemery, & Mullett, 1986; Bernieri, 1988; Chartrand & Bargh, 1999) and physiological linkage (Levenson & Ruef, 1992). Similar mechanisms are also responsible for the phenomenon of emotional contagion (Hatfield, Cacioppo, & Rapson, 1994), detectable in newborns who begin to cry when they hear other babies cry (Simner, 1971), in adults who quietly sit together and adopt one another's moods (Friedman & Riggio, 1981), and in crowds that may break out in violence once a few individuals model violence (Patten, 1999).

The best explanation for synchronization in interaction is that one person's emotional or physiological state is expressed in his or her behavior, this behavior is automatically imitated by the other person, in whom (on the basis of well-practiced associations and perhaps a common neural coding system, Jeannerod & Frak, 1999; Meltzoff, 2002)

similar internal states are generated. If needed, the second person could then correctly represent the first person's emotional or bodily state merely by relying on the default assumption that, all else equal, others will be in a similar state as he himself or she herself is. In this case at least, the assumption leads to accurate judgments because the perceiver's "evidence" for the other's mental state was caused by that very mental state (mediated by expressing and imitating bodies). Such active representation, however, will often be superfluous; the very synchronization may suffice to guide the interaction, leading to rapport and cooperation (Barsade, 2002; Bernieri, 1988). In a sense, mindreading has occurred without anybody really trying to read the other's mind.

The second case comes from conversation. Having a successful conversation requires a person to continuously track the other person's beliefs, goals, intentions, and emotional reactions. Some of this tracking consists of explicit perspective taking (conscious and deliberate reasoning about the other mind), such as when the pining teenager sits next to the class beauty and wonders whether she likes him. Some of the tracking occurs unconsciously but results in the conscious ascription of a specific mental state, as when it dawns on you that a friendly stranger on the street actually wants to sell you something; or that your politely nodding colleague doesn't actually know what the acronym HLM stands for. But, as Barker and Givón (Chapter 14, this volume) argue, much of this tracking in ordinary conversation occurs unconsciously. There are uncountable examples of speakers adjusting word choice and grammatical forms depending on what they think the audience understands (Fussell & Krauss, 1992; Givón, 2005). Here is one: "When you are ready to leave, just knock on *the door*," the homeowner says, disappearing from the repair person's view behind a sliding door. The homeowner can say "the door" because she can safely assume that the repair person knows which door to knock on—the one behind which she disappears while the repair person is watching. If she had called out that utterance from the other side of the house, she would have had to specify something like "the sliding door at the end of the hallway."

In general, acts of reference (to objects, actions, locations, times, etc.) are subtle exercises in perspective taking, as the speaker must consider what the other knows. Now, we shouldn't expect that speakers do some sort of calculations in every case. In fact, there are at least three sources of information that might preempt actual mental state inferences.

First, one's own perspective can serve as the standard. For example, if I recall in my conversation that I previously mentioned that my wife's name is Lara, I will subsequently refer to her simply by her name, Lara, implicitly assuming that my conversation partner also remembers that I

said who "Lara" is. These kinds of situations are of course open to error—because it is so natural for me that the name *Lara* refers to my wife, it may not even occur to me that my current conversation partner does not know that. (For further examples and evidence of such errors, see Barr & Keysar, Chapter 17, this volume.)

A second preemptive source of information is generic knowledge—knowledge that normally all members of a given community share (e.g., "*The* sun is coming out" vs. "*A* storm is rolling in). This will predictably get you into trouble when you misjudge your conversation partners' community membership (Fussell & Krauss, 1992).

Third, well-practiced scripts can support an action that seems to imply a certain mental state inference but is not actually made, at least not consciously. For example, after taking the order, the waiter reaches his hand out toward the guest, "assuming" that the guest knows what the waiter wants, and the guest correctly "infers" that the waiter wants the menu, rather than, say, the napkin or a tip. This well-practiced action may originally have been paired with the question "May I take your menu, ma'am?" (which considers the guest's knowledge and preferences), but over time this utterance and its attendant inferences became superfluous, as myriad interactions led to the desired result. The reaching action has acquired a powerful role in the whole script, triggering the guest's desired response and making mental state inference unnecessary—unless there is some doubt. If the menu lies on the far side of the guest's place setting, the script will have to be adjusted, and the waiter may reintroduce some sort of question, because he believes the guest might not know what he wants. Such adjustments can, in the end, become part of an alternate script, again relieving the agent of any (explicit or perhaps even implicit) mental state inferences.

Let me apply these three sources of preemptive information (self, generic knowledge, scripts) to a fascinating example offered by Bavelas and Coates (1992). Two strangers, involved as participants in an experiment, are told by the experimenter to give an opinion on some topic. After the experimenter leaves the room, A says, "You go ahead." Before A's last word is completed, B already smiles. A immediately laughs, B says, "Gee, thanks," and A responds with "You're welcome." Within a couple of seconds, the two have conducted a sophisticated conversation in multiple channels and correctly decoded ironic meaning. In particular, B's inference that A is being ironic is almost instantaneous, and so is A's recognition that B understands the irony. It seems likely that B didn't want to "give her opinion" and (implicitly) assumed that A didn't either; she also may have quickly searched for a way to get out of being the first to speak. So when A said "You go ahead" (perhaps with slightly exaggerated generosity), B could use her own reluctance as the basis of inferring

A's reluctance and therefore code A's utterance as a not-so-generous offer and the specific formulation as displaying generosity with irony. That is, the context and B's own feelings allowed her to interpret A's move as one of copping out, but the specific words he used were of a different script (the "generous offer") and were therefore understood as ironic. The rest of the conversation, with the irony mutually known, is played out according to the cultural script of responding to a generous offer ("Thanks"—"You're welcome"), with smiles, laughs, tone, and added words (e.g., "Gee . . . ") confirming the continued irony.

If facile mindreading occurs between strangers, we shouldn't be surprised that longtime couples can perform apparent feats of mindreading—such as one partner completing the other's sentence or answering a question before the other has even posed it. Long-term relationships benefit from improvements in all three sources of information. Through shared experiences and converging preferences, one's own mind becomes a more reliable indicator of the other's mind; one gains not only generic but agent-specific knowledge; and there are literally thousands of scripts that are practiced in the relationship every day. Long-term relationships also remove a notorious limitation of interactions among strangers: One learns in which contexts the other person does *not* feel or want the same thing as oneself does or will *not* act as other people do.

So, how can we reconcile the apparent low level of many mental states inferences with the seeming "high-level" character that many prototypical mental state inferences show? Are the processes that subserve this rapid system the same as the ones that engage conscious, deliberate reasoning about others' minds? I believe that the only way we can account for the full range of mindreading is by postulating not one mechanism that comes in degrees of conscious awareness, but a whole set of psychological tools that serve mindreading functions (see Ames, Chapter 10, this volume; Fernandez-Duque & Baird, Chapter 5, this volume). Some are fast and general, others are slow but aim at precision even in new situations. Some rely on stored knowledge of trends and patterns; others rely on the perceiver's own mental states in the specific context.

CONCLUSION: MINDREADING AS A MANIFOLD

All three puzzles considered here suggest that the processes underlying mindreading form a *manifold,* a complex array of related but distinct elements. In light of the first puzzle (the role of behavior in mental state inferences), I argued that there were multiple entries into the noncircular inferential relationships between behavior and mental states: the inten-

tionality concept, transparent behaviors, and "effect" states. In light of the second puzzle (how to talk about the mind), I suggested that there are multiple criteria for appropriate mental state ascriptions (introspection, observation, memory, logic, etc.). None of these criteria "defines" the meaning of mental state terms, but the convergence and social negotiation of the evidence at hand determines the successful application of such terms. In light of the third puzzle, finally, I concluded that people rely on multiple tools when dealing with other minds—tools that include explicit mental state inferences but also more implicit processes, such as emotional contagion, behavioral mimicry, and assumptions in conversation.

The picture that emerges is one of mindreading as a diverse toolbox that covers a broad range of stimuli, information processing mechanisms, and outputs. This toolbox includes a conceptual framework (e.g., the intentionality concept and distinctions among representational states as well as emotions); behavior observation capacities (e.g., for eye gaze, basic actions, and emotional expressions); and the ability to imitate and synchronize behaviors, emotions, and physiology. Add to that capacities I haven't discussed here—joint attention and joint action, imagination and pretense, and explicit perspective taking. The list could easily be expanded, and in many cases we don't yet know the fundamental processes or mechanisms that support the specific functions. But it seems clear that no simple notion of a "mindreading module" or an all-encompassing "theory" of mind will do the job of accounting for what people do when they make sense of other minds. As a social species, humans have evolved a large number of paths to other minds that provide both redundancy and flexibility to achieve their interaction goals in many different contexts and under many different demands. Even though the words we use to describe this fascinating phenomenon (*mindreading, mental state inference, theory of mind*) suggest a singular, bounded process or ability, only the recognition of manifolds at all levels—functional, cognitive, and neurological—will help us understand this unique and wondrous characteristic of human nature.

REFERENCES

Astington, J. W. (1993). *The child's discovery of the mind*. Cambridge, MA: Harvard University Press.

Baron-Cohen, S. (1999). The evolution of a theory of mind. In M. C. Corballis & S. E. G. Lea (Eds.), *The descent of mind: Psychological perspectives on hominid evolution* (pp. 261–277). New York: Oxford University Press.

Baron-Cohen, S., Tager-Flusberg, H., & Cohen, D. (2000). *Understanding other*

minds: Perspectives from developmental cognitive neuroscience (2nd ed.). New York: Oxford University Press.

Barsade, S. G. (2002). The ripple effects: Emotional contagion and its influence on group behavior. *Administrative Science Quarterly, 47*, 644–675.

Bavelas, J. B., & Coates, L. (1992). How do we account for the mindfulness of face-to-face dialogue? *Communication Monographs, 59*, 301–305.

Bavelas, J. B., Black, A., Lemery, C. R., & Mullett, J. (1986). "I show how you feel": Motor mimicry as a communicative act. *Journal of Personality and Social Psychology, 50*, 322–329.

Bernieri, F. J. (1988). Coordinated movement and rapport in teacher–student interactions. *Journal of Nonverbal Behavior, 12*, 120–138.

Blackburn, J., Gottschewski, K., George, E., & L—, N. (2000, May). *A discussion about theory of mind: From an autistic perspective.* From Autism Europe's 6th International Congress, Glasgow. Retrieved April 10, 2003, from www.autistics.org/library/AE2000–ToM.html

Bogdan, R. J. (2000). *Minding minds: Evolving a reflexive mind by interpreting others.* Cambridge, MA: MIT Press.

Buss, A. R. (1978). Causes and reasons in attribution theory: A conceptual critique. *Journal of Personality and Social Psychology, 36*, 1311–1321.

Carpenter, M., Akhtar, N., & Tomasello, M. (1998). Fourteen- through 18–month-old infants differentially imitate intentional and accidental actions. *Infant Behavior and Development, 21,* 315–330.

Carver, C. S., Ganellen, R. J., Froming, W. J., & Chambers, W. (1983). Modeling: An analysis in terms of category accessibility. *Journal of Experimental Social Psychology, 19*, 403–421.

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception–behavior link and social interaction. *Journal of Personality and Social Psychology, 76*, 893–910.

Chartrand, T. L., Maddux, W. W., & Lakin, J. L. (2005). Beyond the perception–behavior link: The ubiquitous utility and motivational moderators of nonconscious mimicry. In R. Hassin, J. S. Uleman, & J. A. Bargh (Eds.), *The new unconscious* (pp. 334–361). New York: Oxford University Press.

Friedman, H. S., & Riggio, R. E. (1981). Effect of individual differences in nonverbal expressiveness on transmission of emotion. *Journal of Nonverbal Behavior, 6*, 96–104.

Fussell, S. R., & Krauss, R. M. (1992). Coordination of knowledge in communication: Effects of speakers' assumptions about what others know. *Journal of Personality and Social Psychology, 62*, 378–391.

Gergely, G., Nádasdy, Z., Csibra, G., & Bíró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition, 56*, 165–193.

Givón, T. (2005). *Context as other minds: The pragmatics of culture, sociality and communication.* Amsterdam: Benjamins.

Goldman, A. I. (2001). Desire, intention, and the simulation theory. In B. F. Malle, L. J. Moses, & D. A. Baldwin (Eds.), *Intentions and intentionality: Foundations of social cognition* (pp. 207–225). Cambridge, MA: MIT Press.

Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1994). *Emotional contagion.* New York: Cambridge University Press.

Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.

Jeannerod, M. (1994). The representing brain: Neural correlates of motor intention and imagery. *Behavioral and Brain Sciences*, *17*, 187–245.

Jeannerod, M., & Frak, V. (1999). Mental imaging of motor activity in humans. *Current Opinion in Neurobiology*, *9*, 735–739.

Levenson, R. W., & Ruef, A. M. (1992). Empathy: A physiological substrate. *Journal of Personality and Social Psychology*, *63*, 234–246.

Levenson, R. W., & Ruef, A. M. (1997). Physiological aspects of emotional knowledge and rapport. In W. Ickes (Ed.), *Empathic accuracy* (pp. 44–72). New York: Guilford Press.

Malle, B. F. (1999). How people explain behavior: A new theoretical framework. *Personality and Social Psychology Review, 3*, 21–43.

Malle, B. F. (2002). The relation between language and theory of mind in development and evolution. In T. Givón & B. F. Malle (Eds.), *The evolution of language out of pre-language* (pp. 265–284). Amsterdam: Benjamins.

Malle, B. F., & Knobe, J. (1997a). The folk concept of intentionality. *Journal of Experimental Social Psychology, 33*, 101–121.

Malle, B. F., & Knobe, J. (1997b). Which behaviors do people explain? A basic actor–observer asymmetry. *Journal of Personality and Social Psychology, 72,* 288–304.

Malle, B. F., Moses, L. J., & Baldwin, D. A. (Eds.). (2001). *Intentions and intentionality: Foundations of social cognition*. Cambridge, MA: MIT Press.

Malle, B. F., & Pearce, G. E. (2001). Attention to behavioral events during social interaction: Two actor–observer gaps and three attempts to close them. *Journal of Personality and Social Psychology, 81,* 278–294.

McClure, J. (2002). Goal-based explanations of actions and outcomes. In W. Stroebe & M. Hewstone (Eds.), *European review of social psychology* (Vol. 12, pp. 201–235). New York: Wiley.

Meltzoff, A. N. (2002). Elements of a developmental theory of imitation. In A. N. Meltzoff and W. Prinz (Eds.), *The imitative mind: Development, evolution, and brain bases* (pp. 19–41). New York: Cambridge University Press.

Meltzoff, A. N., & Brooks, R. (2001). "Like me" as a building block for understanding other minds: Bodily acts, attention, and intention. In B. F. Malle, L. J. Moses, & D. A. Baldwin (Eds.), *Intentions and intentionality: Foundations of social cognition* (pp. 171–191). Cambridge, MA: MIT Press.

Patten, S. B. (1999). Epidemics of violence. *Medical Hypotheses, 53*, 217–220.

Perner, J. (1991). *Understanding the representational mind*. Cambridge, MA: MIT Press.

Povinelli, D. M. (2001). On the possibilities of detecting intentions prior to understanding them. In B. F. Malle, L. J. Moses, & D. A. Baldwin (Eds.), *Intentions and intentionality: Foundations of social cognition* (pp. 225–248). Cambridge, MA: MIT Press.

Read, S. J. (1987). Constructing causal scenarios: A knowledge structure approach to causal reasoning. *Journal of Personality and Social Psychology*, *52*, 288–302.

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, *3*, 131–141.

Russell, J. (1996). *Agency: Its role in mental development*. Hove, UK: Erlbaum.

Ryle, G. (1949). *The concept of mind*. London: Hutchinson.

Sabini, J., Garvey, B., & Hall, A. (2001) Shame and embarrassment revisited. *Personality and Social Psychology Bulletin, 27*, 104–117.

Sabini, J., & Silver. M. (2005).Why emotion names and experiences don't neatly pair. *Psychological Inquiry, 16,* 1–10.

Silver, M., & Sabini, J. (1978). The perception of envy: An experimental analysis. *Social Psychology, 41,* 105–118.

Simner, M. L. (1971). Newborn's response to the cry of another infant. *Developmental Psychology, 5*, 136–150.

Wellman, H. M. (1990). *The child's theory of mind*. Cambridge, MA: MIT Press.

Wellman, H. M., Hickling, A. K., & Schult, C. A. (1997). Young children's psychological, physical, and biological explanations. In H. M. Wellman & K. Inagaki (Eds.), *The emergence of core domains of thought: Children's reasoning about physical, psychological, and biological phenomena* (pp. 7–25). San Francisco: Jossey-Bass.

Whiten, A. (1999). The evolution of deep social mind in humans. In M. C. Corballis & S. E. G. Lea (Eds.), *The descent of mind: Psychological perspectives on hominid evolution* (pp. 173–193). New York: Oxford University Press.

Wittgenstein, L. (1953). *Philosophical investigations* (G. E. M. Anscombe, Trans.). Malden, MA: Blackwell.

Woodward, A. L. (1999). Infants' ability to distinguish between purposeful and non-purposeful behaviors. *Infant Behavior and Development*, *22*, 145–160.