

## SEVEN BILLION SOLITUDES

### Meaning and Truth in Personal Relations

Ronald de Sousa

<http://www.chass.utoronto.ca/~sousa>

*La parole a été donnée à l'homme pour dissimuler sa pensée*

Charles Maurice de Talleyrand

*Sieh dir die Liebenden an,*

*Wenn erst das Bekennen begann*

*Wie bald sie lügen*

R.M. Rilke

*If anyone explains Quantum Logic to me so that I can understand it, I know they haven't explained it right.*

Anon.

*The more I think about language, the more it amazes me that people ever understand each other at all!*

Kurt Gödel

## ABSTRACT

### **Seven Billion Solitudes: Why it's a wonder that we understand one another**

For a message to get through, it takes two. But no two speakers speak the same *idiolect*. At every one of many levels of analysis, a specific potential for misunderstanding arises. One level opens up into the consequences of communication: we might call it the hyper-pragmatic level. It is indefinitely extensible, and it touches, in particular, on emotional resonances, forming another multifarious source of potential disruptions of communication. The hyper-pragmatic level comprises a bewildering variety of nonstandard effects of language on cognition, such as those involving "priming", which exploits direct and unconscious effects of words or images on attitudes and behaviour. In addition, in the context of personal relations, further mechanisms generate both the motivation and the means for both innocent misunderstanding and intentional manipulation. Some of the mechanisms involved are traceable to the very nature of emotions as determinants of perceptual and attentional salience; and to their origins in paradigm scenarios shaped by both temperamental and cultural factors. When all these potentially disruptive influences on communication are taken into account, it is a wonder that we ever understand each other.

## RÉSUMÉ

### **Sept Milliards de Solitudes: Pourquoi il est étonnant qu'on se comprenne.**

Pour qu'un message passe, il faut être deux, qui partagent la même langue. Mais il n'y a pas deux idiolectes semblables. À chaque niveau de l'analyse du langage, il s'élève de nouvelles occasions de malentendu. Au delà du vocabulaire, de la syntaxe et de la sémantique, on débouche sur les multiples formes de la communication: on peut parler d'un niveau hyper-pragmatique. Ce dernier domaine est indéfiniment extensible ; il touche en particulier aux associations émotionnelles, qui fournissent en elles-mêmes toute une foule de répercussions possibles sur la communication. Il y a aussi une étonnante variété de facteurs qui débordent du cadre du sens linguistique normal. Plusieurs de ces effets s'apparentent au "priming", c'est-à-dire à l'effet direct et inconscient que peuvent avoir les mots et les images sur le comportement et les attitudes. Dans le contexte des relations intimes, il est encore d'autres mécanismes qui engendrent à la fois les mobiles et les moyens de manipuler les autres à des buts plus ou moins innocents. Certains de ces mécanismes relèvent de la nature même des émotions, dans la mesure où celles-ci déterminent ce qui retient l'attention et sont formées par des expériences de vie uniques. Ils relèvent aussi de différences individuelles au niveau du tempérament et de la culture. Tout compte fait, il est plutôt surprenant qu'on s'entende jamais.

For a message to get through, it takes two — and a whole lot more. Both sender and receiver must speak the same language. But not necessarily the same *dialect*. What counts as the same language? What the books count as a language<sup>1</sup> involves a number of different levels of analysis. At each level, a specific potential for misunderstanding attends communication. But the level that opens up into the consequences of communication, what we might call the hyper-pragmatic level, is indefinitely extensible. In particular, emotions form another multifarious source of potential disruptions of communication. That level comprises a bewildering variety of what I shall call nonstandard effects of language on cognition. In addition, in the context of personal relations, further mechanisms generate both the motivation and the means for communication to break down, either by causing innocent misunderstanding or by fostering manipulation, hijacking the message, by one or the other of the interlocutors. Some of the mechanisms involved are traceable to the very nature of emotions as determinants of perceptual and attentional salience; to their origins in paradigm scenarios shaped by both temperamental and cultural factors; and to complex motivational factors. According to some recent findings, it appears that political disagreements are sometimes the more intractable because of personality differences traceable to specific genetic factors. And then there is the converse of the *Strangers on a Train* phenomenon: the tangled configuration of close relationships makes many people more ready to confide in strangers than intimates. A uninterested interlocutor, one comfortingly presumes, is a disinterested one. With our intimates, by contrast, emotions are engaged; more is at stake, and the truth is more dangerous. Being heard aright can have worse consequences than being misunderstood. And even when the truth is told, it is, sometimes, impossible to hear.

-----

<sup>1</sup>I've heard it said that among themselves linguists define a language as “a dialect with an army and a navy”. Although this characterization is not directly pertinent to my topic, its metaphorical potential might usefully be kept in mind in the sequel.

## 1. Levels of meaning

Before tackling the emotional hazards of communication, we need to take a brief look at some of the levels of language on which various disruptive factors can operate. Linguists distinguish phonology, morphology, syntax semantics and pragmatics. The characteristics of a person's speech at each level define that person's unique personal dialect, or *idiolect*, and the potential for manipulation or misunderstanding exists at every level. Here is a trivial experiment at the phonetic level which you can try in a supermarket checkout line. Just murmur "Tickle your arse with a feather". When asked "What?" in a tone of slight alarm, utter loudly and clearly, "Particularly nasty weather!" and watch anxious expressions switch to doubtful relief.

There is a presumption that if you and I speak the same language, we understand in the same way the words we use, the sentences we construct, and the speech acts we perform. To convey one's meaning, then, it should suffice to say what you mean, in sentences constructively structured, using words in standard meanings, and in reasonable conformity with Grice's rules for "conversational implicatures." Grice's "Doubt or Denial" condition, for example, dictates that when one utters a declarative sentence, at least in the mode of factual assertion as opposed to idle greeting, our choice of words normally presupposes that the statement made might have been doubted or denied. Hence a tautology such as "boys will be boys" gets interpreted in such a way as no longer to be tautologous—conveying, perhaps, that young males are apt to behave badly. Another conversational implicature is that assertions are believed by the speaker. In this way Grice explains what is anomalous about Moore's paradox "P, but I don't believe P": although it is not a contradiction to

suppose that I believe a falsehood, my assertion of both conjuncts, insofar as it *implicates* my believing what I say, would imply a contradiction (Grice 1989).

### *Objective and Subjective*

No less important, and less often stressed, is the condition that the hearer be capable of understanding the message in a manner suitably related to the speaker's intentions. Whether that suitable relation is one of identity, and if so how it is to be assessed, and if not, whether it admits of degrees—these are all questions fraught with waiting traps. As illustrated by my epigraph on the explanation of Quantum Mechanics, there is an objective and a subjective side to each successful communication. Objectively, there are standards of truth and correctness (at least so I shall take for granted here). But unless the subject is subjectively capable of hearing and understanding what is said, that truth cannot be conveyed, however accurate its formulation. In a conversation about some abstruse domain of physics, understanding rests on the accessibility to the hearer of a background of mathematical and scientific knowledge. In the case of personal relations, that capacity rests on background assumptions of a different sort. It is influenced by personality, temperament, ideology, and microculture. For example: with the current reliance on hurried written communication taking the place of oral exchanges, we have become very familiar—often to our cost—with the importance of intonations and inflections that express seriousness, irony, or other attitudes. When these are missed by one who merely looks at a written text, major misunderstandings may result. Emoticons are dragged in, in the attempt to remedy the situation; but they don't slice very thin. Neither are problems of intonation the only ones that interfere with our capacity to understand one another. There are differences of tone or diction that very difficult to formalize but nonetheless very easy to identify—but

only by those in the know. Adam Gopnik gives a nice illustration of this in a recent article in the *New Yorker*, commenting on some of the difficulties encountered by Artificial Intelligence. About the expressive power of a certain 11-year old American girls' dialect, he writes:

What sounds to the outsider limited and repetitive is to the knowing listener nuanced as Henry James. When one eleven-year old girl says to another eleven-year-old girl, "So then, like, the teacher got all, like, all of you, I guess, are, like, going to have to do a, like, I don't know, a makeup test. So! Like, *yeah*," she means: " The teacher becoming heated (that's why he "*got*, like," rather than "*said*, like") announced, in effect, that many of us (I suppose, at a first approximation, all) will, at some point in, as it were, the near future, have to take what actually amounts to, when all is said and done, a secondary makeup test. I have indignant feelings about this—as who among us would not?—but I recognize their essential futility." All of this is completely clear to the knowing listener, but it has been impossible, so far, to teach a machine to, you know, like, really, like, *get it*." (Gopnik 2011).

Without Gopnik's help (or that of an articulate member of the relevant speech community) the full subtleties to which he draws attention might remain inaccessible.

In what follows I sketch four broad classes of facts that generate misunderstandings in general. One results from the way in which our idiolects are learned. Another lies in the very similar ways in which each of us acquires our emotional idiolects, that is, our unique repertoire of emotions. A third class relates specifically to the rhetorical aspects of communication. And finally, I will suggest,

there are factors that may aggravate these problems when the communication involved pertains specifically to personal and intimate relations.

## 2. Paradigm Scenarios: Emotions as codeterminants of salience

We seldom retain an episodic memory of the specific occasion when we first encountered any particular word. Psychologists distinguish "episodic" from "semantic memory": the latter is so called because it is precisely the kind we have for meanings, but it applies to everything we know without associating the knowledge with a particular dated episode (Tulving 1972). We know how to use a word without remembering where or how we came to know it. Nevertheless, we can speculate that the specific circumstances in which we first learned to speak of "airplanes", "apricot", or "assignation" influence the connotation, if not the Fregean sense and reference of those words.<sup>2</sup> I have suggested elsewhere that a similar process underlies our emotional repertoire, which is constituted by "paradigm scenarios" (de Sousa 1987).

The idea of a paradigm scenario is related to the Freudian idea at the heart of the concept of transference (Freud 1915). The quality of every emotion is rooted in a dramatic, situation or episode type, associated with a characteristic feel (in the broad sense, not in some specific sense limited to bodily feelings), and also with "objects" in a confusing diversity of senses of that word. These typically (but not invariably) include a *target* (some particular object that the current emotion is

-----  
<sup>2</sup>(Frege 1980) distinguishes two objective properties of meaningful units: reference (the actual object or class of objects designated by an expression), and sense (the set of specifications that allow us to zero in on the reference). Both are contrasted with "colouring and shading", which "are not objective, and must be evoked by each hearer or reader according to the hints of the poet or the speaker." Subjective "colouring and shading", or connotations, are not pertinent to the determination of semantic truth value, but are of crucial importance in the context of both rhetoric and therapy.

directed at), a *motivating aspect* (that characteristic in virtue of which the target has triggered the emotion), and *motivational aim* (what the emotion moves me to do, by way of immediate expression or long-term aim). Once acquired (often in early childhood), the script for the drama or scenario specific to a given emotion is played out by agents cast in certain determinate roles. The hypothesis that forms my starting point is that these paradigm scenarios define the meaning for us of the emotions of which they constitute the prototypes; but in adult life the target is different; the motivating aspect may be partly imagined; and the motivational aim, insofar as it is matched to the old situation, might be irrelevant to the current one.

On this view, human emotions are not those simple reactions on the basis of which we attribute emotions to infants or animals—tears of rage, smiles of joy. Those are only the raw materials of full-fledged emotions. Nor do I refer to the “instincts” associated with basic bodily needs associated with food, safety, dominance and sex. Human emotions could not arise without innate predispositions to such reactions, but they are more complex, learned and structured patterns of response to experienced situational configurations. A standard example: a child desires the gaze and touch of the mother; the mother’s attention is diverted to the other parent or another child; there is screaming or whining—and the whole performance is called *jealousy*. When its name becomes known, it will have already crystallized into a recognizable pattern. It will therefore seem natural. ‘Jealousy’ will seem to be the name for some real, private feeling, quite independent of any social facts. It will have acquired the structure of a story, with a beginning, a middle and an end. But despite the common name, the original course of the story will have been somewhat different in every case, depending both on the specific social context in which it was learned and on individual temperament. Thus common names

and associated generalizations may conceal wide discrepancies. Emotional repertoires, like idiolects, can vary. The commonalities between them, suggested by the use of the same emotion words, will often not suffice to keep interlocutors "on the same page".

One important feature of the picture I am drawing is that paradigm scenarios mediate between the present and the past. That raises an inevitable question about the authenticity of any particular emotion and its appropriateness to the present target. A paradigm scenario challenges us to perceive the present experience with its present singular object, even while the emotional vocabulary it has provided to describe it is rooted in the past. This enhances emotional experience by endowing it with significance beyond the present particular instance. But that enhancement is an ambivalent one. For that additional significance may obfuscate the present reality as well as enrich it. We are usually unaware of the past's influence on our interpretation of the present. Lack of appropriateness to the current situation may remain concealed, exacerbating the potential for confusion or misunderstanding.

A striking problem stemming from the role of the common language in the shaping of individual consciousness is expressed in the following passage from Nietzsche's *Gay Science*:

consciousness does not really belong to man's individual existence but rather to his social or herd nature; . . . it has developed subtlety only insofar as this is required by social or herd utility. Consequently, given the best will in the world to understand ourselves

as individually as possible, ... our thoughts are continually . . . translated back into the perspective of the herd. (Nietzsche 1974, §354).<sup>3</sup>

Nietzsche seems to be arguing that our consciousness is constrained by what we can talk about, which in turn is limited by the language we share with others. But my concern here is different. From my point of view, the language of the herd would be a very desirable convenience: it would enable people actually to understand each other, even though it might limit what they can communicate. The problems that concern me here arise precisely insofar as Nietzsche is wrong. If my hypothesis about the way we learn our emotional idiolects is right, it offers a partial explanation for the fact that emotional intercourse between any two individuals is likely to be riddled with potential mismatches and misunderstandings. A lover's argument—using this designation metonymically, to refer to any conversation between intimates in which there are emotional issues and competing points of view at stake—is fraught with pitfalls, because each participant expects the other to act out a role from a script the other has not read. Even if the *speaker* confines herself to the categories sanctioned by the herd, there is no guarantee that she understands those in the same way as the hearer. Furthermore, by refining her expression of emotion in a way that is capable of conveying nuances beyond the crude categories of the "common herd", the general tendency will

---

<sup>3</sup> The whole passage in the original reads: "Mein Gedanke ist, wie man sieht: dass das Bewusstsein nicht eigentlich zur Individual-Existenz des Menschen gehört, vielmehr zu dem, was an ihm Gemeinschafts- und Heerden-Natur ist; dass es, wie daraus folgt, auch nur in Bezug auf Gemeinschafts- und Heerden-Nützlichkeit fein entwickelt ist, und dass folglich Jeder von uns, beim besten Willen, sich selbst so individuell wie möglich zu verstehen, "sich selbst zu kennen", doch immer nur gerade das Nicht-Individuelle an sich zum Bewusstsein bringen wird, sein "Durchschnittliches", - dass unser Gedanke selbst fortwährend durch den Charakter des Bewusstseins - durch den in ihm gebietenden "Genius der Gattung" - gleichsam majorisirt und in die Heerden-Perspektive zurück-übersetzt wird."

remain in the *hearer* to classify whatever is said into one of those standard categories; but there will be no guarantee that the hearer has not also embellished the categories, in incompatible ways. So the existence of common categories is not enough to prevent misunderstandings.

Milan Kundera's *Unbearable Lightness of Being* offers persuasive illustrations in a "Dictionary of Misunderstood Words". *Family, country, love, music, fidelity*, are words the associations of which for the characters are so remote from each other that they constitute guarantees of systematic misunderstanding:

[for Franz] fidelity deserved pride of place among the virtues: fidelity gave the unity to lives that would otherwise splinter into thousands of split-second impressions.....Franz assumed that she would be charmed by his ability to be faithful, that it would win her over. What he did not know was that Sabina was charmed more by betrayal than by fidelity. The word "fidelity" reminded her of her father, a small-town puritan, who spent his Sundays painting away at canvases of woodland sunsets and roses in vases.... After completing school she went off to Prague with the euphoric feeling that now at last she could betray her home. (Kundera 1984, 96).

### **3. The multifarious interaction of emotion and language**

In the present section, I describe half a dozen cases of what might best be called non-standard routes from words to emotional or behavioural change. Three are built around the experimental psychological paradigm of *priming*, which involves exposing a subject to a stimulus (a verbal one, in the cases that interest me here) too

briefly for the subject to be aware of it. The fourth case is built on the principles of operant conditioning, and appears to have therapeutic uses; but it could equally well function in destructive ways. A fifth consists in a direct challenge to the thesis of the informational encapsulation of direct perception. And the sixth further illustrates how social influence can affect our ability to make objective judgments, directly, without the mediation of any conscious thought for the need to conform. These vignettes illustrate some of the many ways in which apparently rational argument is threatened by factors extraneous to the conventional semantic content of language. After this brief catalogue, I shall say a little about how these sorts of obstacles to clear communication can be made use of in manipulative strategies of intimate contention.

*(i) Priming: Slow Walking.*

In a much discussed experiment, John Bargh and colleagues applied the priming technique to subjects who were asked to unscramble some sentences. In the experimental group, the sentences in question contained a number of words previously found to be associated with the stereotype of old age. These were: "*worried, Florida, old, lonely, grey, selfishly, careful, sentimental, wise, stubborn, courteous, bingo, withdraw, forgetful, retired, wrinkled, rigid, traditional, bitter, obedient, conservative, knits, dependent, ancient, helpless, gullible, cautious, and alone.*" (Bargh, Chen, and Burrows 1996, 236). The control group had a similar task, but without any of the stereotype words. Subjects were told that the experiment concerned verbal fluency. The actual experiment, however, began only after that partial debriefing:

Waiting until the participant had gathered all of his or her belongings, the experimenter told the participant that the elevator

was down the hall and thanked him or her for participating. Using a hidden stopwatch, a confederate of the experimenter, who was sitting in a chair apparently waiting to talk to a professor in a nearby office, recorded the amount of time in seconds that the participant spent walking a length of the corridor starting from the doorway of the experimental room and ending at a broad strip of silver carpet tape on the floor 9.75 m away. Participants in the elderly priming condition ( $M = 8.28$  s) had a slower walking speed compared to participants in the neutral priming condition ( $M = 7.30$ s)...  $p < .01$ , as predicted (Bargh, Chen, and Burrows 1996, 237).

A separate test confirmed that the participants had no idea of the influence of the priming words on their behaviour (though one out of nineteen subjects did seem to notice the occurrence of the stereotypical words). The influence of the priming words on their behaviour was therefore direct and entirely unconscious; it does not seem to have been mediated by any reasoning, inference, or articulable belief.

*(ii) Priming: The direct influence of unconscious feeling on moral judgment.*

In experiments conducted by Thalia Wheatley and Jonathan Haidt, arbitrary words ('take' and 'often') were associated with disgust in hypnotised subjects. In a later phase, without recalling the hypnotic episode, subjects read a simple anecdote, in one of two versions differing only in that one contained the words 'take' or 'often'; the other used paraphrases.

Each vignette described a moral transgression and was followed by two rating scales, one for rating "how morally wrong" and the second for rating "how disgusting" the behavior was. . . . After making their

ratings, participants were asked to briefly explain their morality ratings . . . . Participants found moral transgressions to be more disgusting when their hypnotic disgust word was embedded within the vignettes than when this word was absent. Moreover, the disgust word caused participants to rate transgressions as more morally wrong. Apparently, participants used their feelings of disgust (attached only to a word, not to the act in question) as information about the wrongness of the act. (Wheatley and Haidt 2005, 781).

Some of the vignettes included acts that might have elicited disgust in many subjects anyway, such as someone eating his dead dog, or cousins who had a sexual relationship. But another vignette was

added to provide a story with no violation of any kind: 'Dan is a student council representative at his school. This semester he is in charge of scheduling discussions about academic issues. He [tries to take/often picks] topics that appeal to both professors and students in order to stimulate discussion.' [A third version contained neither trigger word] Here those subjects whose versions of the story contained the trigger words 'take' or 'often' invented reasons to find fault with Dan, even though nothing in the story actually suggests that he did anything wrong. One subject just wrote, 'It just seems like he's up to something,' without specifying anything in particular. (ibid.)

What Wheatly and Haidt conclude from this experiment is that our moral judgements are at least in part driven by emotions unconnected to reason. My interest here is different: if we can imagine, taking the place of the posthypnotic suggestions in the original experiment, that two partners in a relationship have very different associations at the level of gut feeling with certain inherently neutral

concepts, then we can also expect that in an argument involving those concepts, each might be driven to different judgments of value, including moral value. And since these priming effects are unconscious, they cannot easily be corrected for.

In this regard they differ from the notorious Müller-Lyer illusion, the standard paradigm case of a visual perception that is impervious to correction from knowledge of the facts. I shall return to this illusion, and to the question of the “penetrability” by cognition of the deliverances of perceptual modules. But for the moment I want to draw attention to the fact that anyone who has once been made aware of the Müller-Lyer illusion will be suspicious of their own first impression based on purely visual cues, and can then stop to check. That will make them more or less immune to false beliefs triggered by the illusion, even if the illusion remains robust. In the case of emotional illusion, such immunity is not to be so easily achieved. We can test our sensory impression against objective measurement; but in most cases of emotional disagreement, such recourse to objective measurement will itself be contentious if not impossible. And one can expect the effect to be magnified in personal relations, on the ground that people are often inclined to be particularly *judgmental* in relation to those who are close to them.<sup>4</sup>

*(iii) Complex priming: Derren Brown's manipulations.*

Derren Brown has demonstrated, on British television episodes widely available on YouTube, priming techniques by which people's desires and behaviour can be manipulated in spectacular ways. Until later debriefed, subjects have no

---

<sup>4</sup>There is a compelling association between certain classes of emotions, generally shared by our close primate cousins, and each of five domains of moral concern (de Sousa 2009) (or six domains, in J. Haidt's updated version as expounded in (Haidt 2012). In this scheme disgust is specifically associated with the moral domain identified as pertaining to *purity*. But it has been found that moral judgments in domains unrelated to purity elicit micro-expressions that involve the neural and motor circuitry of disgust (Chapman, Kim, Susskind, et al. 2009).

awareness of his manipulations. An example of this technique is illustrated in a short video available at <http://tinyurl.com/2ujy4zu>.

In this experiment, Brown repeatedly plays a public announcement in a shopping mall. This announcement contains a number of words connoting gestures in an upward direction, such as *uplifting*, "details of our specials are *handily* near the *elevator*". After this priming he says "shoppers who wish to *reach up and grab* this exciting opportunity should do so NOW!". Thus prompted, an impressive proportion of shoppers do indeed reach up, and to their surprise notice everyone else holding their hands up. In other experiments, Brown gets people to do even more surprising things, including in one extreme case getting participants in a motivational seminar to commit armed robbery (Janson 2009).

The techniques used by Derren Brown can be regarded as particularly manipulative rhetorical devices. They are of a piece with techniques which, in an interview with Richard Dawkins to be found on the latter's website (<http://tinyurl.com/3pyc2g7>), he describes as the stock and trade of so-called "cold readings" by psychics. Among these classic techniques is the reliance of the speaker on drawing a link between, for example, a random proper name, and someone of interest to the hearer. When the psychics says "I'm getting a name, I'm getting the name Harry", the hearer is almost bound to know someone of that name, and in accordance with the basic presupposition that referring expressions typically refer to something or someone in particular, will assume that the Harry putatively mentioned by the psychic is indeed the very Harry they know. In this way the hearer is in fact doing all the work in creating the perception to which they are being subjected. A similar strategy consists in making statements that are vacuous disjunctions: "You're the sort of person that is usually very calm, but you can get

quite excited sometimes." Or one can even resort to outright contradictions: "I feel you embody an intriguing contrast between calm and a capacity for engagement." In either case, the listener can again be counted on to hear only the applicable disjunct or conjunct, particularly if the statement is framed by flattering qualifiers such as 'intriguing', 'remarkable' etc. which effectively encourage the interlocutor's cooperation. Another technique used by Brown is negative suggestion: in one case he gets a young woman to push a button that she believes will kill a kitten by electrocution. Brown frames this in such a way as to make it into an act of childish "naughtiness". The mood is prepared by various memory exercises intended to evoke the childhood thrill of resisting authority, and it is specifically triggered by Derren Brown's injunction NOT to kill the kitten. The negation is ignored, just as Freud claimed is characteristic of the ways of the Id. The power of negative suggestion seems to be boosted by the fact that a negative specification of a state of affairs provides nothing for the imagination to focus on, except the very state of affairs that is being negated: the injunction *Don't think of a white rabbit*, will guarantee that most people will think of a white rabbit. Clearly, while such mechanisms are particularly favoured by charlatans who use them for deliberate manipulation, they can be equally effective when the priming has happened by chance and is not shared among sender and receiver of a given message.

#### *(iv) Orectic Penetrability*

The Müller-Lyer illusion is a standard illustration of the "cognitive impenetrability" or "encapsulation" of the visual module: the agreed line, since Fodor's *Modularity of Mind*, is that beliefs do not influence sensation. That is clearly an unsatisfactory formulation. What is true is that one can be fully cognisant of the actual facts about the two lines in the Müller-Lyer illusion, while still seeing them as

having different lengths. But there are also many cases where a belief about what one is seeing entirely changes the subjective *Gestalt*. My favourite example of the cyclist wearing a Mexican hat seen from on top: Upon simply seeing a circle with two segments sticking out on either side, one may not see it as anything in particular. But on hearing the caption, "Mexican riding a bicycle", the sensation, the actual experience is changed. The McGurk effect may provide a comparable auditory illusion, although the influence comes from the sense of sight, not from verbal information<sup>5</sup>, but it seems obvious from unaided observation that a belief about what one is hearing may change an auditory experience. My first phonological example of a misunderstanding relied on something like this mechanism. As for taste, there may be some phenomena that are analogous to the Müller-Lyer illusion. I have heard it reported that according to Charles Spence, whose research explores synaesthetic phenomena, stale chips taste all right if they are chewed to the accompanying sound of crunchy crisp chewing. Similarly, experiments on the effect of belief about price on the pleasure taken in different wines seem conclusively to show that the effect in question is both direct and real at the level of brain function. The evidence is that in tasting identical wines while under the impression that one was worth \$10 while the other was a \$90 wine, subjects whose brains were monitored showed increasing activity in the pleasure centres, suggesting that their report of greater enjoyment did not need to be mediated by snobbery. Taste, it seems, was directly affected by cognition. (Plassmann, O'Doherty, Shiv, et al. 2008).

In a recent paper, Dustin Stokes has defended the thesis of *orectic penetrability*, namely that perception can be influenced internally and directly by

-----

<sup>5</sup> This may be experienced at <http://www.youtube.com/watch?v=G-IN8vWm3m0>

desire (Stokes 2012). He is interested only in cases where the influence is both internal and direct. Thus he would set aside many of the cases I have been concerned with, where the influence is indirect, including cases where our judgment of what we see is influenced as a consequence of a pattern of salience induced by an emotion. If I want to see you, I will be attending with particular intensity to anything that may be construed as a sign of your imminent arrival. If I am jealous, I will, like Othello, notice things I would not otherwise have noticed, and construe their significance in new ways. In that sense our perception is certainly influenced by desire, or at least by emotion, which typically includes desire. But these are neither pure nor clear cases of unmediated orrectic penetration of perception. Stokes is concerned to show that some cases clearly refute the thesis of the cognitive impenetrability of visual perception in its starkest form. His most compelling piece of evidence is a classic of psychological research over 60 years old. Jerome Bruner and Cecile Goodman asked children to estimate the relative sizes of two disc shaped objects. In both conditions one of the object was simply a projected circle of light; but in one condition the other object was a simple cardboard disc, while in the other condition it was a coin. The result was that the children significantly overestimated the size or proximity of the coins in relation to the neutral object. Equally significantly it was found that children of poor families overestimated the size of the coins to a greater extent than children of well-to-do families (Bruner and Goodman 1947). Assuming that coins are more desirable than cardboard discs, and that a modest amount of money is more intensely desired by someone who lacks it than by someone who does not, Stokes takes this to establish that there is indeed both direct and internal causal penetration of the perceptual mechanism by desire. This is "orectic penetrability".

As Stokes points out, this has a number of philosophically and practically significant consequences. Not the least of those is that it throws further doubt on the already shaky status of eyewitness reports in forensic contexts. For my purposes, however, the significance of the phenomenon of orrectic penetration lies in providing one more example of a mechanism whereby two individuals engaged in debating a matter of fact for which there is visual evidence may honestly disagree, without any conscious attempt at distortion, about the nature of the evidence that is supposedly "plain for all to see" .

*(v) Attention Bias Modification Treatment (ABMT)*

A new therapy based entirely on interaction with a computer screen has recently attracted attention. On the basis of standard Skinnerian learning theory, it is intended to minimize attention to anxiety-causing stimuli. In this way, "ABMT" (or "CBMT", for "Cognitive Bias Modification Therapy") aims to modify directly subcortical pathways that promote anxiety and are not affected by regular talk-based Cognitive Behaviour Therapy.

The basic idea is simple. The subject is looking at a screen, and set the task of locating a given probe stimulus, such as a colon in the case schematized by Figure 1. The subject is exposed for 500ms to simultaneous words, one of which is neutral while the other is disturbing. The target then consistently appears in the same area as the non-threatening stimulus. Subjects learn to attend to that area, away from the threatening stimulus. The authors write: "Compared with the explicit training techniques of CBT, attention perturbation might be more easily shaped by ABMT, with its use of repetitive, computer-based training methods targeting implicit, subcortical processes. As such, ABMT might represent a novel treatment that directly targets perturbed neural circuitry function." (Hatanama et al. 2010).

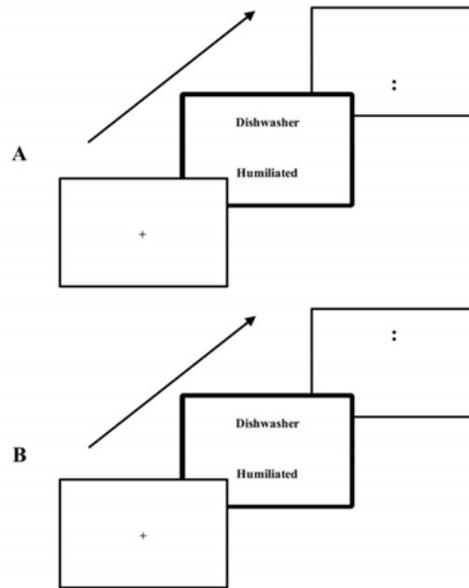


Figure 1: Set-up for ABT (from Hatanama et al., 2010, p. 983)

There is nothing remarkable about the fact that one can be trained to look away from the threatening stimulus. What is surprising is that the effect can generalize in such a way as to cause a significant improvement in the tendency of anxious subjects to focus on whatever tends to make them anxious. Some of the studies included in the meta-analysis were pictorial. But others involved only words, and in those cases what is remarkable is that it works at all. For the fact that the training generalizes to areas not directly involved in the training phase suggests that the actual stimulus property is the relatively high-level property of *being anxiety provoking*. For since that is the only thing shared by the experimental stimulus and real-life topics that trigger the subject's anxiety, training to the former could not otherwise affect the latter. So this presents another illustration of what I have been calling a non-standard way that words penetrate through to system one processes and influence emotion and behaviour.

*(vi) Social Influence: Berns's replication of Asch's Experiment.*

In some well-known experiments on social conformism, Solomon Asch had found that when asked to make a judgment of a visual quantity, some 40% of subjects went along with the false judgment of stooges posing as fellow subjects (Asch 1962). In a new variant on these experiments, using fMRI data, Berns and his colleagues found a "highly suggestive... lack of concomitant activity changes in more frontal areas," where one might have expected activity if the subject's judgment had resulted from a decision to override one's own judgment in favour of the majority response. (Berns et al. 2005, 7). The surprising aspect of their findings is that *no special cognitive activity was detected in the cortex of those who conformed to others' false opinion*. Instead, "the effects of social conformity are exerted on the very same brain regions that perform the task". In other words, far from being an inference required by the need to resolve cognitive dissonance ("So many others can't be wrong, I must revise my verdict"), the influence of others' judgments seems to act *directly* on perception. The distorting effect of conformity did not require any calculation of costs and benefits: it was those who saw and stood up for an independent truth who endured emotional cost. This finding is as intriguing as it is discouraging, for it suggests that the power of social conformism does not so much overcome perceptual judgment as bypass it altogether. When we think of it in relation to Stokes' discussion of orectic penetrability, we might interpret it as offering another example of the phenomenon exemplified in Bruner and Goodman's experiment: a desire to conform seems to influence perception directly, without going through any inference or judgment. On the other hand, there is nothing in particular that requires us to say that subjects are influenced by desire as such. Imitation is a very basic strategy of human learning (Henrich and Boyd 1998). It

could just as well be that subjects are being influenced by something best described as an *imitation module* that doesn't necessarily offer any grip to the attempt to decompose it into something in the nature either of desire or of judgment.

#### 4. Goals of conversation

The six mechanisms described in the previous section are relatively specific; but their role in lovers' arguments is likely to be randomly disruptive rather than targeted in conscious manipulation. The same is true of cultural variables and of individual temperamental differences. These—particularly in combination—are liable to impact on personal relationships in unpredictable ways. Just one example: in a narrative account of her “double life” as a Chinese immigrant to Australia, Veronica Ye notes that she was puzzled by her Australian husband's habit of saying 'Thank you'. In her culture, that expression had no place within the family: one just does what one is meant to do, and gratitude is reserved for strangers. Her husband, on the other hand, was disconcerted by the fact that she *didn't* say thank you for routine favours. She had to “learn to be polite” when politeness seemed to her simply out of place (Ye 2003). Obviously this sort of discrepancy could cause friction until it is defanged by its recognition as a cultural variable. But the case can be further complicated by temperamental and personality difference. A person low on extroversion and high on neuroticism might not ever get to the point of bringing the vague dissatisfaction caused by such differences into the open.

I shall say nothing more here about temperamental and cultural differences. Their impact on intimate relations is powerful, but not news. Instead, I shall turn to yet another level at which conversation can be manipulated, on a level that can involve a measure of self-deception, according to the goals of conversation.

When we tell a story about ourselves, we shape the telling in the light of our aim in telling it. This could involve merely conveying facts, or inducing sympathy, or providing entertainment, or avoiding and sometimes provoking conflict. (Marsh and Tversky 2004). The range of goals served by our utterances are often specific to the occasion. One wants to seduce, or placate, or take revenge, and so on. Whatever it may be, one might be playing different "games" (Berne 1964).

Berne's "Transactional Analysis" (TA) reveals yet another level at which the devious intentionality of "games people play" subverts the surface meaning of what appears as rational argument. The TA model brings some interesting modifications to the Freudian structural model of Id, Superego and Ego: the "Ego-State Parent" replaces the Superego, and the "Ego State Child" replaces the Id. Neither is so ineluctably locked into the unconscious as it is in the Freudian model, and this reflects the fact that with a certain effort of reflection those patterns are discernible to the naked eye of self-observation. That isn't to say they are easily evaded, but their unveiling doesn't require quite such a lengthy expenditure of psychic archeology. Here are some of Berne's names for " games which characteristically evolve into their most full-blown forms in the marital relationship": 'Corner'; 'Courtroom'; 'Frigid Woman'; 'Frigid Man'; 'Harried'; 'If It Weren't for You'; 'Look How Hard I've Tried'; and 'Sweetheart'.

These names suggest a certain whimsy, but it would be a mistake not to take them seriously. They capture insightful descriptions of intimate manipulative strategies. Their rhetorical tactics serve goals that tend to make sense only in the context of personal relationship. In other circumstances, (although other types of games might be played) these would get little traction.

Consider, for example, a representative illustration of "Corner".

1. Mrs. White suggests to her husband that they go to a movie. Mr. White agrees. 2a. Mrs. White makes an "unconscious" slip. She mentions quite naturally in the course of conversation that the house needs painting. This is an expensive project, and White has recently told her that their finances are strained.... This is therefore an ill-chosen moment to bring up the condition of the house, and White responds rudely. 2b. Alternatively: White steers the conversation around to the house, making it difficult for Mrs. White to resist the temptation to say that it needs painting. As in the previous case, White responds rudely. 3. Mrs. White takes offense and says that if he is in one of his bad moods, she will not go to the movie with him, and he had best go by himself. He says if that is the way she feels about it, he will go alone. 4. White goes to the movie (or out with the boys), leaving Mrs. White at home to nurse her injured feelings.... White knows very well from past experience that he is not supposed to take her pique seriously. What she really wants is to be honeyed out of it; then they would go off happily together. But he refuses to play, knowing that his refusal is dishonest: he knows she wants to be coaxed, but pretends he doesn't. He leaves the house, feeling cheerful and relieved, but looking wronged. She is left feeling disappointed and resentful. *In each of these cases the winner's position is, from a naive standpoint, irreproachable; all he or she has done is take the other literally.* This is clearer in (B), where White takes Mrs. White's refusal to go at face value. They both know that this is cheating, but since she said it, she is cornered. [my emphasis] (Berne 1964, 92).

What is characteristic in this interaction is not only that it is shaped by emotion, but that emotion itself is used strategically. Robert Frank (Frank 1988) and Jonathan Schell (1984) have also noted an important class of examples of the strategic use of emotion, in the deterrent effect of allowing emotion visibly to overwhelm rationality. This is clearly as effective a tool in personal relations as in political ones, except that the long-term effects of its excessive use might be more rapidly to diminish its effectiveness.

### **5. Belief perseverance and counterproductive evidence.**

One last observation about the perils of intimate conversation will take us away from unconscious mechanisms of influence at the microlevel. It belongs instead to the wider sphere of the dynamics of belief.

The phenomenon in question involves experimental evidence that concern prejudices adduced to support political attitudes. Their bearing on personal relations is indirect. But it seems reasonable to speculate that the same mechanisms are active in the dynamics of intimate relations, *mutatis mutandis*. For in that domain, no less than in politics, one is often favourably or negatively disposed towards another in ways that have little to do with rational consideration of evidence.

The psychological literature on the "perseverance of belief" confirms in the lab what we all know from ordinary experience: persuading anyone of anything by rational argument is extraordinarily difficult. (Ross and Lepper 1980). If we happen to believe humans are "rational animals", that is both surprising and depressing.<sup>6</sup>

---

<sup>6</sup>From the point of view of evolutionary psychology, it might make sense in terms of the thesis recently defended by Hugo Mercier and Dan Sperber, that the basic function of argument is not to establish

More recently experiments by Nyhan and Reifler (2010) have shown that when people have factual misconceptions of the sort associated with certain political convictions, evidence of the falsehood of these convictions, far from persuading them to give them up, actually reinforces the original misconceptions. Part of what is at stake here is the ineffectual rhetorical effect of mere negation, which is already been remarked on and was exploited by some of the techniques used by Derren Brown. Nyhan and colleagues explored, as one case in point,

the persistent rumor from the 2008 presidential campaign that Barack Obama is a Muslim, comparing the effectiveness of Obama's use of what we call a misperception negation ("I am not and never have been of the Muslim faith.") with what we call a corrective affirmation ("I am a Christian."), which should be more effective at reducing misperceptions. As expected, we find that the misperception negation was ineffective. However, our hypothesis that the corrective affirmation would successfully reduce misperceptions was only supported when a non-white experimental administrator was present—a finding that is consistent with the literature on race of interviewer effects in survey research. In addition, experimental participants who received the corrective affirmation with only white experimental administrators present became more likely to believe Obama is a Muslim and less likely to believe he was being honest about his religion (Nyhan, Reifler, et al. 2009).

any sort of truth, but to convince an interlocutor—or oneself—of a foregone conclusion (Mercier and Sperber 2011).

The effect is complex, and we can expect that it would be no simpler in the analogous case of personal relations. The fact that one's interlocutor disagrees, even if she is in a position of epistemic authority—for example because she is directly describing her own feelings—may cause one to become more entrenched in one's own contrary conviction. "You're just saying that to annoy me." Or, "You're just projecting". Or any number of alternative rationalisations akin to the complex manoeuvres exemplified in Eric Berne's inventory of emotional manipulations.

## **6. Conclusion**

With so many ways that communication can go wrong, it is indeed something of a wonder that we ever manage to understand one another at all. At every level of analysis, language itself seems designed to trip us up. The process of learning to feel complex emotions sets up another layer of potential misunderstanding between any two people who are not indifferent to one another. Random words and other stimuli can exert unsuspected influence on our value judgments. The equally covert raids by irrelevant stimuli on apparently straightforward cases of direct perception threaten to distort the factual inferences we make from sensory evidence. And finally, the imbrication of our temperamental bias, cultural prejudices, and individual emotions with our factual beliefs can block or distort what would, from a rational point of view, appear to be the most straightforward and undeniable inferences. All of these factors acquire an added virulence in the context of personal arguments between people whose cause is not the mere conveying of information, but the influencing of emotion and desire. When exploited by poets, these mechanisms enrich the emotional impact of words; even here, however, that impact is neither universal nor reliably predictable. Here as elsewhere, as we endeavour to shed more light on our

own mental processes in the pursuit of a rational life, mere awareness of the pitfalls all about us will not keep us from stumbling. But as therapists know, or at least hope, awareness may sometimes help, a little, to forewarn, if not entirely to forearm.

\*\*\*\*\*



## References

- Asch, S. (1962). *Social Psychology*. Englewood Cliffs, N.J.: Prentice-Hall.
- Bargh, J. A., Chen, M., & Burrows, L. (1996, Aug). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology*, *71*(2), 230-244.
- Berne, E. (1964). *Games People Play: The Psychology of Human Relationships*. New York: Grove Press.
- Berns, G. S., Chappelow, J., Zink, C., Pagnione, G., Martin-Skurski, M. E., & Richards, J. (2005, online 22 June 2005). Neurobiological correlates of social conformity and independence during mental rotation. *Biological Psychiatry*, *58*(3), 245-253.
- Bruner, J. S., & Goodman, C. C.. (1947). Value and need as organizing factors in perception. *Journal of Abnormal and Social Psychology*, *42*, 33-44.
- Chapman, H., Kim, Susskind, J., & Anderson. A.K. (2009). In bad taste: Evidence for the oral origins of moral disgust. *Science*, *323*, 1222-6.
- de Sousa, R. (1987). *The rationality of emotion*. Cambridge, MA: MIT Press.
- de Sousa, R. (2009). What else is there? Emotions, value, and morality. *Critical Quarterly*, *50*(42), 12-24.
- Frank, R. H. (1988). *Passions within reason: The strategic role of the emotions*. New York: Norton.

- Frege, G. (1980). On sense and reference. *Mind*, 65(259), 51-78. M. Black (trans. and ed.). Reprinted in P. Geach (Ed.), *Translations from the philosophical writings of Gottlob Frege*. Oxford: Blackwell. (Original published 1892)
- Freud, S. (1915). Observations on transference love. In *Standard Edition of the complete psychological works* (A. Tyson, & J. Strachey, eds) vol. 12. London: Hogarth Press, pp. 157-173.
- Gopnik, A. (2011). Get smart: How will we know when machines are more intelligent than we are? (*New Yorker*, April 11: pp. 71-74).
- Grice, P. (1989). *Studies in the way of words*. Cambridge, MA: Harvard U. Press.
- Hatanama, Y., Lissek, S., Bar-Haim, Y., Britton, J. C., Fox, N. A., Leibenluft, E., et al. (2010). A Meta-analysis toward the establishment of novel treatment for anxiety. *Biological Psychiatry*, 68, 982-990.
- Henrich, J., & Boyd, R. (1998). The evolution of conformist transmission and the emergence of between-group differences. *Evolution and Human Behavior*, 19, 215-242.
- Kundera, M. (1984). *The unbearable lightness of being*. New York: Harper and Row.
- Marsh, E., & Tversky, B. (2004, July). Spinning the stories of our lives. *Applied Cognitive Psychology*, 18/5: pp. 491-503.
- Nietzsche, F. (1974). *The gay science*. (W. Kaufmann, Trans.). New York: Random House, Vintage House.

- Nyhan, B., & Reifler, J. (2010). When corrections fail: the persistence of political misperceptions. *Political Behavior*, 30: pp. 303-330.
- Nyhan, B., Reifler, J., & Undergraduate Students. (2009, ). The effects of semantics and social desirability in correcting the Obama Muslim myth, , .
- Plassmann, H., O'Doherty, J., Shiv, B., & Rangel, A. (2008). Marketing actions can modulate neural representations of experienced pleasantness. *Proceedings of the National Academy of Sciences*, 105/3: pp. 1050-1054.
- Ross, L., & Lepper, M. (1980). The perseverance of belief: Empirical and normative considerations. In R. Shweder (Ed.), *Fallible judgment in behavioral research: New directions for methodology of social and behavioral science* (Vol. 4, pp. 17-36). San Francisco: Jossey-Bass.
- Schell, J. (1984). *The abolition*. New York: Knopf.
- Stokes, D. (2012). Perceiving and desiring: A new look at the cognitive penetrability of experience. *Philosophical Studies* 158/3: pp. 477-492.
- Wheatley, T., & Haidt, J. (2005). Hypnotic disgust makes moral judgments more severe. *Psychological Science*, 16(10), pp. 780-784.
- Ye, Z. V. (2003). La double vie de Veronique: Reflections on my life as a Chinese migrant to Australia. *Mots Pluriels*, pp. 23.