Chapter One:

Intention Recognition and its Psychological Underpinnings

Daniel W. Harris

The hypothesis of this book is that humans regularly engage in a distinctive form of communication called intention recognition, and that this is what explains many of the features of human communication that make it unique. Intention recogniton happens when one person reveals to another person an intention to change their mind in a particular way.[1] Suppose, for example, that I intend to get you to believe that your fly is down. Acting on this intention, I could say, 'your fly is down', or I could get your attention and then pantomime the action of zipping up my fly. Either of these actions might lead you to form the belief that I intend you to have. How? By giving you evidence that this is what I intend for you to believe, and—assuming that you trust me—thereby giving you a new reason to believe it.

Humans routinely communicate in this way, but other organisms do not. The reason is that intention recognition relies on psychological capacities that are either unique or uniquely powerful in humans. My focus in this book is on three of these capacities that I take to be particularly important. When communicating, we rely on our *mindreading* capacity to understand what others intend by their communicative acts, and to predict how our communicative acts will affect the minds of others. We rely on our *planning* capacity to reason about what to communicate and how, and to coordinate our goals and activities in such a way that allows for us to communicate in an organized and cooperative way. And we rely on our capacity

[1] The idea of intention recognition originates with Grice (1957), though he did not use the term 'intention recognition'. Others have used the terms 'Gricean communication' (Moore, 2018) or 'ostensive–inferential communication' (Sperber and Wilson, 1995) for the same phenomenon.

for *language* to encode and decode richly structured evidence of what we intend by our communicative acts. In this chapter, I will delve into each of these capacities in turn, sketching how they make intention recognition possible, and how this sets us apart from nonhuman communicators.

1 Mindreading

To say that humans are mindreaders, in the relevant technical sense, is to say that we have a powerful capacity and compulsion to think of some parts of the world around us as agents, rather than as mere things, and to attribute mental states to them in order to understand their behavior.[2]

We routinely predict and infer the inner lives of everyone we encounter, including their beliefs, desires, intentions, emotions, and their thoughts about our own and others' thoughts. Suppose that you witness the following scene on the subway: A man and a woman are both walking toward the only empty seat. His gaze stays fixed on the seat, but hers momentarily raises toward him. At this moment she slows down, then stops walking toward the seat. How do you understand this event? Simple: both wanted to sit down and were walking with the intention of taking the last seat; the man didn't notice that the woman also intended to sit, but she came to believe that he did. Since her desire to avoid potential conflict outweighed her desire to sit, she abandoned her plan. You don't have to do any unusual thinking in order to figure all of this out. It's the sort of thing that you continually intuit about the social world around you. But it is nonetheless an impressive feat: in an instant you formulate a detailed representation of the hidden causes of your fellow passengers' actions, including the woman's higher-order thoughts about the man's thoughts.

Mindreading is a central engine of human social intelligence—a faculty on which we continually depend in our interactions with others. Imagine trying to carry on a normal relationship without being able to monitor your counterpart's moods or desires. Imagine trying to order a cup of tea if you had no ability to tell when the barista was paying attention to you, or when they wanted you to begin

[2] Mindreading is often called 'theory of mind' or 'folk psychology'. Following Apperly (2011, 3–4), I will avoid these labels so as not to prejudge the controversial question of whether our capacity to mindread is grounded in a kind of pre-scientific theorizing.

speaking or hand them your money. Imagine trying to navigate a crowded public place with no ability to formulate quick and mostly accurate hypotheses about where others were trying to get to. Imagine trying to hold down a job with no ability to think about what is expected of you, or about whether your colleagues or customers are pleased or annoyed by your efforts. Without our advanced capacity to predict and infer the mental lives of others, we would find ourselves adrift in a tangle of inscrutable bodies.

We have such a compulsive urge to mindread that we can barely stop ourselves from projecting mental lives onto the inanimate world around us. Heider and Simmel (1944) famously illustrated this compulsion by creating a simple animation in which several monochromatic two-dimensional shapes move around on a screen. We know full well that these simple shapes don't *really* have thoughts, and yet as viewers we can't help but see their movements as the unfolding of a tragic narrative wherein a small triangle and circle have an altercation with and evade an aggressive larger triangle, who then destroys his home in frustration. For generations of psychology undergraduates, the animation has been a palpable demonstration of our tendency to attribute thoughts and feelings at the drop of a hat. As Barrett (2000) puts it, we are prone to 'hyperactive agency detection'—a trait that manifests itself in our tendency to treat accidents as if they were intentional (Rosset, 2008), in our temptation to interpret nocturnal noises as intruders or ghosts, and, some have argued, in our urge to discern the will of supernatural beings behind unusual natural events (Barrett, 2000; Bering, 2002; Gray and Wegner, 2010).

As Heider and Simmel's animation also illustrates, our ability to understand and enjoy even rudimentary works of fiction is a further application of our mindreading capacity. In the same month that Heider and Simmel published their study, Disney released *Donald Duck: Contrary Condor*, in which Donald Duck sneaks into a condor's nest to steal an egg. In order to follow the plot, we must attribute an astonishing array of thoughts to the characters, including complex thoughts about other characters' thoughts: Donald wants a condor egg, but the mother condor is suspicious of him. He tricks her into falsely believing that he is one of her newly hatched chicks. However, the mother has another chick who knows that Donald isn't really its sibling, has figured out Donald's plan to deceive his mother and steal her egg, and intends for his mother to recognize Donald's deception. As viewers, we must take all of this for granted while we make sense of fur-

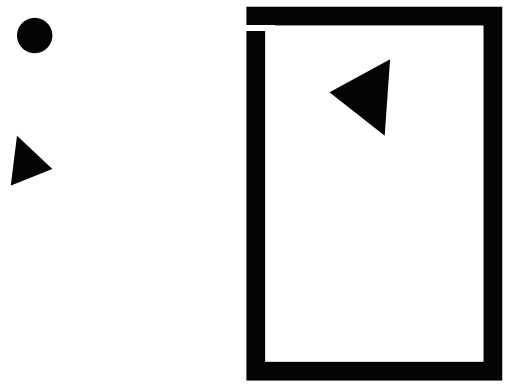

Figure 1: A scene adapted from Heider and Simmel's (1944) video. I find myself tempted to describe this moment in the video by saying that the circle and the small triangle, who are in love, have escaped the wrath of the jealous large triangle, shutting him up in his house. They are about to elope together while he destroys his own home in a fit of rage.

ther interactions and deceptions requiring us to attribute yet more beliefs, desires, intentions, and emotions. The viewer is meant to gather all of this from a 7-minute children's cartoon with no dialogue—something that is surprisingly easy for us to do. Many works of fiction—from Mozart's comedic operas to John le Carré's spy novels—present us with even more tangled webs of higher-order mental states to navigate, and thereby give our mindreading capacities more vigorous workouts. Some recent studies have found that engaging with works of fiction that involve complex, nuanced characters can at least temporarily enhance our mindreading ability (Kidd and Castano, 2013, 2018). This sort of mindreading practice may be an important part of the explanation of why we enjoy consuming certain kinds of fictions at all (Zunshine, 2006).

So far in this section, I hope to have defamiliarized a cognitive capacity that we rely on so heavily that we tend to take it for granted, not noticing how strange and powerful it really is. But this of course raises questions about how it works. These questions have been among the most discussed topics in cognitive science in recent decades: Is mindreading an application of a general-purpose capacity for theoriz-

ing about the hidden causes of observable events (Gopnik and Meltzoff, 1997), or is it grounded in dedicated cognitive structures (Baron-Cohen et al., 1985; Carey, 2009; Leslie, 1994; Scholl and Leslie, 1999)? If the latter, is there a single, unified mindreading system, or is mindreading the product of multiple systems that make distinctive contributions (Apperly, 2011; Butterfill and Apperly, 2013)? Are the cognitive structures underlying mindreading best likened to theories (Gopnik and Wellman, 1992; Perner, 1991; Wellman, 1990), or is mindreading a capacity to use our own decision-making capacity to simulate the thought processes of others (Goldman, 1995, 2006; Gordon, 1986)? Do we gain our mindreading capacity only around the age of four (Wellman et al., 2001; Wimmer and Perner, 1983), or are at least some parts of this capacity already online in infants (Carey, 2009; Carruthers, 2013; Leslie, 1987; Onishi and Baillargeon, 2005)? These questions have spawned enormous literatures, and I will discuss some of them in Chapter Four. However, they are mostly orthogonal to the main argument of this book.

There is one feature of mindreading that I need to highlight, because it plays a particularly important role in explaining human communication: It is that mindreading is, in the vocabulary of Fodor (1983), an informationally unencapsulated capacity. This means that mindreading processes can, in principle, draw on any of the mindreader's beliefs, so long as there is some salient and relevant connection between those beliefs and the behavior to be explained or predicted. Suppose that you happen to see a friend leaving the office of a divorce lawyer. What was he doing there? If you suspect that his marriage has been on the rocks, you might conclude that he is planning to end it. On the other hand, you might happen to know that your friend has been looking everywhere to buy a Nolan Ryan rookie card, and that this lawyer has a sports-collectables side hustle. Or you might happen to know that lawyers' offices in this county are required to offer public restrooms, and your friend's diabetes necessitates frequent pit stops. In these scenarios, your ability to attribute the right motivation to your friend depends on your beliefs about his personal life, but also about divorce law, the values of 1960s baseball cards, local bathroom by-laws, and the symptoms of diabetes. Mindreading inferences, like other explanatory inferences, are informationally omnivorous in this way. This is an important part of what makes mindreading so useful: there are no principled limits on what sorts of beliefs, desires, and intentions might influence another person's behavior—a point to which I will return below—and so an effective tool for

the mentalistic explanation of behavior needs to be capable of making use of any available information as well.[3]

2 Mindreading and Communication

Mindreading is not just for passive observation. It also allows us to actively influence others by helping us to choose actions that will change what they think in the ways that we want. A job applicant who gets dressed up for an interview in order to seem professional, a filmmaker who scores a scene with subtly dissonant music to build tension in their audience, and a viral marketer who gives free beard oil to a celebrity in order to spur acquisitive desire in his fans are all designing their actions to shape others' states of mind. Each of us engages in more mundane versions of the same sort of design every day.

We also act in ways that are calculated to influence what others think about our own thoughts. This requires making accurate predictions about how others will mindread with the aim of providing behavioral evidence that will push them toward particular hypotheses about us. In some cases, we may design intentionally misleading evidence of our thoughts. You could make a show of giving to charity or speaking out for a political cause in the hope that others will think favorably of your motivations, even if you don't actually care about these things, for example. Using reverse psychology, you might give me evidence that you want one thing in order to intentionally make me give you the opposite. These are both forms of manipulation—ways of hijacking others' mindreading capacities in ways that would not work if the targets understood your plan.

More often, we try to make our attempts to change others' minds transparent to them. Imagine that your car has broken down and you find yourself on the side of the highway, attempting to flag down passers by. You intend to get them to pull over, but you also intend for them to understand that this is why you are waving your arms. After all, if they can't figure out why you are waving—for example, if

[3] Of course, this is not to say that we actually search through all of our beliefs for relevant information every time we perform a mindreading inference. This would make mindreading computationally intractable, and so the mindreading process clearly has some imperfect method of limiting its search for relevant information. Very little is know about how it accomplishes this, but I will discuss some of the going hypotheses in Chapter Four.

they misinterpret your wave as a friendly greeting—they probably won't pull over. So, it is by getting them to recognize what you intend to persuade them of that you intend to effect this very persuasion. Successfully communicating, in this case, requires getting a passing driver to understand how you intend to influence them. This kind of communication—intention recognition—is the special kind of human communication that is at the center of this book.

One thing that makes intention recognition a powerful and flexible form of communication is that it depends on the active participation of those with whom we wish to communicate. Assuming that your addressee has a will to cooperate with you—whether this is due to generosity, fear, or habit—then knowing how you intend them to respond to you gives them a new reason to respond in that way. Similarly, if a cooperative addressee can tell that you are attempting to communicate with them in some way, this gives them a reason to put some effort into understanding you. Intention recognition thus depends on an underlying modicum of cooperativity, and leverages it to achieve feats of communication and coordination that would otherwise not be possible.[4] Unlike reverse psychology, which is a way of hijacking the addressee's mindreading capacities without their consent, intention recognition recruits the addressee into a collaborative effort. This is one reason why it is normally much easier to communicate overtly and openly than to manipulate people into thinking what we want them to think. Consider the range of things that you communicate in the run of a given day, and just imagine the sort of plotting and luck that would be required in order to manipulate your addressees into thinking these things without tipping them off to what you were trying to do. Even in situations when addressees are not wholly trusting or cooperative, and may not otherwise be inclined to go along, having the communicator's intentions made transparent may disarm their suspicions. Without revealing my intention to my addressee, I can accrue none of these considerable communicative advantages.

I will use the term 'communicative intention' to refer to any complex intention to produce a response in someone in part by revealing to them the intention to do so.[5] To act with a communicative intention is to perform a communicative

[4]This is not to say that those who communicate must be wholly cooperative. I will say more about the sort of cooperativity that is required—and what adding larger doses of cooperativity gets us—in §5.

[5]Grice introduced the notion of a communicative intention, although he did not use the term.

act.[6] To successfully interpret a communicative act is to recognize what response the speaker intended to have. Getting one's intention recognized in this way is the minimal way of successfully communicating by intention recognition. Thereby producing the response that one intends to produce constitutes a further sort of success, just as convincing someone of something goes beyond getting them to understand what you're saying.

Intention recognition is a cognitively demanding way of communicating. It requires us to quickly and accurately surmise the contents of others' thoughts, including their present and future thoughts about our own thoughts. Although this involves considerable cognitive labor, it is well worth the trouble. In return, we get a way of communicating that is normally much more powerful and flexible than other available options. Much of this power and flexibility is a direct result of the power and flexibility of our capacity for mindreading.

As my divorce-lawyer example illustrates, a single observed behavior can serve as evidence of many different states of mind, depending on the other background information that is available. By the same token, a single signal type can be used with a wide range of different communicative intentions, because we can count on our addressees to interpret it in light of their other information that bear on our beliefs and motivations. This flexibility is what explains several of the unique properties of human communication.

First, consider our use of novel signals. For example, you could do your roommate's chores in order to passive-aggressively communicate your annoyance at their

I follow the usage of Strawson (1964), Recanati 1986, and Neale 1992 in using "communicative intention" to refer to the entire complex plan consisting of an intention to produce a psychological effect in an addressee together with the intention to reveal the effective intention to them. It is worth noting that Sperber and Wilson (1995) use the term "communicative intention" in a different, conflicting way, and their usage has become common as well. They distinguish what they call the "informative intention" (the intention to inform the addressee) from what they call the "communicative intention" (the intention for the addressee to recognize one's informative intention). I use the term to refer to the complex consisting of both parts.

[6]The term 'communicative act' is due to Bach and Harnish (1979). Like them, I take communicative acts to be a genre of illocutionary acts, which are in turn a genre of speech acts. Some illocutionary acts that I don't take to be communicative acts include the act of pronouncing a couple married, the act of testifying in court, and the act of voting on a bill in a legislature. In Bach and Harnish's terminology, these are 'conventional acts'—ritualized moves within social institutions, outside of which they could not be performed.

messiness. You could play a particular song at the right moment in order to communicate that it's time to dance. You can communicate by slamming a book down on a co-worker's desk, by looking your sibling straight in the eye while flushing their goldfish down the toilet, by suddenly turning your car around after your child has once again kicked the back of your seat, or by placing the severed head of someone's horse in their bed while they sleep. Our ability to communicate in these ways not only gives us great flexibility and expressiveness, it also gives us ways to communicate messages for which we have no handy conventional signal. We can communicate with novel signals because there is no requirement that intention recognition be accomplished by means of a signal-type with a pre-established communicative function. When we mindread—in particular, when we infer the intention behind a communicative act—we seek to attribute whatever state of mind best explains an agent's behavior, whether that behavior conforms to a pre-established regularity or not. Even if neither you nor your friend has any history of communicating that it's time to dance by playing 'Girls Just Want to Have Fun', a context in which you are making a decision about what to do together along with the song's inherent danceability might be enough to reveal to your friend your intention to provoke an impromptu dance party. The fact that intention recognition is a kind of guided mindreading is therefore what explains why it needn't exploit prior conventions, and can instead occur idiosyncratically.

The possibility of one-off communication is part of what explains how we create conventions in the first place, and this helps to explain both our ability to add new signal types to our repertoires and the impressive size of our lexicons. The process of creating a new convention that pairs a signal-type with a meaning normally begins with an episode of communication that does not itself rely on a prior convention. This is what happens when we understand someone who has used a word or a gesture that we've never communicated with before: an episode of unconventional communication creates a precedent that can develop into a convention after repetition (Lewis 1969, 35–41; Schiffer 1972, 7, 119–136). This path to convention acquisition presupposes the possibility of non-conventional communication in the first place. So, the fact that intention recognition gives us a powerful method of communicating without relying on pre-existing conventions is an important part of the explanation of our proclivity for picking up new ways to communicate.

It is important to recognize that once we have acquired a repertoire of signal

types with stable meanings, their use in communication remains subservient to the broader process of intention recognition. When someone uses a conventionalized gesture or utters a sentence of their native language, they are not simply encoding their message so that it may be decoded by a recipient. Rather, the signals in our repertoires are stored methods for offering *partial* and *defeasible* evidence of our intentions.

To say that the evidence that we normally give of our intentions is *partial* is to say that the decoded meaning of a signal never fully specifies the intention behind it. There is always more inferential work to do, which requires access to information gathered from elsewhere. This is to say that the signals in our repertoires are semantically underspecified—a property that allows them to be used to reveal different intentions on different occasions. Pointing at a sculpture on someone's front lawn could be a way of communicating that it is ugly, that it is beautiful, that Mr. Johnson has once again violated the neighborhood association's policies, or any number of other things. The act of pointing on its own encodes little more than the information that the pointer intends to communicate *something* that bears *some* relation to the thing pointed at; the addressee must infer the rest of the message using information from other sources. Similarly, imagine that someone begins a conversation by saying, 'he has done it again'. From the sentence alone you can glean something about what the speaker intended for you to think—i.e., that some male has done something or other again—but you'll need other sources of information about the speaker's intentions to work out which male and which activity the speaker was talking about. The fact that our signal types are semantically underspecified in this way is the reason why we must so routinely and effectively bring background information to bear when interpreting signals.

Why are the signals in our repertoires semantically underdetermined in this way? Given our considerable storage capacity, why don't we just build up a collection of signals that could be used to make our intentions fully explicit? It is important to see that the inexplicitness of our signals is a feature rather than a bug. By storing a signal type that can be used to give evidence of many different intentions on different occasions, we vastly increase the expressive power of the signals that we have, allowing us to communicate more with less. The personal pronoun, 'he', can be used to refer to any male, including those whose names we don't know. Likewise, the gradable adjective, 'rich', can be used to pick out many

different degrees of wealth, depending on what counts as rich by the contextually relevant standard. What a speaker thinks of as rich is likely to be something different in a yurt in rural Mongolia than it is in a Silicon Valley boardroom. We could build predicates that would make these different degrees of wealth more explicit, but provided that the addressee can reliably-enough infer the speaker's intended standard, just one semantically underdetermined word can do the work that many fully explicit words could.

In addition to storage efficiency, Levinson (2000, 28–29) has also argued that semantic underdetermination gives rise to processing efficiency. He points out that speech is rather slow, topping out at about seven syllables per second before it becomes too difficult for most listeners to understand. If we were to take the time to fully articulate everything we say, it would take a long time to communicate anything. We overcome this "articulatory bottleneck" by taking advantage of the speed and multitasking capabilities of human thought, including our capacity for mindreading. In face-to-face interactions, we supplement our linguistic signals with signals in other modalities: the rhythmic and sing-song vocal modulations that linguists call 'prosody', myriad facial expressions and body language, and co-speech gestures and gesticulations. And in all communication, we attempt to achieve a delicate balance between, on one hand, encoding enough evidence of our intentions in the signal to give the addressee something to go on, and, on the other hand, leaving enough information out of the signal that communication can be as fast and efficient as possible.

None of this would be possible if mindreading weren't an informationally unencapsulated process—one that can draw on whatever relevant background information is available. Our ability to rely on background information when interpreting communicative acts—an ability that makes human communication enormously more efficient—is therefore a direct result of a basic feature of our capacity for mindreading.

Finally, consider our ability to communicate indirectly by using a signal in a way that goes beyond or conflicts with its usual communicative function (Grice, 1975). Depending on the circumstances, telling a co-worker, "I really need to wake up early tomorrow" could be a way of implying that you want to leave work, of indicating that you want to keep working instead of breaking for dinner, or of ironically conveying that you would like to stay up all night. None of these messages is con-

ventionally linked to to the sentence you've uttered. The key to understanding this phenomenon is to see that the signals in our repertoires are devices for offering *defeasible* evidence of our intentions. When you speak ironically—for example, by saying 'he is a fine friend' in order to imply that someone is treacherous—the sentence you utter encodes evidence that you have one intention when you actually have a different, conflicting intention. The reason that it is possible to communicate ironically is that the linguistic evidence of your intentions that you offer to your addressee can be defeated by the other evidence that they have available to them—such as their knowledge that the "fine friend" has recently betrayed you. But this is just a special case of the non-monotonicity of explanatory inferences, such as those involved in mindreading. When inferring someone's thoughts, any single piece of evidence can, in principle, be outweighed by all of the rest. Since even linguistic communication is powered by intention recognition, the evidence of our intentions that we encode in language obeys this principle too. When we communicate indirectly, we calculatedly exploit this feature of mindreading by offering evidence that we expect to be defeated.

3 Planning

In addition to our capacity for mindreading, intention recognition relies on our advanced capacity for planning.[7] By 'planning', I mean the process by which we form intentions and build them into complex plans.[8] One part of this process involves weighing our desires and inclinations in order to choose from among what we take to be our live practical options. But the live options are themselves constrained by the results of our prior choices, and each choice is normally a choice about how to further develop plans to which one has already committed. Your choice about whether to fly or drive to another city arises in the shadow of your previous decision to go there in the first place. A decision to drive will focus your attention on a variety of possible routes, forcing yet further choices. Most decisions are decisions about how to follow through on intentions formed as a result of prior decisions.

[7] Also known as 'practical reasoning', 'instrumental reasoning, and 'means-end reasoning'.

[8] I use the words 'intention' and 'plan' roughly interchangeably, though I take 'plan' to have a connotation of complexity. As Bratman (1987) puts it, "intentions are, so to speak, the building blocks of…plans, and plans are intentions writ large" (8).

Practical reasoning is normally a matter of reasoning from plans to subplans.

It is in the nature of intentions to play a role in this sort of hierarchical decision-making process. An intention is a relatively stable, action-guiding mental state—a kind of commitment about what to do. It is the result of a choice, and part of its purpose is to constrain the space of options for future choices. It is important for finite agents like us to have mental states of this kind. By constraining our options, they prevent our practical lives from overwhelming us. By virtue of being relatively stable commitments, they allow us to build up complex plans whose foundations don't erode away, and they allow other agents to coordinate with us by planning around what they know about our plans. On the view I have been sketching, which is due mainly to Bratman (1987; 2014), intentions are the lynchpins of both *intra*personally and *inter*personally coordinated action.

Intentions can play these roles because they are subject to principles that dictate which combinations of mental states are minimally rational, and these rational requirements exert pressure on our practical reasoning. If you already intend to drive to Toronto, and you believe that there are only two routes that will get you there by your planned arrival time, then you fail to be fully rational if you don't form an intention to take one of these routes. This follows from the principle that being rational requires us to intend what we take to be the necessary means to our intended ends.[9] Similarly, if you assume that the two routes to Toronto diverge, then rationality demands that you not intend to take both of them: intending to do inconsistent things, just like believing inconsistent things, is irrational.[10] And given that you take it to be impossible to drive across water, it would be irrational for you to intend to take a shortcut across Lake Ontario: it is irrational to intend to do what one takes to be impossible.[11] It is tempting to think of rational requirements as I have just described them as deeply normative constraints on how we *ought* to think and plan. But from my point of view in this book, it is the fact that our minds are designed to conform to them, at least often enough to confer significant benefits on us, that is most important. It is the psychological and social pressure to avoid irrationality, together with our other beliefs and intentions, that pushes us to make

[9]For defenses of variations on this principle, see Kant (1997, AA4:417); Hill (1973); Bratman (1987, 31–2); Broome (1999; 2013, §9.4) Schroeder (2004); and Holton (2008, 51).

[10]Bratman (1987, 31); Broome (2013, 136, 156–7); Holton (2011, 41–2).

[11]Bratman (1987); Broome (1999); Holton (2011).

some plans and steer clear of others.

In order to deliver plans that are rational by the standards of requirements like those I've just mentioned, planning has to be an informationally unencapsulated process, much like mindreading. The rationality of an intention depends on how it fits with all of our other intentions and beliefs, and so the planning process must at least in principle be responsive to all of those intentions and beliefs in order to be capable of delivering rational plans. Of course, this is not to say that humans always or even normally live up to this ideal. There are cognitive biases and performance constraints that nudge us toward irrationality in particular instances. My point is only that if the foregoing account of intentions works, then there can't be any principled architectural limitations on the planning process's access to our intentions and beliefs.

4 Planning and Communication

One reason that planning is an important part of intention recognition is that both the formation of communicative intentions and the subsequent execution of communicative acts tends to involve intricate planning. Consider again the relatively simple example of your car breaking down. Your intention to change what passing drivers do is itself a subplan of your broader plans. With your car broken down, you form an intention to extricate yourself from the side of the highway, and you reason that the best way to do this is to get someone to pull over. But you don't want to force anyone to pull over against their will, and so you need to find a way to get them to decide to pull over—to form an *intention* to help you. But how will you manage *that*? You could try to trick someone—say, by conspicuously pretending to be so badly injured that you don't even notice them driving by. But although this might get them to the side of the road, they might later become annoyed and decline to assist with your rescue. Better to appeal to their cooperativity. You can do this by making your plan understood to them—by revealing that you intend to prompt them to form an intention to pull over. Now you have a communicative intention—an intention to change someone's mind along with a subplan to do so by revealing this plan to them. But your practical reasoning is not over: you still need to decide what to *do* in order to reveal your intention to passing motorists. You don't believe that shouting could work, and so it would be irrational for you

to try that. But you think that waving your arms in the right way just might do the job, and so you settle on that.

Notice how much practical reasoning there is here, how it relies on your background knowledge about your situation and your drive to form a coherent plan, how it depends on your ability to anticipate the effects that your actions will have on the thoughts of passing motorists (a form of mindreading), and how each intention that you adopt raises new questions that must be answered by the adoption of further intentions. The process of forming a communicative intention and then designing evidence of it for an addressee often happens very quickly. Although we may sometimes be aware of ourselves doing it, we often aren't. But this should not lead us to doubt the existence of such reasoning, provided that it is the best explanation of how we manage to design communicative acts in the ways that we do.

These roles of planning in the performance of communicative acts help to explain humans' powerful capacity to design both what they communicate and how they communicate it for their addressees—the main subject matter of Chapter Two. We are able to make much more flexible and situation-independent decisions about what, when, where and with whom to communicate, and this is a direct result of the power and flexibility of the human planning capacity, together with the fact that communicative intentions are themselves the outputs of exercises of this capacity. A communicative intention may be a subplan of any plan that would benefit from altering another person's state of mind, and the informational unencapsulation of planning means that there are no limits on which schemes we might cook up that would so benefit. We are likewise far more adept than any other animal at customizing the signals by means of which we communicate with our addressees in mind. This is a direct result of the power and flexibility of our planning capacity together with the fact that the planning process can take communicative intentions as inputs. Once we have a communicative intention, we reason about how best to act on it by designing a signal that takes account of what we know about our addressees, including what we know about their thoughts.

In general, then, there is considerable explanatory insight to be gained from recognizing that communicative intentions, like other intentions, are elements in complex plans, and that that they serve as both conclusions and premises in practical reasoning. This conclusion naturally follows from the view that humans com-

municate by intention recognition together with the independently motivated view that a role in the planning process is central to the nature of intentions.

5 Intention Recognition, Cooperation, and Collective Planning

There is a second way that human communication is underwritten by our advanced planning capacity. In addition to planning our own actions, we also engage in collective planning with others. Collective planning is essential to the kind of coordinated social activities that humans excel at. To take a simple example, suppose that you and I intend to bake a cake together. We will need to coordinate on when and where to do it, who will bring which ingredients, and who will execute which parts of the recipe. If each of us thinks that the other will bring the flour, things will go badly. Bratman (2014) has argued that joint planning is driven by rational requirements that govern how multiple agents' beliefs and intentions fit together, just as any one agent's planning is driven by requirements on their own beliefs and intentions. In Bratman's terminology, agents with a joint intention—such as an intention to bake a cake together—are under pressure to form "meshing subplans" of that intention, and to make each other aware of these subplans. If I plan to supply the butter and milk, then you should not plan on bringing these things, and this will work best if you have accurate beliefs about my intentions. In general, failure to adhere to these requirements will result in uncoordinated action. And for this reason, there are social penalties for intentionally or negligently misleading others about the contents of any plans with which they may need to coordinate. If I tell you that I'll bring the butter and then don't do it, I make myself subject to rebuke, because I have given you leave to plan around a plan that I either never had or didn't stick to. In the context of collective planning, the pressure to obey rational requirements is, in part, social pressure.

Human communication is an example of the kind of social activity that often involves collective planning, and this is what allows it to be cooperative in the way that intention recognition tends to be.[12] Intention recognition is hard to get off the

[12] We can find earlier versions of this point in Grice (1975), Grosz and Sidner (1986), Clark 1996, Roberts (2012), and Bratman (2014), among others.

ground unless those involved are pursuing a shared conversation plan, minimally involving at least a shared intention to understand each other, and if they form meshing subplans of this plan. If I have no intention of being understood by you, then there is no rational pressure on me to say anything at all to you, or to design my utterances in a way that will make my intentions intelligible. If you have no intention of understanding me, then there is no pressure on you to pay any attention to what I say, or to provide me with feedback that makes further communication likely.

Although reliable intention recognition requires only this minimal degree of cooperativity, it often revolves around much more substantial shared plans. At any time in an ongoing conversation, we will normally have shared intentions about where to take the conversation next—intentions to share certain information, or to coordinate our plans in a certain way, or sometimes just to show interest in one another. These intentions, like others, give rise to questions about how to enact them, and these questions typically play a role in determining the topic of a conversation and the contributions that it is appropriate to make next (Roberts, 2012). If we are discussing our plan to bake a cake together, the question of where and when to do it might arise. This is a question about which subplan of a shared intention we wish to adopt. When it arises, answering it will normally become our immediate conversational goal. In another conversation, you might ask me a question about who directed the film, *Full Metal Jacket*. In a sufficiently cooperative setting, this question will give rise to a new shared intention of getting you the information that you seek. Ongoing communicative interactions, such as conversations, tend to be organized around hierarchical shared plans in this way. In Chapter Three, I will develop a theory of conversation plans that further fleshes out the role of these and other roles of shared planning in human communication.

This sort of organization is a crucial part of what explains the fact that humans are able to knit together our communicative acts into extended communicative exchanges—a feature of human communication that I discussed in Chapter One (§10). Again, this explanation depends on thinking of communicative intentions as elements in broader plans—in this case, shared plans. A communicative intention to answer a question is a subplan of a joint intention to resolve an issue that has arisen in the conversation. We form subplans like this as a result of the same kind of rational and social pressure that drives our other collective planning.

6 Language

What I have said so far applies to both linguistic and nonlinguistic communication, both of which are driven by intention recognition. The difference between the two is that linguistic communication relies on the exercise of an additional psychological capacity whose role in communication is to encode and decode richly structured evidence of our intentions. It is possible to flag down a car by waving your arms, but it would be much more difficult to nonverbally perform communicative acts like inquiring whether a bookseller has a copy of *Middlemarch* or informing your mother that you've dropped out of college to pursue a career as a professional video gamer. Performing these acts would require intending to produce mental states that are extremely specific and divorced from the situation at hand, and so it would be difficult to make yourself understood without offering specific evidence of what you intend.

Language is our most powerful solution to this problem. When you speak, write, or sign, a subsystem within your mind designs the expressions that you utter to serve as evidence of your intentions. When you perceive the products of someone else's language use, the same subsystem (or a closely related one) decodes evidence of their intentions from the expressions they have uttered. In this book, I will use the terms 'meaning' and 'semantic value' to talk about the evidential potential that is encoded in a linguistic expression.[13] The communicative job of your language system, then, is to encode meanings in the expressions that you utter and to decode the meanings from the expressions uttered by others.[14]

The evidence encoded in a sentence is richly structured because a sentence's

[13] It is typical to use the term 'semantic value' (and sometimes 'meaning') to refer to the contents of utterances, which vary from one context to another. This is not how I am using these expressions, which in my terminology denote context-invariant semantic properties of expressions. One way to put this is that, on my view, the meaning or semantic value of an expression is what you can know about what someone has said with the expression without knowing anything about the context or their intentions aside from the assumption that they are speaking literally and directly. I will defend this way of understanding meaning Chapter Seven.

[14] We can also speak of the meanings of the nonlinguistic signal types that we have in our repertoires, provided that the communicators in question have standing dispositions to use them to give and receive certain kinds of evidence of their intentions. But we should keep in mind that the cognitive structures that ground these dispositions are of a different, less specialized kind than those that ground our capacity to encode and decode linguistic meaning.

meaning depends in systematic ways on its syntactic structure and the meanings of its primitive parts. You are able to encode and decode a boundless variety of structurally complex linguistic expressions, and these complex expressions may be built up from the tens of thousands of primitive elements that you have stored in your lexicon. The only plausible explanation for this is that your mind contains a special-purpose subsystem that is capable of executing intricate encoding and decoding algorithms. These algorithms draw on a proprietary database of information that defines the phonological, syntactic, and semantic properties of the language that you speak. The project of generative linguistics—a project begun by Chomsky (1957; 1965) and that has now become a major research program in linguistics—is to reverse-engineer this body of information, which is called a grammar, as well as the algorithms by which we acquire and make use of them.

In Chapter Seven, I will defend a particular position about the nature of the language system: it is an informationally encapsulated and centrally inaccessible input-output system (or collection of such systems) of the kind first posited and described by Fodor (1983). To say that the language system is informationally encapsulated is to say that a person's language system carries out its task in a way that is insensitive to their beliefs, desires, and intentions. The radically false syntactic theory of a mistaken linguist won't interfere with their ability to use language, for example, because the part of their mind whose job it is to encode and decode linguistic meanings operates in a way that is insensitive to their beliefs. Instead, it relies on its own proprietary database of grammatical information. To say that the language system is centrally inaccessible is to say that the central-cognitive system(s), such as the system(s) responsible for mindreading and planning, have no access to the inner workings of the module or to the information stored in its proprietary database. The language system sends representations of meanings as outputs to the mindreading system when we perceive language, and it takes complex utterance plans as inputs from the planning system when we produce language, but the format and content of these boundary-crossing representations is severely limited, because the representational vocabularies of the language system and the central system(s) only partially overlap. Fodor and others (e.g. Frazier 1987) have argued that syntactic processing is accomplished by system that is "modular" in the foregoing, informationally isolated sense. I will also argue that a modular system underlies our capacity to encode and decode meanings, and that the theories

of contemporary compositional semanticists are best construed as theories of the proprietary database of this "semantic module".[15]

7 Language and Intention Recognition

Natural languages are complex codes. But from this fact we must not conclude that linguistic communication can be understood as the mere encoding of a message by a speaker and its decoding by an audience. This is because the linguistic encoding and decoding involved in human communication always subserves a larger process of intention recognition.[16] A speaker never simply takes a message that they want to communicate and encodes it, whole, in a sentence. And successful interpretation never consists in simply decoding a meaning from a sentence and leaving it at that. Upon grasping the meaning of a sentence that someone has uttered, we treat it as one source of information about the intentions with which they spoke—one among many potential sources. For this reason, decoding the meaning of a sentence is never enough on its own to allow us to understand a communicative act that was performed by uttering it. We also need background information in order to bridge the gap of semantic underspecification, and in order to tell whether the speaker was being indirect or nonliteral, and, if so, how.

This division of labor between linguistic encoding and intention recognition is a consequence of how human minds are organized.[17] The language system is an informationally encapsulated input-output system. When you interpret an utterance, it decodes the meaning of the expressions involved, but it lacks access to information about the context and the speaker's intentions. It can therefore only ever provide you with partial evidence of what the speaker intended to communicate. We therefore need to make use of nonlinguistic background information in order to infer the speaker's intended message. On the other hand, mindreading is an informationally omnivorous central-cognitive capacity. It is not in the business of treating any one source of information about a person's mental states as defini-

[15] I have previously defended this view in Harris (2022), of which Chapter Seven is a descendent.

[16] This claim was already implicit in the writings of Grice (1957; 1968; 1969; 1975), but it has been more clearly articulated by Sperber and Wilson (1995), Schiffer (2003), Neale (2004; 2005; 2016), and Scott-Phillips (2014, ch.1).

[17] This point about cognitive architecture is likewise influenced by Sperber and Wilson (1995).

tive when other sources of information are available. To put this another way, the reasoning that we do about others' minds is non-monotonic: any one piece of evidence about what someone thinks, including evidence delivered by the language system, may be overruled by stronger conflicting evidence. And, as rational planners who know that our addressees won't always take what we say as exhaustive or definitive evidence of what we intend, we exploit this fact to communicate in ways that go beyond what our linguistic signals encode on their own. Put another way: because mindreading can be guided by whatever background information is available, and because intention recognition is a kind of guided mindreading, intention recognition is a form of communication that can be accomplished in flexibly unformulaic ways.

8 How intention recognition makes us unique

Our capacity for intention recognition explains a number of features of human communication that make it quite different from any other form of communication that can be found in nature. Although many other animals have their own fascinating and powerful communication systems, they lack the advanced capacities for mindreading, planning, and language that make our form of intention recognition possible. This goes some way to explaining our large and easily expanded repertoires of lexical items, which can in turn be combined into indefinitely many signal types. It also explains our capacity to engage in organized communicative exchanges as well as our proclivities for indirect communication, context-sensitive signals, and our capacity to customize both what we communicate and how we communicate it for our addressees.

Each of the cognitive ingredients in our capacity for intention recognition has been held by some to be completely unique to humans—the result of radical discontinuities lying somewhere on the phyolgenetic tree between us and our nearest relatives.[18] However, these claims have proven controversial, and many others

[18]For arguments that there is no good evidence for nonhuman mindreading, see Lurz (2011); Penn and Povinelli (2007); Povinelli and Vonk (2003). For the claim that planning (a.k.a. 'instrumental reasoning' or 'practical reasoning') is unique to humans, see Papineau (2001) and Millikan (2006), and Camp and Shupe (2018) for a summary of the debate. Bratman does not quite claim that planning is unique to humans, but he sometimes contrasts genuine planning agents with

have claimed to find continuities linking humans to other animals, and in particular to other great apes. In Chapter 10, I will wade into this debate, and argue that although it is plausible that other great apes have homologues of our capacities for mindreading, planning, and language, these capacities nonetheless aren't well enough developed or integrated in other animals to support intention recognition, and so must have evolved to play other cognitive roles.

Of course, mindreading, planning, and language aren't the only things that make human communication unique. One reason that human humans can communicate about such a wide range of topics is that we can *think* about a very wide range of topics. Our conceptual capacities, like our communicative capacities, have greater expressive power than those of other animals. We can have thoughts about distant times and places, transfinite cardinals, moral obligations, the average American family, and—for that matter—the average family of Middle Earth dwarves. We can communicate about these things in part because we can think about them, and there is no good reason to think that non-human creatures can even think about them. With this difference in mind, we might wonder how much of the communicative difference between us and nonhuman animals is due to our capacity for intention recognition and how much is due to our conceptual range. And, in particular, it might be suggested that the very large lexicons that superpowers our linguistic capacities is explained largely by the very large stock of concepts that we use our lexical items to express.

I agree that it would be foolish to downplay the role of our conceptual repertoires in explaining humans' communicative prowess. But we should also be careful not to exaggerate this point. Many non-human animals can think about plenty of things about which they lack the capacity to communicate. Rats construct mental representations of the layouts of mazes (Tolman, 1948), but lack the means to pass this information on to other rats, for example. And although great apes track complex information about the social structure of their groups, and this information influences how they interpret affiliative and aggressive signals (Armstrong, 2018; Cheney and Seyfarth, 2007), they have no way to gossip about the latest po-

merely "purposive agents…who pursue goals in light of their representations of the world" but whose agency is not "embedded in planning structures" (Bratman, 1999, 5, 59). For the claim that language is entirely unique to humans, see Berwick and Chomsky (2015); Chomsky (1966); Hauser et al. (2002).

litical developments with group members who might have missed something. Although it is an interesting question whether humans have any ineffable thoughts, it should be uncontroversial that we are far less limited in which of our thoughts we can communicate than any other creatures. Conceptual expressive power alone does not guarantee communicative expressive power.

Still, we might wonder whether there are particular conceptual domains that play particularly important roles in human communication. One example worth considering is our capacity to think about the future and past, about what might be and what might have been, and about fictions. To borrow a term from linguists, the contents of our thoughts, like the contents of what we say, exhibit "displacement" (von Fintel and Heim, 2011; Hockett, 1960). Natural languages turn out to be quite dense with devices for talking about the non-actual and the non-present (see, for example, Portner 2009). Presumably, these would be useless to us if we could not think about the non-actual or the non-present. More importantly, there is a deep connection between displacement and planning, which is that our ability to choose from a range of options and then plan around the result depends on our ability to conjure up representations of those options in the first place. But to represent a potential goal that requires planning in order to achieve is, ipso facto, to represent an unactualized future possibility. A genuine capacity for planning—and so, it would seem, the capacity for intention recognition—therefore depends on a capacity for displaced thoughts. Although there is evidence that some non-human animals are capable of representing some displaced possibilities,[19] it is plausible that humans' advanced capacity for displacement in thought helps to explain not only our communicative expressive power, but also our capacity for intention recognition itself.

A different kind of worry about my emphasis in this chapter arises from the observation that our communicative capacities are embodied: they depend in their details on aspects of our anatomy that aren't narrowly psychological. We have highly expressive faces, manual dexterity sufficient for sign language and precise gesture, vocal cords capable of producing complex systems of phonemes and visual and auditory perceptual systems capable of picking up on all of these details. There is little doubt that many aspects of our bodies are well suited to the kinds of communicating that we do. However, there are also good reasons to think

[19] See, for example, Pepperberg et al (2019) on African Gray Parrots' capacity to entertain alternative possibilities in reasoning tasks.

that other great apes' communicative capacities are limited more by psychological factors than by anatomical factors. Great apes have enough manual dexterity to learn considerable numbers of sign-language-like gestures, but without most of the syntactic, compositional-semantic, and pragmatic properties of human linguistic communication, for example. Likewise, Fitch argues that great ape's auditory and vocal-tract anatomy would be sufficient for far richer verbal communication systems than we in fact find. "By process of elimination," Fitch concludes, the fact that non-human animals do not possess rich systems for verbal communication "appears to result from differences in neural control of the vocal apparatus, rather than vocal morphology" (Fitch, 2010, 327–328). Although it is true that nonhuman primates lack some of the vocal and manual flexibility of humans, their most significant limitations vis-a-vis humans are psychological.

9 Looking Forward

This chapter has served as an introduction to the theory of human communication that I will develop and defend in greater detail over the rest of this book. Along the way, I have made many controversial claims, noting some of them along the way, but not attempting to defend them in detail. That will be my task in later chapters.

References

Apperly, I. (2011). *Mindreaders: The Cognitive Basis of "Theory of Mind"*. Psychology Press, Hove and New York.

Armstrong, J. (2018). Provincialism in pragmatics. *Philosophical Perspectives*, 32(5–40).

Bach, K. and Harnish, R. M. (1979). *Linguistic Communication and Speech Acts*. MIT Press, Cambridge, Mass.

Baron-Cohen, S., Leslie, A. M., and Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition*, 21:37–46.

Barrett, J. L. (2000). Exploring the natural foundations of religion. *Trends in Cognitive Sciences*, 4(1):29–34.

Bering, J. M. (2002). The existential theory of mind. *Review of General Psychology*, 6(1):3–24.

Berwick, R. C. and Chomsky, N. (2015). *Why Only Us*. MIT Press, Cambridge, MA.

Bratman, M. (1987). *Intention, plans, and practical reason*. Harvard University Press, Cambridge, Mass.

Bratman, M. (1999). *Faces of Intention: Selected Essays on Intention and Agency*. Cambridge University Press, Cambridge.

Bratman, M. (2014). *Shared Agency: A Planning Theory of Acting Together*. Oxford University Press, Oxford.

Broome, J. (1999). Normative requirements. *Ratio*, XII(4):398–419.

Broome, J. (2013). *Rationality Through Reasoning*. Wiley-Blackwell, West Sussex.

Butterfill, S. A. and Apperly, I. A. (2013). How to construct a minimal theory of mind. *Mind and Language*, 28(5):606–637.

Camp, E. and Shupe, E. (2018). Instrumental reasoning in non-human animals. In Andrews, K. and Beck, J., editors, *The Routledge Handbook of Philosophy of Animal Minds*, pages 100–108. Routledge, London.

Carey, S. (2009). *The Origin of Concepts*. Oxford University Press, Oxford.

Carruthers, P. (2013). Mindreading in infancy. *Mind and Language*, 28(2):141–172.

Cheney, D. L. and Seyfarth, R. M. (2007). *Baboon Metaphysics: The Evolution of a Social Mind*. University of Chicago Press, Chicago and London.

Chomsky, N. (1957). *Syntactic Structures*. Mouton & Co., The Hague.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. MIT Press, Cambridge, MA.

Chomsky, N. (1966). *Cartesian Linguistics: A Chapter in the History of Rationalist Thought*. Harper & Row, New York.

Clark, H. H. (1996). *Using Language*. Cambridge University Press, Cambridge.

von Fintel, K. and Heim, I. (2011). *Intensional Semantics*. Unpublished Lecture Notes, online at http://web.mit.edu/fintel/fintel-heim-intensional.pdf, spring 2011 edition.

Fitch, T. (2010). *The Evolution of Language*. Cambridge University Press, Cambridge.

Fodor, J. A. (1983). *The Modularity of Mind*. MIT Press, Cambridge, MA.

Frazier, L. (1987). Theories of sentence processing. In Garfield, J., editor, *Modularity in Knowledge Representation and Natural-Language Understanding*, chapter 15, pages 291–307. MIT Press, Cambridge, MA.

Goldman, A. (1995). In defense of the simulation theory. In *Folk Psychology*, pages 191–206. Blackwell, Oxford.

Goldman, A. (2006). *Simulating Minds*. Oxford University Press, Oxford.

Gopnik, A. and Meltzoff, A. (1997). *Words, Thoughts, and Theories*. MIT Press, Cambridge, MA.

Gopnik, A. and Wellman, H. M. (1992). Why the child's theory of mind really is a theory. *Mind & Language*, 7(1–2):145–171.

Gordon, R. M. (1986). Folk psychology as simulation. *Mind & Language*, 1:158–171.

Gray, K. and Wegner, D. (2010). Blaming god for our pain: Human suffering and the divine mind. *Personality and Social Psychology Review*, 14(1):7–16.

Grice, H. P. (1957). Meaning. *The Philosophical Review*, 66(3):377–388.

Grice, H. P. (1968). Utterer's meaning, sentence-meaning, and word-meaning. *Foundations of Language*, 4(3):225–242.

Grice, H. P. (1969). Utterer's meaning and intention. *The Philosophical Review*, 78(2):147–177.

Grice, H. P. (1975). Logic and conversation. In Cole, P., editor, *Syntax and Semantics 3: Pragmatics*. Academic Press, New York.

Grosz, B. and Sidner, C. L. (1986). Attention, intentions, and the structure of discourse. *Computational Linguistics*, 12(3):175–204.

Harris, D. W. (2022). Semantics without semantic content. *Mind & Language*, 37(3):304–328.

Hauser, M. D., Chomsky, N., and Fitch, W. T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science*, 298:1569–1579.

Heider, F. and Simmel, M. (1944). An experimental study of apparent behavior. *The American Journal of Psychology*, 57:243–259.

Hill, T. E. (1973). The hypothetical imperative. *The Philosophical Review*, 82(4):429–450.

Hockett, C. (1960). The origin of speech. *Scientific American*, 203:89–96.

Holton, R. (2008). Partial belief, partial intention. *Mind*, 117(465):27–58.

Holton, R. (2011). *Willing, Wanting, Waiting*. Oxford University Press, Oxford.

Kant, I. (1997). *Groundwork of the Metaphysics of Morals*. Cambridge University Press, Cambridge.

Kidd, D. and Castano, E. (2013). Reading literary fiction improves theory of mind. *Science*, 342(6156):377–380.

Kidd, D. and Castano, E. (2018). Reading literary fiction can improve theory of mind. *Nature Human Behaviour*, 2:604.

Leslie, A. M. (1987). Pretense and representation: the origins of 'theory of mind'. *Psychological Review*, 94:412–426.

Leslie, A. M. (1994). ToMM, ToBy and agency: Core architecture and domain specificity. In Hirschfeld, L. and Gelman, S., editors, *Mapping the Mind*, pages 119–148. Cambridge University Press, Cambridge.

Levinson, S. C. (2000). *Presumptive Meanings: The Theory of Generalized Conversational Implicatures*. MIT Press.

Lewis, D. K. (1969). *Convention: A Philosophical Study*. Harvard University Press.

Lurz, R. W. (2011). *Mindreading Animals: The Debate Over What Animals Know about Other Minds*. MIT Press, Cambridge, MA.

Millikan, R. G. (2006). Styles of rationality. In Hurley, S. and Nudds, M., editors, *Rational Animals?* Oxford University Press, Oxford.

Moore, R. (2018). Gricean communication, language development, and animal minds. *Philosophy Compass*, 13(12):1–13.

Neale, S. (1992). Paul Grice and the philosophy of language. *Linguistics and Philosophy*, 15:509–559.

Neale, S. (2004). This, that, and the other. In Bezuidenhout, A. and Reimer, M., editors, *Descriptions and Beyond*, pages 68–182. Oxford University Press, Oxford.

Neale, S. (2005). Pragmatism and binding. In Szabó, Z. G., editor, *Semantics versus Pragmatics*, pages 165–285. Oxford University Press, Oxford.

Neale, S. (2016). Silent reference. In Ostertag, G., editor, *Meanings and Other Things: Essays in Honor of Stephen Schiffer*, pages 229–342. Oxford University Press, Oxford.

Onishi, K. and Baillargeon, R. (2005). Do 15-month-olds understand false beliefs. *Science*, 308:255–258.

Papineau, D. (2001). The evolution of means-end reasoning. *The Royal Institute of Philosophy Supplement*, 49:145–178.

Penn, D. C. and Povinelli, D. J. (2007). On the lack of evidence that non-human animals possess anything remotely resembling a 'theory of mind'. *Philosophical Transactions of the Royal Society B*, 362:731–744.

Pepperberg, I. M., Gray, S. L., Mody, S., Cornero, F. M., and Carey, S. (2019). Logical reasoning by a grey parrot? a case study of the disjunctive syllogism. *Behaviour*, 156(5–8):409–445.

Perner, J. (1991). *Understanding the Representational Mind*. MIT Press, Cambridge, MA.

Portner, P. (2009). *Modality*. Oxford University Press, Oxford.

Povinelli, D. J. and Vonk, J. (2003). Chimpanzee minds: Suspiciously human? *Trends in Cognitive Sciences*, 7(4):157–160.

Recanati, F. (1986). On defining communicative intentions. *Mind & Language*, 1(3):213–242.

Roberts, C. (2012). Information structure in discourse: Toward an integrated formal theory of pragmatics. *Semantics and Pragmatics*, 5:1–69.

Rosset, E. (2008). It's no accident: Our bias for intentional explanations. *Cognition*, 108:771–780.

Schiffer, S. (1972). *Meaning*. Oxford University Press, Oxford.

Schiffer, S. (2003). *The Things We Mean*. Oxford University Press, Oxford.

Scholl, B. J. and Leslie, A. M. (1999). Modularity, development, and 'theory of mind'. *Mind & Language*, 14:131–153.

Schroeder, M. (2004). The scope of instrumental reason. *Philosophical Perspectives*, 18:337–364.

Scott-Phillips, T. (2014). *Speaking Our Minds: Why Human Communication is Different, and How Language Evolved to Make it Special*. Palgrave Macmillan.

Sperber, D. and Wilson, D. (1995). *Relevance: Communication and Cognition*. Blackwell, Oxford.

Strawson, P. F. (1964). Intention and convention in speech acts. *The Philosophical Review*, 73(4):439–460.

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review*, 55(4):189–208.

Wellman, H. M. (1990). *The Child's Theory of Mind*. MIT Press, Cambridge, MA.

Wellman, H. M., Cross, D., and Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development*, 72(3):655–684.

Wimmer, H. and Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13:103–128.

Zunshine, L. (2006). *Why We Read Fiction: Theory of Mind and the Novel*. The Ohio State University Press, Columbus, OH.