The logic of indirect speech

Steven Pinker*[†], Martin A. Nowak[‡], and James J. Lee*

Department of Psychology, and ‡Program for Evolutionary Dynamics, Departments of Mathematics and Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138

When people speak, they often insinuate their intent indirectly rather than stating it as a bald proposition. Examples include sexual come-ons, veiled threats, polite requests, and concealed bribes. We propose a three-part theory of indirect speech, based on the idea that human communication involves a mixture of cooperation and conflict. First, indirect requests allow for plausible deniability, in which a cooperative listener can accept the request, but an uncooperative one cannot react adversarially to it. This intuition is supported by a game-theoretic model that predicts the costs and benefits to a speaker of direct and indirect requests. Second, language has two functions: to convey information and to negotiate the type of relationship holding between speaker and hearer (in particular, dominance, communality, or reciprocity). The emotional costs of a mismatch in the assumed relationship type can create a need for plausible deniability and, thereby, select for indirectness even when there are no tangible costs. Third, people perceive language as a digital medium, which allows a sentence to generate common knowledge, to propagate a message with high fidelity, and to serve as a reference point in coordination games. This feature makes an indirect request qualitatively different from a direct one even when the speaker and listener can infer each other's intentions with high confidence.

People often don't blurt out what they mean in so many words but veil their intentions in innuendo, euphemism, or doublespeak. Here are some familiar examples:

- Would you like to come up and see my etchings? [a sexual come-on]
- If you could pass the guacamole, that would be awesome. [a polite request]
- Nice store you got there. Would be a real shame if something happened to it. [a threat]
- We're counting on you to show leadership in our Campaign for the Future. [a solicitation for a donation]
- Gee, officer, is there some way we could take care of the ticket here? [a bribe]

This phenomenon poses a theoretical puzzle. Indirect speech is inefficient, vulnerable to being misunderstood, and seemingly unnecessary (because only a naïf could fail to see past the literal meaning). Yet politeness and other forms of indirectness in speech appear to be universal or nearly so (1). We all play this game and may be offended at those who don't, setting the stage for the hypocrisy and taboo in social life that are ubiquitously decried, yet ubiquitously obeyed.

Indirect speech also has considerable practical importance. It figures in the design of computer language understanding systems, which need to be programmed not to take indirect requests, such as "Can you tell me . . ." or "Do you know . . . ," literally. It is also a major bone of contention in the framing and interpretation of diplomatic agreements, and in the prosecution of bribery, extortion, and sexual harassment.

For 50 years, indirect speech has been intensively studied by linguists, philosophers, and psycholinguists, and the processes by which speakers veil their requests and hearers recover them have been well documented (1–3). However, the reason people engage in these maneuvers in the first place (as opposed to saying what they mean clearly and succinctly) is still largely unexplained. In this Perspective, we apply ideas from the analysis of signaling in evolutionary biology and from evolutionary game theory to illuminate possible advantages of indirect speech (4).

Existing theories of indirect speech are based on the premise that human conversation is an exercise in pure cooperation, in which conversational partners work together toward a common goal—the efficient exchange of information, in the influential theory of H. P. Grice (5), or the maintenance of "face" (esteem and autonomy) in Penelope Brown and Stephen Levinson's Politeness Theory (1). Yet a fundamental insight from evolutionary biology is that most social relationships involve combinations of cooperation *and* conflict (6, 7). This insight applies to communication among organisms no less than to physical actions, and indeed animal signaling has been found to involve exploitative manipulation as well as the cooperative exchange of information (8). In the human case, one has to think only of threats (the proverbial "offer you can't refuse"), dangerous secrets (hence the need for witness protection programs), contaminating leakage (such in blind refereeing, sealed bids, and clinical trials), and incriminating questions (for which one answer might be damaging, the other a lie, and a refusal to answer a de facto confession that those are the respondent's two options) (ref. 9; see also Gerd Gigerenzer's Law of Indispensable Ignorance, www.edge.org/q2004/page2.html#gigerenzer). The very existence of indirectness in language suggests that such adversarial dynamics might be in play in human communication. In cases of pure cooperation, one expects maximally efficient conspiratorial whispers; in cases of pure conflict, one expects a shouting match (8). The complex, coded communiqués that characterize human language bespeak a mixture of cooperation *and* conflict. This conclusion is reinforced by considering that most of the practical applications of indirect speech (diplomacy, extortion, bribery, and sexual harassment) take place in arenas of conflict.

Indirect speech takes many forms, including gestures of sympathy and deference (termed "positive politeness" and "negative politeness" in Politeness Theory). The phenomenon we address here is sometimes called "off-record indirect speech acts." We propose a theory in three parts, which apply to interactions with successively more subtle cost–benefit structures (4). The first part is the logic of plausible deniability. In a simple case like bribing a police officer, the appeal of a veiled bribe is intuitively clear: If some officers are corrupt and would accept the bribe, but others are honest and might arrest the driver for bribery, an indirect bribe can be detected by the corrupt cop while not being blatant enough for the honest cop to prove it beyond a reasonable doubt. A simple game-theoretic model can delineate the circumstances in which indirect speech is an optimal solution to this problem.

The second part of theory extends the game-theoretic logic to social situations in which there are no fines or other tangible costs and benefits, such as a diner without a restaurant reservation who bribes a maitre d' for a quick table, or a person who tenders a sexual invitation to a friend after a dinner. Unlike the driver and the officer, the speaker would incur no financial or judicial penalty were the hearer to turn down a blatant proposition, so the question here is why speakers still resort to innuendo.

The third part addresses scenarios in which people use indirect speech even when the degree of uncertainty about the other's intentions is low—either because variance among listeners' values is low (so speakers' confidence in their values is high), or because the listener is astute enough to understand the intent of a speaker's innuendo with high confidence. Why, in such cases, is a thinly veiled proposition still more acceptable than a naked one? The answer must pertain to some property of overt language itself, as opposed to the processes of social inference that power the interpretation of innuendo.

Part 1: Plausible Deniability

Consider a speaker whose speech obeys Grice's maxims (5) of efficient communication and is thereby always succinct, truthful, direct, and relevant. He is pulled over for running a red light and is pondering whether to bribe the officer. His choice is whether to remain silent or to say, "If you let me go without a ticket, I'll pay you $50."

Unfortunately, he doesn't know whether the officer is corrupt and will accept the bribe or is honest and will arrest him for attempting to bribe an officer. The game-theoretic conundrum where one actor does not know the values of the other has been explored by Thomas Schelling, who calls it the Identification Problem (9). The payoffs are as follows:

	Dishonest officer	Honest officer
Don't bribe	Traffic ticket	Traffic ticket
Bribe	Go free	Arrest for bribery

If the driver doesn't try to bribe the officer (first row), either way he gets a ticket; if he does offer the bribe (second row), the stakes are much higher either way: going free with just the cost of the bribe if he is facing a dishonest cop, or an arrest for bribery if he is facing an honest one.

Now consider a driver who knows how to use an "implicature" to convey an ambiguous bribe ("So maybe the best thing

would be to take care of it here"). Suppose he knows that the officer can recognize it as an intended bribe, and that the officer knows that he couldn't make a bribery charge stick in court because the ambiguous wording would prevent a prosecutor from proving his guilt beyond a reasonable doubt. The driver now has a third option:

	Dishonest officer	Honest officer
Don't bribe	Traffic ticket	Traffic ticket
Bribe	Go free	Arrest for bribery
Implicate bribe	Go free	Traffic ticket

The payoffs in this third row combine the very large advantage of bribing a dishonest cop with the relatively small penalty of failing to bribe an honest one. In these circumstances, indirect speech is the rational option. Note how this analysis is inconsistent with the traditional idea that indirect speech is an implementation of pure cooperation: The driver here is using indirect speech not to help the honest officer attain his goal (viz., to enforce the law) but rather to confound that goal.

The intuition that indirect speech can be an optimal strategy can be confirmed in a simple model of a Rational Briber. The expected cost of a bribe y can be calculated from (i) the proportion of officers that are honest, q; (ii) the cost of the bribe, c_0; (iii) the cost of the ticket, c_1 (which must be greater than the cost of the bribe, or it would never pay to bribe); (iv) the cost of an arrest for bribery, c_2 (which must be greater than the cost of the ticket; otherwise, it would always pay to bribe); and, the crucial psychological variable, (v) the probability p that an officer will treat a statement with a given degree of directness d as an attempted bribe. Directness is a semantic variable that corresponds to the degree of vagueness of the proposition (the number of readings) and the proportion of those readings that are consistent with its being a bribe rather than an innocent remark. An *if–then* proposition ("If you let me go, I will give you $50") is most direct; a leading question ("Is there some way to take care of it here?") is less direct; and a generic remark ("I've learned my lesson; you don't have to worry about me doing this again") is least direct. (In practice, p can be estimated empirically by asking people their degree of confidence that a given sentence was intended as a bribe.) Finally, the target of the indirect speech must de-

cide how to react to the proposition; this tendency can be captured by a decision function, L, which monotonically relates the directness of the proposition to the probability that the officer will treat it as an attempted bribe and act accordingly. Putting these together, the expected cost to a driver facing a corrupt cop is $y_c = c_0 p + c_1(1 - p)$; the cost when facing an honest cop is $y_h = c_2 p + c_1(1 - p)$; altogether, the driver's expected cost is $y = qy_h + (1 - q)y_c$.

Now, if the corrupt and honest cops share a single linear decision function L and hence have the same p for any proposition, the optimal level of directness will simply be determined by the fraction of honest officers. If $q > (c_1 - c_0)/(c_2 - c_0)$, then the optimum strategy for the driver is not to make any bribing attempt at all: $d = 0$. If the fraction of honest officers is less than this critical value, then the optimum strategy for the driver is to make the most direct and unambiguous bribing attempt, $d = 1$. In this model, so far, indirect speech is never an optimum strategy. The reason is that the cost functions are linear in d (Fig. 1).

For an indirect bribe to be advantageous to the driver, his overall cost function must be nonlinear. This situation could come about if the honest and corrupt officers employ nonlinear decision functions L_h and L_c relating the probability they will act (p) to the directness of the bribe (d), and if the two decision functions are distinct. That is, even if honest and dishonest officers interpret indirect speech the same way, and thus have identical assessments as whether an invitation to "settle it here" is an attempted bribe, the honest cop may be more hesitant to arrest the driver than the dishonest cop is to accept the implicit bribe, because of the burden of proof in a prosecution. In general, the expected cost for the driver is as follows:

$$y = q\{c_2 L_h(d) + c_1[1 - L_h(d)]\} + (1 - q)$$
$$\cdot \{c_0 L_c(d) + c_1[1 - L_c(d)]\}$$

In the simple case in which L_h and L_c are step functions, the scenario may be displayed as in Fig. 2.

The expected cost for the driver is as follows:

$$y = c_1 \text{ if } d < d_c$$
$$y = qc_1 + (1 - q)c_0 \text{ if } d_c < d < d_h$$
$$y = qc_2 + (1 - q)c_0 \text{ if } d_h < d$$

The intermediate region, $d_c < d < d_h$, has the lowest expected cost for the driver and, therefore, represents the optimum
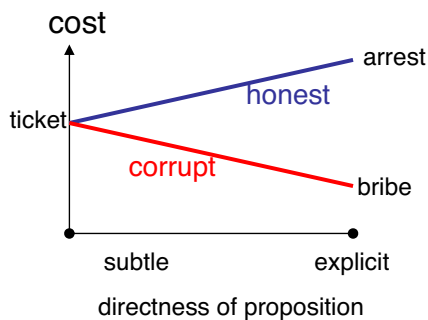
cost

Fig. 1. The expected cost to a driver of tendering bribes to honest and corrupt police officers with linear decision functions.

level of directness. This result confirms that there are plausible circumstances in which indirect speech is an optimal strategy and is a formal implementation of the everyday concept of plausible deniability.

The result does not depend on the officers' decision rules L_h and L_c being step-functions. They could also be sigmoid functions such as logistic or normal-ogive. As long as corrupt officers have lower threshold parameters than honest officers in their sigmoid function, that is, $L_h(d) > L_c(d)$ over some appreciable interval, it is easy to show that a wide range of parameter values yields a minimum of y with respect to d between the boundaries of silence at one end and bald directness at the other. The results also hold when there are distributions of the threshold parameters in the two populations of police officers. Yet another plausible extension to real life is the use of a sequence of statements escalating in directness (2), thereby probing the reaction of the officer ("What a beautiful morning. I'm very sorry for speeding. I know I'll have to pay for my mistake. I admire officers doing their duty. Can I make a contribution to the policeman's benevolent association? Is there some way we could avoid the paper-work and settle it here?").

How plausible is the key assumption that the decision functions of honest and dishonest officers differ? The answer depends on the determinants of *their* cost functions. Take the honest officer: why wouldn't he arrest anyone who offered a veiled bribe, exactly as a dishonest officer would accept all such bribes? The reason is that even if all dishonest drivers offer remarks that can be interpreted (correctly) as implicated bribes, some honest drivers make those remarks too, as innocent observations (this is inherent to the definition of indirectness), so any arrest might be unsuccessful. An unsuccessful arrest might be costly to the officer, exposing him to a charge of false arrest and the police department to punitive damages. The cost to the honest officer of

arresting the driver will thus depend on the proportion of dishonest and honest drivers who utter a remark with that level of directness, and on the professional rewards for successful arrests and the penalties for false ones. Conversely, for a dishonest cop, the cost depends on the amount of the bribe, the chances of *his* being arrested in a sting operation, and the penalty for being convicted of accepting a bribe. It is unlikely that the two decision functions would have the same shape.

Part 2: Relationship Negotiation

The second puzzle of indirect speech is why people use it in nonlegal situations, where there are no financial or legal payoffs and penalties. Consider bribery in everyday life, such as bribing a maitre d' at a popular restaurant to be seated immediately despite having no reservation. A restaurant critic, given the assignment to write about such an experience for *Gourmet* magazine (10), reported that the prospect of being turned down filled him with anxiety and that he resorted to indirect speech to tender the bribe, such as "I hope you can fit us in" or "I was wondering if you might have a cancellation." The second part of the theory thus seeks to explain why speech would be indirect in a nonlegal context such as a restaurant bribe or in a sexual overture among peers.

A key to this puzzle comes from Politeness Theory (1), which proposes that language serves two purposes: to convey a proposition (e.g., a bribe, a command, an offer) and to negotiate and maintain a relationship. People achieve these dual ends by using language at two levels. The literal form of a sentence is consistent with the *safest* relationship between speaker and hearer. At the same time, by implicating a meaning between the lines, the speaker counts on the listener to infer its real intent, which may initiate a different relationship. For example, in a polite request, such as "If you could pass the salt, that would be awesome," the literal content violates Grice's maxims of efficient conversation (5) because it is irrelevant and untruthful (i.e., an overstatement). The hearer implicitly reasons: "The speaker is saying that an outcome of an action by me is good. Therefore, he must be requesting it." The overall effect is that the intended content—an imperative—gets through, but without the presumption of dominance that would ordinarily accompany an imperative, with its tacit assumption by the speaker that he can expect the hearer's compliance.

This reticence raises the question of what kind of relationships, other than dominance, people are mindful of when choosing their words. Alan Fiske (11) has advanced the strong claim that human

relationships in all cultures fall into only three distinct types and that most of the complexities of social life within and across cultures may be explained in terms of variation as to which relationship type applies to a given dyad. (Fiske's theory also posits a fourth relationship type, "market pricing," but holds that it is specific to industrial and postindustrial societies.) Each relationship type is characterized by an ethos governing the distribution of resources between participants, and a straightforward evolutionary basis that specifies which kinds of dyads fall into that type of relationship by default. However, each relationship type can be extended to *other* dyads by negotiation and manipulation:

- The *dominance* or *authority* relationship is governed by the ethos, "Don't mess with me." It has a basis in the dominance hierarchies common in the animal kingdom, although in humans, it is based not just on brawn or seniority but on social recognition: how much others are willing to defer to you (12). It is generally communicated by psychophysical cues to strength and resolve (13).
- The *communality* or *communal sharing* relationship conforms to the ethos, "What's mine is thine; what's thine is mine." It naturally arises among kin, who are bound by shared genes (14, 15), within monogamous pair-bonds, who are bound by their shared children (16), and by close friends and allies, who are bound by shared interests (17). It can be extended to others by nonverbal cues of solidarity such as physical contact, communal meals, and coordinated movements and experiences (13).
- The *reciprocity* or *equality-matching* relationship obeys the ethos, "You scratch my back; I'll scratch yours." It has an evolutionary basis in reciprocal altruism (18, 19). It is usually signaled by tit-for-tat exchanges or division into equal portions but, unlike the other two relationship types, can be negotiated by people via explicit verbal contracts (13).

The assumed relationship type among a pair of individuals has dramatic effects on the behavior that is acceptable between them; not surprisingly, given that the relationship type governs the acceptable distribution of resources in a given social setting (13, 17). Behavior that is acceptable in one relationship type can thus be anomalous in another. For example, at a party, one might help oneself to a shrimp off the plate of one's spouse or sibling or close friend (communality), but not off
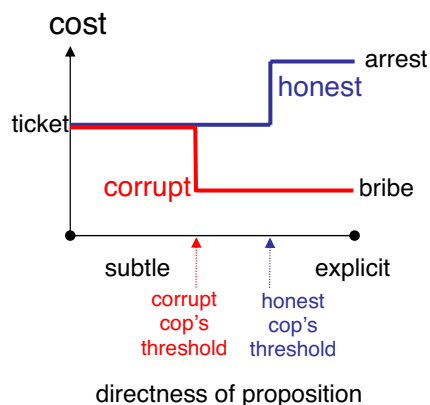
Fig. 2. The expected cost to a driver of tendering bribes to honest and corrupt police officers with nonlinear decision rules.

the plate of one's boss (dominance). Also, a guest at a dinner party (communality) would be perceived as crass, not fair, if at the end of the meal he pulled out his wallet and offered to pay his host for the meal (reciprocity).

When relationships are ambiguous, a divergent understanding between the parties can lead to the aversive emotion we call "awkwardness." There are awkward moments in a workplace or university when an underling or student makes a transition from a subordinate (dominance) to something closer to a friend (communality). Good friends (communality) are advised not to engage in a business transaction (reciprocity), like the sale of a car or a house, which can endanger the friendship. The ambiguity between dominance and sex (a kind of communal relationship) is the battleground of sexual harassment conflicts, and the ambiguity between friendship and sex gives rise to the frisson of dating.

The ultimate reason that relationship mismatches impose emotional costs (and hence motivate the parties to align their behavior with one relationship type or another) is the familiar principle from biology that cooperation is evolutionarily fragile (because it is vulnerable to cheaters), and that each variety of cooperation depends on a particular set of circumstances being in place for the cooperation to be adaptive (18, 19). Depending on their ecological niche and evolutionary antecedents, different organisms may cooperate via nepotism, mutualism, or reciprocity, or they may not cooperate at all, ceding resources via contests of dominance. Humans avail themselves of all of these options, facultatively choosing among them on the basis of their current social context rather than on a fixed, phylogenetically typical one. The anxiety surrounding relationship mismatches is the price we pay for having multiple, context-

specific ways of allocating resources available, with the consequence that a given type of behavior can vary radically in its adaptive value depending on which scheme is currently in effect. For example, helping yourself to a person's food or other resources can be a prerogative in the context of one relationship type but a case of theft in another. Ordering someone around can be a requirement of your job in one setting but a case of extortion in another.

The fraught nature of divergent relationships gives rise to a *social* identification problem that is isomorphic to the legal identification problem described by Schelling (9). In this case, the emotional cost of a relationship mismatch duplicates the payoff matrix of other risky propositions. Take the ambiguity between the authority relationship ordinarily commanded by a maitre d' (in which he assigns tables as he pleases) and the reciprocity relationship suggested by a diner offering a bribe (in which the maitre d' would be bound to offer a table in exchange for the bribe). The payoff matrix is identical in structure to the one for bribing a police officer, but the payoffs are reckoned by matches or mismatches in the assumed relationship between the two parties rather than in dollars and cents:

	Dishonest maitre d'	Honest maitre d'
Don't bribe	Long wait (dominance/ dominance)	Long wait (dominance/ dominance)
Bribe	Instant seating (reciprocity/ reciprocity)	Awkwardness (reciprocity/ dominance)
Implicate bribe	Instant seating (reciprocity/ reciprocity)	Long wait (dominance/ dominance)

An implicated bribe (third row), such as "I was hoping you might have a cancellation," combines the best of the first two rows: the high payoff of an overt bribe (a quick table) with the low penalty of not bribing at all (a long wait).

Part 3: Language as a Digital Medium
There is one remaining problem, which arises in cases where speakers use indirect speech despite a lack of genuine uncertainty on the part of one or both of the conversational partners. One such case is when the Identification Problem does not arise because the values of the listener are known—for example, if all of the police officers or maitre d's in a given city are known to be bribable. Another occurs when an indirect proposition is so conventional or transparent as to leave no doubt in the hearer's mind as to what was intended (the probability

that it is interpreted as intended is close to 1). That is, in many circumstances, both parties know when an overture has been made by innuendo. Can any adult really claim to be fooled by the etchings, the offer to "settle it here," or the advisory on the accidents that can befall a store? Any "deniability" in these cases is really not so plausible after all. The puzzle, then, is why, in cases of low or zero uncertainty about the agents' values and intentions, an obvious indirect overture should still feel less awkward than a direct overture that is sensed to be "out there" or "on the record."

This circumstance is addressed by the third part of the theory, which holds that language is tacitly perceived to be a *digital* medium. Discreteness is in the very design of human language. Features, morphemes, words, and phrases are concatenated, not blended, and each one's contribution to the meaning of the whole is qualitative: A sound that is halfway between *to bat* and *to pat* does not refer to an action that is halfway between batting and patting. Moreover, in all languages, real-world continua of space, time, and matter are digitized into discrete categories such as *in* versus *on*, past versus nonpast, and *pebbles* versus *gravel* (4, 20). Propositions with discrete truth

values, too, may be conveyed through the choice and arrangement of words. This does not imply that in practice linguistic information is transmitted with perfect fidelity (the phenomenon of indirect speech is an example to the contrary), but the intuition that language is a reliable medium is widespread as a folk theory of language (21), and belief in it may affect how people choose and interpret their words. This hypothesis has (at least) three corollaries.

First, overt propositions are perceived as *certain*, as opposed to merely highly likely. The relevance to relationship negotiation is that the signals of many relationship types may be analogue and highly ambiguous. How close does a man have to sit to a woman, how lavishly can he compliment her looks, and how secluded are the locales he leads her to before she con-

Pinker *et al.*

cludes that his intentions are sexual, and must be accepted or rebuffed? When faced with these ambiguities, people may treat certainty as a "focal point" in Schelling's sense (9). A focal point is a salient location that two rational agents can agree on when they would be better off coordinating their behavior than acting independently, and when there is no rational basis for choosing one value over another. Examples include a prominent landmark in a city where separated spouses agree to meet, or a round number that a buyer and seller in a negotiation can settle on. Although the preference for a focal point (when one exists) is an inherently rational strategy that does not depend on details of the agents' psychology, the question of whether a focal point exists, and where in the problem space it lies, depends on details of the local environment and on common properties of the agents' perceptual and cognitive systems that predictably single out the same point as psychologically salient. Daniel Kahneman and Amos Tversky, in their work on Prospect Theory (22), have shown in studies of people's tolerance for risk that human psychology categorically distinguishes 100% certainty from all other probability values. We suggest, then, that overt propositions are perceived as certain and act as focal points, whereas implicatures from indirect speech are perceived as being some measure short of certainty (even 99%). With the lack of a focal point to trigger a change of relationship, the speaker is given "the benefit of the doubt," and the relationship can remain unchanged. The deniability, then, doesn't have to be plausible, only possible.

A second corollary of language as a digital medium invokes Erving Goffman's hypothesis (23) that people always behave as if they are playing to an audience. A crucial feature of indirect speech is that it can be interpreted only in context. The hearer's interpretation taps into many aspects of the context, such as the lead-up to the speech act, the speaker's body language, or his tone of voice. Overt propositions, in contrast, may be perceived as *context-free*—their intent can be interpreted by eavesdroppers—and *lossless*—their intent can be transmitted perfectly along chains of gossipers (in the same way that other digital media, such as files of music or images, can be transmitted losslessly). According to this hypothesis, the deniability is plausible to the virtual audience, even if it is not particularly plausible to the hearer, and people tacitly consult the reactions of a virtual audience in retaining or switching relationship types. The ultimate reason is that relationship types are not solely dyadic but depend in part on the tacit ratification of a community, which may be necessary to back up

the authority of a dominant figure; yoke the interests of couples, families, and other communal groups; or disseminate information about the probity of reciprocating partners (24).

The third corollary invokes a concept that linguists, logicians, and economists have called common knowledge, mutual knowledge, and common ground (2, 9, 25–30). In common knowledge, not only does A know x and B know x, but A knows that B knows x, and B knows that A knows x, and A knows that B knows that A knows x, ad infinitum. As with other phenomena in linguistics in which a person is said to "know" an infinite number of things, the knowledge is not enumerated as an infinite list, of course, but is implicit in a finite recursive formula. In this case, it could be the formula y: "Everyone knows x, and everyone knows y" (2). Moreover, common knowledge can be ascertained perceptually, by observing that x is perceptible or broadcasted in public circumstances.

The paradigm illustration of common knowledge is the story of the Emperor's New Clothes. When the boy called out, "The emperor is naked!" he was not telling the onlookers anything they didn't already know. Yet he was conveying knowledge nonetheless: Now everyone knew that *everyone else* knew, and that everyone else knew that they knew, and so on, and that common knowledge licensed the people to challenge the dominance relationship commanded by the emperor. The moral for the present theory is that language is an efficient way of generating common knowledge.

This corollary of the language-as-digital-medium hypothesis, then, is that indirect speech merely provides *shared individual* knowledge, whereas direct speech provides *common knowledge*, and relationships are maintained or nullified by common knowledge of the relationship type. Imagine that Harry says, "Would you like to come up and see my etchings?" and Sally demurs. There is little or no uncertainty about Harry's intent, and none about Sally's: Sally knows that she has turned down an overture, and Harry knows that she has turned down an overture. However, Sally doesn't necessarily know that *Harry* knows; she might think to herself, "Maybe Harry thinks I'm naïve." In addition, Harry doesn't necessarily know that *Sally* knows that he knows; he might think to himself, "Maybe Sally thinks I'm obtuse." Although there is individual knowledge, there is no common knowledge, and they can maintain the fiction of friendship. In contrast, if Harry were to have said, "Would you like to come up and have sex?" then Harry knows that Sally knows that Harry knows that Sally knows, and so on. With this

common knowledge, they *cannot* maintain the fiction of a friendship, and they would have the sense that "it's out there," and that "he can't take it back."

The ultimate reason that common knowledge would mandate a relationship change is that the alternative—maintaining the relationship in the presence of common knowledge that contradicts it—would entail the listener relinquishing her claim to rationality, resolve, or honesty. She would be admitting that, despite possessing knowledge that is incompatible with the presumptions of pure friendship, she is willing to pay the costs or exploit the perquisites that go with it (e.g., accepting or offering favors without expectation of reciprocation). Similar tacit confessions would be entailed by other failures to change a relationship following a commonly known direct overture, as in bribes and threats.

This three-part theory (plausible deniability, relationship negotiation, and language as a digital medium) makes numerous empirical predictions about how people interpret the intent and perceive the relationship between a speaker and a hearer depending on the level of indirectness in the speaker's words. The plausible-deniability hypothesis predicts that the directness of speakers' wording of a veiled bribe or other overture (assessed on linguistic grounds) is not an arbitrary social ritual, like saying "Please" and "Thank you," but is predictable from strategic factors affecting its expected utility, such as the proportion of honest and dishonest officers in an area, the cost of a bribe, the cost of a ticket, and the cost of a bribery charge. For the listener's part, the directness of a speech act should predict their subjective estimates of the likelihood that the speaker intended the fraught proposition as opposed to making an innocent remark. The relationship-negotiation hypothesis predicts that indirect speech should be judged as generating less awkwardness and discomfort, as being more respectful, as better acknowledging the expected relationship with the hearer (such as affection, deference, or collegiality), and as making it easier for the participants to resume their normal relationship should the offer be rebuffed. The digital-medium hypothesis predicts that overt speech should be judged as *certain* in intent—that listeners and overhearers should assess the probability of a direct proposition as 1.0, with negligible variance among people, whereas even the most direct innuendo should be perceived as admitting of nonzero uncertainty, and with substantial variance. Finally, overt speech, compared with indirect speech, should be perceived as lossless in chains of gossip (the last link in a chain should be as confident in the speaker's intent as

the first) and in recursive embeddings of knowledge (the speaker and hearer should be as confident in their interpretation of the other's interpretation of their interpretation as they are in the simple interpretation itself).

The theory leads to further hypotheses about even higher-order psychological processes that might be triggered by people's experience in tendering and interpreting indirect speech. For example, hearers may credit a speaker with consideration for their dignity and feelings, or with greater social intelligence, if he uses indirect speech in a skillful way.

Most generally, indirect speech is a unique window into human social life (4).

If these analyses are correct, the phenomenon reveals a number of important facts about social behavior in *Homo sapiens*. Humans employ several, mutually incompatible, modes of cooperation and, as a result, are extremely touchy about their relationships. With some of their fellows (typically kin, lovers, and friends), they freely share and do favors; with others, they jockey for dominance; with still others, they trade goods and favors. People distinguish these relationships sharply, and when one person breaches the logic of a relationship with another, they both suffer an emotional cost. Nonetheless, humans often have to risk these breaches to get on with the business of life, and they often

use language to do so. In exploring the boundaries of relationship types, humans anticipate what *other* humans think about the relationship: what the other party in the relationship thinks; what overhearers and gossipers think; and what the other party thinks about what they think about what the other party thinks about what they think, and so on. The need to preserve their relationships while transacting the business of their lives can thus explain humans' tendency to fill their social life with innuendo, hypocrisy, and taboo.

1. Brown P, Levinson SC (1987) *Politeness: Some Universals in Language Usage* (Cambridge Univ Press, New York).
2. Clark HH (1996) *Using Language* (Cambridge Univ Press, New York).
3. Holtgraves TM (2002) *Language as Social Action* (Erlbaum Associates, Mahwah, NJ).
4. Pinker S (2007) *The Stuff of Thought: Language as a Window into Human Nature* (Viking, New York).
5. Grice HP (1975) Logic and conversation. *Syntax & Semantics*, eds Cole P, Morgan JL (Academic, New York), Vol 3.
6. Williams GC (1966) *Adaptation and Natural Selection: A Critique of Some Current Evolutionary Thought* (Princeton Univ Press, Princeton).
7. Trivers R (1985) *Social Evolution* (Benjamin/Cummings, Reading, MA).
8. Dawkins R, Krebs JR (1978) Animal signals: Information or manipulation? *Behavioral Ecology*, eds Krebs JR, Davies NB (Blackwell, Oxford), pp 282–309.
9. Schelling TC (1960) *The Strategy of Conflict* (Harvard Univ Press, Cambridge, MA).
10. Feiler B (2000) Pocketful of dough. *Gourmet* October:99–101.
11. Fiske AP (1992) The four elementary forms of sociality: Framework for a unified theory of social relations. *Psychol Rev* 99:689–723.
12. Maynard Smith J (1982) *Evolution and the Theory of Games* (Cambridge Univ Press, New York).
13. Fiske AP (2004) Four modes of constituting relationships: Consubstantial assimilation; space, magnitude, time, and force; concrete procedures; abstract symbolism. *Relational Models Theory: A Contemporary Overview*, ed Haslam N (Erlbaum Associates, Mahwah, NJ).
14. Hamilton WD (1964) The genetical evolution of social behaviour. I. *J Theor Biol* 7:1–16.
15. Hamilton WD (1964) The genetical evolution of social behaviour. II. *J Theor Biol* 7:17–52.
16. Alexander RD (1987) *The Biology of Moral Systems* (Aldine de Gruyter, Hawthorne, NY).
17. Tooby J, Cosmides L (1996) Friendship and the Banker's Paradox: Other pathways to the evolution of adaptations for altruism. *Proc Brit Acad* 88:119–143.
18. Trivers R (1971) The evolution of reciprocal altruism. *Q Rev Biol* 46:35–57.
19. Nowak MA (2006) Five rules for the evolution of cooperation. *Science* 314:1560–1563.
20. Pinker S (1994) *The Language Instinct* (HarperCollins, New York).
21. Reddy M (1993) The conduit metaphor: A case of frame conflict in our language about language. *Metaphor and Thought*, ed Ortony A (Cambridge Univ Press, New York).
22. Kahneman D, Tversky A (1979) Prospect theory: An analysis of decisions under risk. *Econometrica* 47:313–327.
23. Goffman E (1959) *The Presentation of Self in Everyday Life* (Doubleday, New York).
24. Nowak MA, Sigmund K (1998) Evolution of indirect reciprocity by image scoring. *Nature* 393:573–577.
25. Vanderschraaf P, Sillari G (Winter 2005) Entry for ''common knowledge.'' *The Stanford Encyclopedia of Philosophy*, ed Zalta EN (Metaphysics Research Lab, Center for the Study of Language and Information, Stanford University, Stanford), http://plato.stanford.edu/archives/win2005/entries/common-knowledge/.
26. Schiffer SR (1972) *Meaning* (Oxford Univ Press, New York).
27. Lewis DK (1969) *Convention: A Philosophical Study* (Harvard Univ Press, Cambridge, MA).
28. Smith NV, ed (1982) *Mutual Knowledge* (Academic, New York).
29. Clark HH, Marshall CR (1991) Definite reference and mutual knowledge. *Elements of Discourse Understanding*, eds Joshi AK, Webber BL, Sag IA (Cambridge Univ Press, New York), pp 10–63.
30. Chwe MS-Y (2001) *Rational Ritual: Culture, Coordination, and Common Knowledge* (Princeton Univ Press, Princeton).

Pinker *et al.*