Part 2

Cognition and Communication in Dyadic Encounters

Taylor & Francis
Not for distribution
Models of Communication, Epistemic Trust, and Epistemic Vigilance

MIKHAIL KISSINE AND OLIVIER KLEIN

INTRODUCTION

Because most social interactions involve routine use of language, one of the questions that stands prominently on the agenda of social psychology is of how people come to believe what they are told. In particular, it is the bread and butter of persuasion research. But we also come to believe many things others tell us without their necessarily pursuing a persuasive goal. When your neighbor brings up the persistent rain during his holidays in France, you will probably unquestioningly consider his description of the weather as accurate, and so even if you were yourself in Indonesia at the time. In such mundane examples, although the speaker is not pursuing any specific persuasive strategy, for the listener, believing the communicated information is a routine activity that constitutes the fabric of social interaction—and makes it possible. And yet, in spite of their importance to social life, such ordinary instances of belief validation have largely fallen out of the scope of social psychology.

The most straightforward issue such unexceptional validation processes raise is that of the connection between grasping the content of a statement, and believing it. Obviously, this distinction is not only conceptual as one can mentally represent the reference of false statements (e.g., Brussels lies under the Mediterranean sun) while knowing that they are false. The question is rather how do hearers switch from one to the other. Is grasping without believing always prior to believing? Or do we automatically believe whatever we understand, such that realizing that a statement is false entails unbelief?

One might think this issue occupies a central place in linguistics, and more specifically in pragmatics, a subfield devoted to the study of language use. Many,
if not most, linguists adhere to a dominant view, inspired by Grice’s (1989), according to which verbal understanding follows a tortuous inferential route, based on assumptions about speaker’s communicative intentions. As shown below, such a take on utterance interpretation entails that any communicated information has to be assessed before being translated into belief. This consequence went largely unquestioned, except by a minority who argue that, just as you (generally) believe what your eyes see, you (generally) believe what you are told until proven otherwise. Recently, however, proponents of the classic, Grice-inspired position have undertaken to actively endorse and defend the hypothesis that there are no communication-based beliefs without prior filtering (Sperber et al. 2010).

Although it remains largely unsettled, the question of validation of communicated information is thus central both to social psychology and to pragmatics. In this paper, we attempt to clarify the debate. The position that we defend is essentially the minority view. Yet, we do so in a nuanced way. The main claim of the paper is the following. Acquiring beliefs from speech is as direct as perception; however, this process may be mediated by a series of domain-specific and independent filters, so that in some cases, information is rejected without having been previously integrated among the hearer’s beliefs. In the next section, we expose in more detail two competing conceptions of communication, the inferential and the direct perception models, and focus on their consequences regarding vigilance toward communicated information. The relationship between communication and epistemic vigilance can be assessed from at least two standpoints. First, there exists experimental research relevant to this issue. Second, any claim about such a link should be evolutionary plausible. In Section 3, we propose a critical survey of the relevant experimental evidence and argue that existing data are incompatible with the inferential model. Next, we argue that, from an evolutionary point of view, the most plausible model of communication is that of direct perception, gradually supplemented with various epistemic filters. Finally, we flag two important programs of research in social psychology that could be profitably reinterpreted in the light of the direct perception model.

TWO MODELS OF COMMUNICATION

Let us now consider the two conflicting pragmatic models that we mentioned above. The first—and by far the most widespread one—stems from Grice (1989). The key idea is that interpretative processes can be reconstructed as an attribution of complex communicative intentions to the speaker. Although it is not clear that Grice himself conceived of this inferential mechanism as a psychologically valid model, rather than as a mere rational reconstruction of speaker’s meaning (see Saul, 2002), it has been subsequently promoted as a cognitive claim. This later shift brought a very strong, albeit often unquestioned, hypothesis about interpretative processing into contemporary psychology and cognitively oriented linguistics. So, it is perhaps worth emphasizing that our aim
here is not to defend any particular exegesis of Grice’s work; it is rather to high-
light the consequences of endorsing cognitive theories of Gricean inspiration.

The most paradigmatic transposition of Grice’s ideas into a cognitive model
is Sperber and Wilson’s (1995) relevance theory. According to this theory, the
hearer infers the meaning communicated through a linguistic utterance (e.g.,
this chair is blue) by attributing to the speaker the informative intention to
make manifest to the hearer a certain piece of information. The hearer infers
the informative intention he attributed to the speaker (i.e., speaker intends
hearer to believe that this chair is blue) by attributing to her the communicative
intention to make the speaker’s informative intention manifest to the hearer
(i.e., speaker intends hearer to realize that, by saying that this chair is blue, she
intends to make him believe that this chair is blue). In what follows, we refer to
this model of interpretation as the inferential model—IM, for short.

The IM puts communication apart from other information channels, such as
perception. Visual perception is direct and devoid of any epistemic gap. When
you see that there is a chair in front of you, (a) you do not come to the conclu-
sion that there is a chair in front of you by inference from the mental processes
underlying your perceiving the chair, (b) your believing that there is a chair in
front of you is not preceded by an internal deliberation about your acceptance
or not of this information (although you can subsequently reassess this belief in
the light of other information, and, perhaps, reject it ). In other words, nothing
comes between the visual experience of the chair and the belief that there is a
chair over there; visual perception is direct.2 By contrast, according to the IM,
extracting meaning from an utterance is inherently indirect. From this stand-
point, understanding that the content of the speaker’s utterance is a piece of
information amounts to grasping the speaker’s informative and communicative
intentions. An extra step is needed to arrive at the belief of that piece of infor-
mation: understanding that a speaker wants us to believe that piece of information
does not automatically cause us to believe that piece of information. The out-
put of the interpretation mechanism is limited to the content of the speaker’s
communicative intentions, and besides this, it does not provide any informa-
tion about the world. Integrating communicated content within one’s belief
box would thus never be automatic. As emphasized by Sperber et al. (2010),
this means that epistemic vigilance is part and parcel of the IM. Believing, or
not believing, the content that has been communicated depends on a filtering
mechanism of some sort, which checks received information for accuracy and
consistency with other beliefs. In the absence of epistemic assessment, grasping
the communicated content fails to lead to the belief that it is true.

The second view of information transfer through language we examine
can be called the direct perception model (DPM). According to the DPM, the
mechanisms allowing hearers to derive the literal meaning of an utterance are
subconscious and as direct as those underlying visual perception—that is, they
are not adequately modeled as inferences to what the utterance means. Accord-
ing to the DPM, when told that piece of information, you directly form the
belief of that piece of information without any Gricean reasoning about the
speaker’s intentions. A central prediction of the DPM is therefore that once the contents of communicative stimuli are grasped, they do not need to go through another assessment mechanism to get into the interpreters’ belief boxes. As far as the cognitive mechanisms underlying the retrieval of literal meaning go, you believe everything that you are told.

It is important to note that the DPM does not imply the absence of epistemic barrier filtering out the communicated contents that can get into the belief box. The crucial difference between the DPM and the IM is that for the IM the gap between interpretation and belief is presupposed by the very cognitive mechanisms assumed to underpin belief acquisition from linguistic stimuli. For the DPM, by contrast, any epistemic filtering of hearsay information is independent from the interpretation process.

**EXPERIMENTAL EVIDENCE**

The DPM has a respectable, if somewhat minority, tradition in contemporary philosophy (see, for instance Burge, 1993; Millikan, 2004, 2005; Recanati, 2002), but there also exists an experimental side to the debate. An important set of empirical evidence that may be invoked in favor of the DPM comes from experiments by Gilbert and colleagues (Gilbert, Krull, & Malone, 1990; Gilbert, Tafarodi, & Malone, 1993). The aim of these experiments was to evaluate two competing models of belief acquisition. According to the first, dubbed “Cartesian” by Gilbert, validation of a statement is never concomitant with its comprehension: A filter—one kind of internal deliberation—is necessary before endorsing communicated content as a valid description of the world. The IM is very similar, at least in spirit, to this Cartesian model. The contrasting, Spinozean, model predicts that any belief is acquired automatically; if there is assessment, it takes place after information has been integrated within the belief box. The DPM is fully compatible with the Spinozean view.

A central prediction of the Spinozean model is that belief rejection thus operates post hoc and is an effortful process. By contrast, the Cartesian model predicts that subjects can prevent a proposition from getting within the belief box and that, if no deliberation takes place, being told the piece of information will never result in the belief of that piece of information. Because a central prediction of the DPM is that acquiring information through the communicative channel is not intrinsically mediated by epistemic filtering, results favoring the Cartesian model would disprove the DPM.

In a paradigmatic study confronting these two models (Gilbert et al., 1990, Study 1), participants read statements about the meaning of words in an unknown language (Hopi), such as, for example, *monishna is a star*. After an 8-second presentation of each statement, the words *true* or *false* appeared on the screen. For some of the statements, participants had to perform a secondary task (i.e., responding to a tone), which mobilized additional cognitive resources. In a subsequent, testing phase, participants were presented with a list of statements (some of which were true and the others false in view of the earlier statements)...
and asked to identify their truth-value. The variable of primary interest is the rate of correct recognition as a function of the truth of the statements and the presence of an interfering task in the learning phase. When identifying statements that had not been interrupted during the learning phase, people did not make more errors on false than on true statements. This secondary task did not influence performance on true statements either but, crucially, it led to more errors on false statements, which were more often identified as true than in the absence of such interruption. The latter result suggests that cognitive resources are necessary to correct the default endorsement of the sentence’s content as true. By taxing these resources, the secondary task prevents such correction to take place. Such a pattern of findings cannot be properly explained if a Cartesian filter operates between the encoding of the statement and a (hypothetical) subsequent judgment of truthfulness.

In this initial set of studies, it is unclear whether participants’ judgments reflected their memory of the information presented in the learning phase or actual belief in such information. In their second set of studies, Gilbert et al. measured more reliable indicators of actual beliefs. For example, in one experiment, they presented information about a defendant in the context of a criminal case, such that some bits of this material were explicitly tagged as false. Furthermore, the information thus communicated was either exonerating or aggravating for the defendant. In addition, half of the participants had to simultaneously perform a secondary task, whereas the other half was not interrupted. In the interrupted condition, when aggravating information was false, participants proposed a more severe penalty than when exonerating information was false. This was not the case when they were not interrupted. Thus, when interrupted, people assimilated the false information and failed to correct it. Taken together, these findings suggest only that people encode false information as true when their cognitive resources are taxed but that they also act on it.

As is suggested by the title of the 1993 paper “You can’t not believe everything you are told,” Gilbert et al. seem to consider that people are incapable of suspending belief. As hearers, we would be capable of considering statements to be false only after having previously endorsed them. This—extreme—position has been challenged by Hasson et al. (2005). A straightforward concern about Gilbert’s experiments is that, from the participant’s point of view, false statements used in the experiment (such as the ones about the meaning of Hopi words) are uninformative. Accordingly, knowing that they are false (e.g., that it is not true that moshina means star in Hopi) has no informative value for the participant. But of course, being told that a statement is false may prove informative per se. For instance, when a statement like Tom is generous is tagged as false, the participant may directly encode the information Tom is greedy. That is, it is not the statement itself that would be encoded—as a strict application of the Spinozean model would suggest—but rather, an inference drawn from it being tagged as false. If this happens, the sentence Tom is generous should be readily identified as false in the second phase of the experiment even when cognitive resources are depleted. To test this hypothesis, Hasson et al.
replicated the first experiment of Gilbert et al. but manipulated the informa-
tiveness of false statements. They obtained the same results as Gilbert et al. when
the false statements were uninformative (viz., the distracting task during the
exposure phase led to consider false statements as true). However, when the
false statements were informative, recognition performance was not altered by
the secondary task. This result allowed Hasson et al. to conclude that recalling
the truth-value—be it true or false—of informative statements is not affected
by parallel cognitive overload.

Drawing on research in psycholinguistics, Richter et al. (2009) have posed
another challenge to Gilbert’s claims. They suggest that people routinely (i.e.,
in the absence of an explicit goal) and effortlessly rely on validation processes
when comprehending sentences. Contrary to both the Spinozean and the Car-
tesian view, it would be impossible to divorce validation from comprehen-
sion. Validation, they argue, is grounded in background assumptions related to the
topic at hand. Because the material used in Gilbert’s experiments consisted of
statements about unknown topics, lack of any relevant background may have
therefore prevented the participants from performing routine validation. Rich-
ter et al. reproduced the Hopi experiment in Gilbert et al., but with statements
half of which were perceived to be true or false with a high certainty in a pilot
study, whereas the truth-value of the other half of the stimuli was seen as uncer-
tain, in the same pilot study. Richter et al. replicated Gilbert’s pattern for the
latter group of stimuli (interruption during the learning phase yielded weaker
recognition performance for false, but not true, statements). However, they also
found that for statements with strong background beliefs (be they true or false),
interruption did not affect performance.

In a more direct test of the presence of an implicit validation, Richter et al.
(Experiments 3 and 4) relied on an epistemic Stroop paradigm, in which sub-
jects had to evaluate (by rapidly clicking on one of two buttons) the spelling of
words belonging to sentences that were either consistent or inconsistent with
strong background beliefs. Richter et al. assume that, if belief validation is
routinely triggered, it should interfere with orthographical judgments. Accord-
ingly, people should experience difficulties both with approving the spelling
of words within statements that contradict strong background beliefs and with
disapproving the spelling of words within statements that conform to such
beliefs. These predictions were corroborated on measures of error rates and
reaction times. Participants made fewer errors and (in Experiment 4) took less
time to respond when words within true sentences were correctly spelled and
when words within false sentences were incorrectly spelled than in the two
incongruent conditions.

These and Hasson’s findings suggest that the radical version of the Spino-
zean view is hardly tenable. In some cases false information is not rejected a
posteriori, but filtered straight at the entrance of the belief box. However, the
presence of a filter in some context does not imply that it is a necessary condi-
tion for acquiring hearsay beliefs. If anything, these studies suggest boundary
conditions for the operation of a filter. But this falls short of invalidating the
DPM. What remains uncontroversial about Gilbert’s results is that epistemic filtering is not inherent in communication. In certain circumstances, you believe directly what you hear (or read). What the IM predicts is that there is no believing of communicated meaning without epistemic check—and this is not the case.

This line of thought receives support from data on the ontogenesis of epistemic vigilance. The capacity to assess the reliability of a communicator is quite a precocious one. From the age of four, children are capable to discriminate between a reliable and an unreliable puppet (Clément, Koenig, & Harris, 2004). Likewise, four-year-olds tend to distrust a puppet characterized as a liar (Mascaro & Sperber, 2009). Yet, this capacity is by no means part and parcel of the way communicative behavior is processed (as the IM would have it). To begin with, the same studies also revealed that at the age of three, children fail to adopt such selective trust. Furthermore, Vanderbilt, Liu, and Heyman (2011) show that explicitly identifying an adult as an unreliable deceiver in three consecutive communicative exchanges does not prevent four-year-olds from trusting the information communicated by this same person right afterward.

The point is not that young children are blindly gullible. They are not. For instance, children below five years tend to privilege first-hand, perceptual information over verbal claims made by an adult (Robinson, Mitchell, & Nye, 1995). However, there is no evidence that acquisition of such skepticism strategies is inherent in the development of the capacity to interpret communicative stimuli.

**THE EVOLUTION OF EPISTEMIC VIGILANCE: SOME SPECULATIONS**

A major finding of the neo-Darwinian paradigm has been that cooperative behavior proves evolutionarily rewarding. It is also widely accepted that such strategies must encompass a mechanism aimed at the exclusion of noncooperative cheaters from interaction (e.g., Axelrod, 1984; Axelrod & Hamilton, 1981; Dawkins, 1986; Kitcher, 1993). It is also widely accepted that cooperative behavior can be found, in some form or another, all over the animal kingdom, and especially among great apes (see e.g., De Waal, 2006). In this light, it is not too risky a conjecture that human communication emerged among groups whose members could already reasonably expect each other to be helpful, and from which cheaters were ostracized.

Let us indulge now in some evolutionary speculation. Imagine two hypothetical groups: the direct perceivers and the inferentialists. Members of both groups can communicate to share information, but while the former acquire information from speech in accordance with the DPM, the latter have to go through the inferential strategy posited by the IM. Assume, furthermore, that members of each group pass their interpretative strategy to their offspring.

It seems obvious that within the kind of cooperative niche just described, direct perceivers would be clearly advantaged over inferentialists. Because any communicated content is directly added to the direct perceivers’ belief boxes,
they would acquire information much faster and in a more effortless way than inferentialists, who need, every time, to go through assessment and consistency checking before validating what they have been told. Provided that communicators are benevolent and competent, communicated information will be accurate often enough to privilege direct perceivers, because accurate information acquisition would then mobilize fewer resources than those needed by the inferentialists—resources which can thus be profitably allocated to another task. In such an environment, direct perception through speech would be evolutionary stable. Subpopulations endowed with it would rapidly take over individuals who cannot communicate altogether—and also over a hypothetical inferentialist, Cartesian subgroup.

However, being a direct perceiver makes interactions with unreliable speakers very costly. The DPM predicts there is a great risk that the contents of misleading statements will automatically get into the direct perceiver’s belief box, so that an exclusion or assessment process will be necessary (at the risk of letting a false belief influence his or her decisions). But such processes take time and cognitive energy that could be better employed. Therefore, when the quantity of false utterances exceeds a certain threshold, direct perceivers become disadvantaged, and inferentialism starts to look like a more promising strategy.

Even though it is very plausible to assume that these hypothetical ancestors of ours evolved in small groups, bound by kinship and in-group cooperation links, direct perceivers were exposed to two sources of misinformation. First, encounters with out-group deceivers remained possible; second, even within their own, reliable group, direct perceivers must have had to count with benevolent but mistaken communicators. That is, direct perceivers were advantaged only if they had evolved independent means to overcome deception and unintentional misinformation.

Regarding the first type of risk, the best evolutionary strategy is clearly to supplement the mechanism that ensures quick and effortless integration of communicated information with efficient and automatic epistemic filters that activate vigilance with respect to certain speakers (or in certain conditions). Such an evolutionary scenario entails that epistemic vigilance is not of one piece—it is a patchwork of adaptive strategies, shaped by heterogeneous environmental pressures. The supplementation of interpretive mechanisms with epistemic filtering is a classic example of what Krebs and Dawkins (1984) call the evolutionary arms race, concomitant with the development of communication systems. The aptitude to inform begets misinformation and deception, which in turn increases the adaptability of filtering mechanisms. Such an adaptation would be hard to explain in the absence of environmental pressure to control the ingress of communicated information within the belief box. It is precisely because the cognitive processes that allow us to interpret utterances as conveying informative contents do not come with an inherent epistemic safeguard that such filtering mechanisms have been selected.

To be sure, the domain-specific epistemic filters still do not shield direct perceivers from misinformation from their benevolent fellows (viz., from being
misinformed in a kind of cooperative situation where no specific vigilance should be triggered). Recall that misinformation has a higher cost for direct perceivers than for inferentialists. Therefore, the result is that, ceteris paribus, it is better for direct perceivers to avoid interaction with unreliable speakers altogether. Communicating false information, intentionally or not, should be seen as a noncooperative behavior, worthy of ostracism from the group.

At this point, it may be objected that this last feature of our evolutionary scenario renders the DPM evolutionarily implausible after all, for under such a view, speaking seems to be quite a risky business. If saying something false is assimilated to noncooperative behavior, punishable by exclusion from interaction—with all the dramatic consequences this entails—the most evolutionarily stable strategy would be to remain silent unless one is absolutely certain about the truth of her utterance (for a related discussion, see Hurford, 2007, pp. 276–277).

However, this risk may precisely be at the cornerstone of the emergence of human communication. Indeed, recent theories predict that, being prone to exchanging information—with the risk of being mistaken—has a higher evolutionary value than remaining silent in most cases, and thus avoiding any risk of being treated as a cheater. One such model is defended by Dessalles (1998), according to whom by providing reliable (and relevant) information, speakers seek to increase their social prestige, and hence, their reproductive success. Another compatible position is Miller (2000, Chapter 10), who argues that the human propensity to communicate is explained in great part by sexual selection: Verbal display raises the chances for mating.

Dessalles's and Miller's rationales assume that communicating has a certain cost; otherwise, no prestige would be attached to such behavior. They both appeal to the handicap principle (Zahavi & Zahavi, 1997). Some traits that constitute a prima facie handicap provide the organism that displays them with a higher chance of reproduction. Roughly, the idea is that by exhibiting a handicap, the individual demonstrates ability to survive despite this handicap, which, in turn indicates a high degree of fitness. For instance, male bowerbirds build elaborate bowers of twigs whose only function is to serve as a stage for courtship displays. The more adorned and the bigger this stage is, the more energy consuming is its construction, but also the higher is the likelihood for the builder to be chosen by a female. The female evolved this preference because bowers constitute a reliable indication of fitness. A male that can waste time and energy to build a big and elaborate bower, at the expense of looking for food, is likely to be more fitted to the environment than the one that cannot afford such a costly behavior.

The crucial component of the handicap principle is that to be a reliable indicator of fitness, the handicap (e.g., building a complex bower) must be hard to fake—otherwise any individual, not only the fittest one, could afford it. Therefore, if, as Dessalles (1998) and Miller (2000) claim, providing information through speech benefits the speaker by increasing social prestige and sexual attractiveness, the speaker's task should not be easy.
Now, in communities of direct perceivers one such risk is obvious: Unreliable speakers run the danger of being excluded from interaction as cheaters. In other words, the DPM predicts the emergence of the policing mechanism that theories of language evolution based on the handicap principle need to get off the ground.

Another important issue, which we can only mention in passing here, is that the IM requires viewing the appearance of linguistic communication as a twofold and simultaneous evolutionary emergence of a new channel and a new way of information acquisition. Accordingly, it needs to posit a double—and simultaneous—environmental pressure to explain the emergence of linguistic communication: one factor that explains the selection of complex communicative behaviors and another that selects for an inferential acquisition strategy, with an inherent gap between understanding and believing. The DPM, by contrast, views language as a new channel to feed information into the belief box in exactly the same way as perception. This, in itself, makes the DPM more plausible from a phylogenetic point of view. To be sure, hearsay beliefs are not as reliable as perception-based ones. But, as we have argued in this section, linguistic communication proves maximally efficient when appended with domain specific epistemic filters.

IMPACT ON SOCIAL PSYCHOLOGY

Before concluding this paper, we would like to allude to some central issues in social psychology that may benefit from a critical assessment of the IM.

A program of research that coheres with the DPM relates to the saying-is-believing effect (Higgins & Rholes, 1978). In this paradigm, participants read a description of a person (the target). This description is crafted in such a way that the statements it comprises can be interpreted as reflecting either relatively desirable or undesirable traits. Participants are then asked to describe the target to an audience (who is supposedly already acquainted with this person) to allow this audience to identify the target. Crucially, speakers are informed of the audience’s attitude, which can be either favorable or unfavorable to the target. It is not surprising that communicators tend to describe the target in a more flattering light when the audience is said to hold a positive, rather than a negative, attitude. More interesting, however, is the fact that when communicators’ impressions about the target are later probed, it turns out that their own memory is biased as well and in the same direction. This does not happen in a control condition in which speakers are exposed to the audience’s attitude but do not have to communicate. Decades of research on this phenomenon have led to consider it as driven by a desire to establish a shared understanding of the target with the audience (shared reality; Echterhoff, Higgins, & Levine, 2009). This shared reality is contingent on trust in the audience, and especially in its capacity to form an accurate opinion of the target. When this trust is present, there seems to be no barrier to simply incorporating the audience’s attitude into one’s own. In other words, when trusting an audience, the latter’s view of the
target seems to translate into believing that this view is correct. However, this process seems to be mediated by the active construction of an understanding of this target through communication. In other words, everything happens as if once verbalized, descriptions of the target influenced by the audience’s attitude come to be perceived as true. This is a very reasonable prediction that can be made from the DPM; presumably, no filter is activated for one’s own statements, which, therefore, automatically end up in the speaker’s belief box.

Another phenomenon that speaks to the DPM is belief perseverance (Anderson, Lepper, & Ross, 1980). In research on this topic, people are presented with facts which are later discredited. Yet, people keep believing in these facts to a greater extent than control subjects who have not been exposed to them. Obviously, this effect is not necessarily incompatible with an inferential model because what is at stake here is not belief validation but how one can come to disbelieve information that has been initially accepted as true. However, the difficulty in undoing such beliefs suggests that they are incorporated relatively automatically, without conscious control (as a Spinozian model would rather suggest).

GRICEAN INFLUENCES IN SOCIAL PSYCHOLOGY

Grice’s theory of meaning is probably the most important influence from pragmatics on social psychology (cf. Wänke, 2007; for an illustration, see Chapter 11, this volume). Especially, it has been used to (re)interpret research on a variety of cognitive biases. As we have argued in favor of the DPM against the limitations of the IM, which was deeply inspired by Grice’s theory, the question naturally arises: Doesn’t this work precisely show the superiority of the IM?

To address this question, we shall consider one bias that has been extensively studied under a Gricean light: base rate neglect (Kahneman & Tversky, 1973). Let us first illustrate how this bias has been experimentally demonstrated. Participants receive information regarding a person displaying traits typical of a social category A (e.g., engineers) or of a stereotypically opposite category B (e.g., lawyers). For example, Jack loves mathematical puzzles would be construed as more typical of engineers than lawyers. Participants are asked to estimate the likelihood that the target person belongs to category A. In addition, the target is presented as drawn from a sample containing either a majority of members of social category A (e.g., 70% engineers versus 30% lawyers) or only a minority (e.g., 30% engineers versus 70% lawyers). Typically, people’s estimates of the target’s membership are little affected by this statistical information. Rather, it is the stereotypicality of the target that explains most of the variance. Tversky and Kahneman (1974) famously analyzed such biases as the effects of simple heuristics (e.g., representativeness) by opposition to more elaborate and rational calculations. However, Schwarz (1994) pointed out that the conversational context in which these biases arose had been neglected. Conversational moves are governed by certain expectations—among which the ones Grice identified as “conversational maxims”—and these are often implicitly violated
in experimental settings. Typically, hearers expect speakers’ contribution to be relevant, but in the base-rate paradigm described above, a central part of the experimenter’s contribution (viz., target’s stereotypical traits) should not be relevant if participants were to behave rationally. This effect can be eliminated, or attenuated, when conversational expectations are neutralized. Thus Schwarz, Strack, Hilton, and Naderer (1991) showed that when information about the target was presented as selected by a computer, rather than a psychologist, participants’ estimation of the target’s membership is more influenced by statistical information. This is so because we usually do not expect computers to be sophisticated enough to select all, and only, relevant information about the psychological profile of a person; hence, not every bit of information is automatically taken to be relevant. This kind of Gricean explanation has been conclusively applied to a great number of other experimental paradigms (for reviews, see Holtgraves, 2010; Schwarz, 1994).

Although Gricean in spirit, this analysis of cognitive biases is actually consistent with our rejection of the IM. In the foregoing, we claimed that epistemic vigilance is not inherent in our capacity to retrieve information from speech. What Schwarz’s results reveal is that how communicated information is integrated, and hence influences other beliefs and decisions, depends on the context (e.g., the nature of the problem, the identity of the speaker, etc.). That some information selected by a computer is not taken into account does not show that when the same information is provided by a psychologist, and is used within a decision process, it necessarily undergoes inferential epistemic filtering. The DPM does not deny the influence of contextual factors such as expectations of relevance and of cooperation (and therefore allows the applications of Grice’s maxims as is typically done in social psychology); what it does reject is the view that such factors are necessarily mediated by the attribution of complex, communicative intentions to the speakers.

**CONCLUSIONS**

Social psychologists tend to consider comprehension and validation as two independent processes and are actually much more interested in validation than in comprehension, which is best left to (psycho)linguists. What we hope to have shown above is that it is impossible to remain agnostic about the cognitive processes underlying utterance interpretation if one undertakes to explain how communication-based beliefs are formed. Grice’s work helped researchers to realize that communication is an intersubjective activity, whose many aspects are influenced by an expectation of cooperativeness. However, it is a mistake to adopt the reconstruction of speaker’s meaning in terms of intention attribution, operated by Grice, as a monolithic model of language comprehension. Such a theoretical choice forces one to posit that no belief can be drawn from linguistic stimuli without having gone through an epistemic filter. We have argued that this consequence is hard to accommodate with available experimental data.
Much more plausible, both from an empirical and an evolutionary point of view, is the direct perception through language model. Contrary to the extreme position some proponents of this model might advocate, it does not compel us to assume that no information can be rejected without having been previously held as true. The cognitive equipment that allows us to acquire hearsay beliefs is supplemented with a variety of epistemic filters, which may be easily and automatically activated under certain conditions. A promising direction for future research is to understand better the typology of these filters. Meanwhile, it seems fair to conclude that epistemic vigilance is not inherent in our capacity to understand others’ statements.\(^7\)

**ACKNOWLEDGMENTS**

We would like to thank Philippe De Brabanter, Marc Dominicy, Mark Jary, the members of the Linguistics reading workshop, and the members of the ULB social psychology unit for their feedback on an earlier version of this chapter.

**NOTES**

1. An exception is the work on communicational grounding (Clark 1996), which considers how people elaborate a common ground to pursue cooperative projects. However, in this kind of research, the focus is not on an interpersonal level of analysis, viz. on how people manage to incorporate new knowledge through communicational behavior. The cognitive underpinnings of this incorporation, by contrast, are hardly considered.

2. To be sure, with some optical illusions—those that you recognize as such—you do not believe what you see. But note that conscious effort is needed: You eliminate the belief you acquired. Moreover, understanding that what you see is an illusion requires explicit training or at least, a time-consuming, detailed examination of the stimulus from different points of view and/or through different perceptual modalities.

3. Children below four are selective about informational medium (visual or tactile perception, or hearsay), and in cases where two sources provide contradicting information, they favor the most reliable one; however, they have difficulties in reporting correctly the source of their beliefs, which suggests that once a belief is acquired no trace of its provenance subsists (Gopnik & Graf, 1988; Mitchell, Robinson, Nye, & Isaacs, 1997; Whitcombe & Robinson, 2000; for a related discussion, see Millikan, 2005, p. 209–210).

4. Of course, the hypothesis here is not that two such groups coexisted at some point; the idea is rather to speculate how each group would have fared.

5. Handicaps may also serve to deter predators; see Zahavi and Zahavi (1997, Chapter 1).

6. Furthermore, the opposite pattern of results was obtained when the problem was presented as statistical, and not psychological (presumably, computers are better with statistical tasks than psychologists).

7. Origgi and Sperber (2000) and Sperber et al. (2010) invoke massive contextual dependence of the literal meaning to discard the DPM and argue that verbal understanding necessarily conforms to IM. However, there are very strong reasons, mostly from typical and atypical language and cognitive development, to dismiss the idea that all such pragmatic processes require the complex mind reading posited by IM (see, Breheny, 2006; Kissine, 2012, 2013, Chapters 3 and 5; Recanati, 2002).
REFERENCES


