

Gloria Origgi
Italian Academy
7th Annual Roundtable of Philosophy of Social Science
Barnard College, Columbia, New York, April 11-13 2005

What Does it Mean to Trust in Epistemic Authority?

*“There must be a minimal degree of trust in communication for language and action to
be more than stabs in the dark”*

Sissela Bok, *Lying*

*“Mais qu’y a-t-il donc de si périlleux dans le fait que les gens parlent, et leurs discours
indéfiniment prolifèrent?”*

Michel Foucault, *L’ordre du discours*

*“I learned an enormous amount and accepted it on human authority, and then I found
some things confirmed or disconfirmed by my own experience”*

Ludwig Wittgenstein, *On Certainty*

Consider this case. At high-school in Italy many years ago I heard my teacher of Latin say: “Cicero’s prose is full of synecdoches”¹. I had a vague idea of what a synecdoche was, and ignored until then that one could characterize Cicero’s writing in this way. Nevertheless, I relied on my teacher’s intellectual authority to acquire the belief that Cicero’s prose is full of synecdoches, and today have a more precise idea of what my teacher was talking about. Was I justified in any sense in uncritically accepting that pronouncement by deferring to my teacher’s authority? Let us have a closer look at the example. Many things were going on in this apparently trivial case of belief acquisition. I was sitting in a classroom, aware of being in a social institution – school- dedicated to knowledge transmission, and I had been properly instructed to believe what people say in school. While listening to the teacher, I was simultaneously learning a fact, that Cicero’s prose was full of synecdoches, and acquiring a linguistic concept, that is, the word

¹ This example is a reformulation of a Francois Recanati’s example in his paper: “Can We Believe What We Do Not Understand?” *Mind and Language*, 1997, that I have discussed at length in another paper: “Croire sans comprendre”, *Cahiers de Philosophie de l’Université de Caen*, 2000. The problem of deferential beliefs was originally raised by Dan Sperber in a series of papers: “Apparent Irrational Beliefs”, “Intuitive and Reflexive Beliefs” *Mind and language*, 1997.

“synecdoche” (or, if not acquiring it, at least acquiring a rule about its appropriate use, or, better, enriching its meaning). I was learning a fact and learning a language meaning at the same time. My reliance on Italian educational institutions was strong enough to accept this on pure deferential bases.

Or consider another example. I was born in Milan on February 8th 1967. I believe this is true because the office of Vital Records in the Milan Municipal Building registered few days after that date the testimony of my father or my mother that I was indeed born on the 8th of February in a hospital in Milan, and delivered a birth certificate with this date on it. This fact concerns me, and of course I was present, but I can access it only through this complex, institution-mediated form of indirect testimony.

Or else: I know that smoking causes cancer, I’ve been told this and it was enough relevant information for me to make me quit cigarettes 10 years ago. I don’t have the slightest idea of the physiological process that a smoker’s body undergoes from inhaling smoke to developing a cellular process that ends in cancer. Nevertheless, the partial character of my understanding of what it really means that smoke causes cancer doesn’t refrain me to state it in conversations and to rule my behavior according to this belief.

Our cognitive life is pervaded with partially understood, poorly justified, beliefs. The greater part of our knowledge is acquired from other’s people spoken or written words. The floating of other people’s words in our minds is the price we pay for thinking. Traditional epistemology warns us of the risks of uncritically relying on other people’s

authority in acquiring new beliefs. One could view the overall project of classical epistemology - from Plato to the contemporary rationalist perspectives on knowledge - as a normative enterprise aiming at protecting us from credulity and ill-founded opinions. Various criteria, rules and principles on how to conduct our mind have been put forward as a guarantee to preserve the autonomy and freedom of thought necessary to the acquisition of knowledge. Just as an example, a great part of Locke's *Essay concerning Human Understanding* is an attempt to establish principles of regulation of opinions, stated in terms of obligations on one's own "epistemic conduct", that strengthen our intellectual autonomy. According to Locke, four major sources of false opinions threaten our mind:

- I. Propositions that are not in themselves certain and evident, but doubtful and false, taken up for principles
- II. Received hypotheses
- III. Predominant passions or inclinations
- IV. Authority

(Locke, *Essay*, Book 4, XX, 7)

Reliance on other people's authority is thus viewed as a major threat to the cognitive autonomy that distinguishes us as rational thinkers. Exposure to received beliefs increases our risk of being "infected by falsity", the worst danger against which the overall epistemological enterprise was built.

Yet, the massive trust of others that permeates our cognitive life calls for an epistemic treatment, and has become a central issue in contemporary debates in philosophy of

knowledge and social epistemology. A number of approaches have been put forward in order to account for the epistemic reliability of the “division of cognitive labour” so typical in contemporary, information-dense societies.

Most analyses that have been recently proposed in social epistemology, concentrate on the evidential grounds for trusting other people’s authority: Trusting someone’s authority on a given matter means assessing her trustworthiness on that matter. Trustworthiness depends on both competence and benevolence. In order to assess other people’s trustworthiness one needs evidential criteria of their competence and their benevolence. For example, a scientist who trusts the authority of a colleague on a certain experimental data grounds her judgment in her knowledge of her colleague’s previous records in that scientific domain (such as the number of publications in the relevant reviews of the domain, or the number of patents, etc.) plus the beliefs that she is self-interested in being truthful for the sake of their future collaborative work.² Yet, this “reductionist” analysis, that I will detail later, misses some central intuitions about the presumptive character of our trust in others and its motivational dimension. Trust in testimony has a spontaneous dimension that doesn’t seem to be based on rational assessment of other people’s truthfulness. Also, an evidential analysis of epistemic authority doesn’t account for cases of partial understanding, as in the examples above, in which the overt asymmetry between the epistemic position of the authoritative source and the interlocutors is such that it cannot be treated by appealing to evidential criteria only. Here, my aim is to explore some treatments of the more familiar notion of trust in social sciences, moral and political philosophy in order to understand to what extent the notion of epistemic trust

² Another possible rational motivation to be trustworthy in the case of science is the high cost of cheating in the scientific community and the fear of risking permanent exclusion (see M. Blais [1987])

may be illuminated by these analyses. I will contrast evidential vs. motivational analyses in social sciences and claim that motivational analyses can find their places in an epistemology of trust. Motivational analyses have often been described as non-cognitive. Take, for instance Lawrence Becker's distinction between cognitive vs. non cognitive treatments on trust: "Trust is 'cognitive' if it is fundamentally a matter of our beliefs or expectations about others' trustworthiness; it is non cognitive if it is fundamentally a matter of our having trustful attitudes, affects, emotions, or motivational structures [...]" To say that we trust others in a non cognitive way is to say that we are disposed to be trustful of them independently of our beliefs or expectations about their trustworthiness" [Becker 1996, 44, 60].³ I will oppose this distinction by arguing that in the case of epistemic trust a motivational analysis of trust can be cognitive, that is, it can shed some light on our mental processes of acquisition of beliefs and knowledge. In particular, I will try to ground the cognitive bases of our epistemic trust in our communicative practices. My purpose here is to explore a broader notion of epistemic trust, one that could account of what is common in cases as different as the blind trust of the patient in her doctor, the trust needed in collaborative intellectual work and the everyday trust needed to sustain our ordinary conversations.

Intellectual trust is a central question of contemporary epistemological concerns. Yet, most debate around this notion fails to provide a proper analysis of the notion and only superficially bridge it to the parallel social, political and moral treatments of trust. The result is a lack of explanatory power of this notion in epistemology. Often one has the

³ It is interesting to notice that Becker liquidates much of the recent debate around the epistemic role of motivational trust by introducing *credulity*, as the disposition to believe what another person says and to banish skeptical thoughts, and *reliance*, as a disposition to depend upon other people in some respects (pp. 45-46), both of them that lie outside the reach of a rational motivation to accept other people's intellectual authority.

feeling that talking of trust in epistemology is just a way of evoking the need to varnish our study of knowledge with some moral and social considerations.⁴ My belief is that intellectual trust deserves more attention, and its intricate relation with the notion of trust in use in social science needs to be better disentangled.

On the other hand, sociological and moral theories of trust in authority fail to make the distinction between epistemic vs. political authority and present themselves as simultaneously accounting for the two concepts.

There are some obvious parallels between the notion of epistemic trust and that of social and political trust. Trust in authority poses a similar puzzle in both cases. How can someone - an institution or an individual - legitimately impose her/its will on other people's and have a right to rule over their conducts? How is this compatible with freedom and autonomy? And why should we trust an authority to impose us a duty to obey for our own good?

Much ink has been spilt on this apparent paradoxical relation between trust in authority and freedom. And of course an equivalent puzzle can be reformulated in the case of intellectual trust: How can it ever be rational to surrender our reason and accept what another person says on the basis that she is saying this? What does it mean to grant intellectual authority to other people?

The very notion of 'authority' in philosophy is notoriously ambiguous between the authority that someone exercises on other people's beliefs and the authority that someone

⁴ Take for example Hardwig analysis in his paper: "The Role of Trust in Knowledge". There are exceptions to this critics, as for example R. Foley's book *Intellectual Trust in Oneself and Others* (Cambridge UP, 2001) in which a detailed analysis of trust in the authority of others is provided in ch.4.

exercises on other people's actions.⁵ As Friedman has rightly pointed out: "A person may be said to have authority in two distinct senses: For one, he may be said to be 'in authority', meaning that he occupies some office, position or status which entitles him to make decisions about how other people should behave. But, secondly, a person may be said to be 'an authority', meaning that his views or utterances are entitled to be believed" [Friedman, 1990, p. 57].

In both cases, the appeal to authority calls for an explanation or a normative justification of the legitimacy of the authoritative source, a legitimacy that must be acknowledged by those who submit to it. Still, I think that trust in epistemic authority and in political authority are two distinct phenomena that deserve a separate treatment.

As I said above, most accounts of epistemic trust ground its legitimacy in the evidential bases we have to assess other people's trustworthiness. Motivational accounts in the case of knowledge seem desperately unable to avoid the risk of credulity and irrationality that accompanies *prima facie* any a priori trust in others as a source of knowledge.

In what follows, I will briefly sketch evidential vs. motivational approaches to trust as they are discussed in social sciences and then try to use this distinction to gain a better understanding of epistemic trust.

Evidential accounts of trust

A common view of trust in contemporary social science reduces it to a set of rational expectations about the likely behaviour of others in a future cooperation with us. Take the definition that Diego Gambetta gives in his influential anthology on trust: "Trust (or,

⁵ For an analysis of this ambiguity, cf. R. B. Friedman (1990) "On the Concept of Authority in Political Philosophy" in J. Raz (ed.) *Authority*, New York University Press.

symmetrically, distrust) is a particular level of the subjective probability with which an agent assesses that another agent or group of agents will perform a particular action, both *before* he can monitor such action (or independently of his capacity ever to be able to monitor it) *and* in a context in which it affects *his own* action. When we say we trust someone or that someone is trustworthy, we implicitly mean that the probability that he will perform an action that is beneficial or at least not detrimental to us is high enough for us to consider engaging in some form of cooperation with him. Correspondingly, when we say that someone is untrustworthy, we imply that that probability is low enough for us to refrain from doing so” [Gambetta, 1988, p. 218]. Thus trust is a cognitive notion, a set of beliefs or expectations about the commitment of the trusted in behaving in a determinate way in a context that is relevant to us. Following the literature, I call these analyses *reductive* and *evidential*. They are reductive accounts because they don’t set trust as a primitive notion, but reduce it to more fundamental notions such as beliefs and expectations. They are evidential because they make trust depend on the probabilities we assign to our expectations towards other people’s actions towards us. I may trust or distrust on the basis of some evidence that I have about someone else’s future behaviour. As it has been stressed by contemporary literature on trust in social science, trust must be distinguished from pure reliance. Trust is an interesting notion in social sciences only insofar as it explains the implicit commitment that it imposes on a relationship. If it were just a matter of assessing probabilities of another person’s behaviour without taking into account the effect of her behaviour on our own actions, it would not be so different from general inductive reasoning. I trust on a certain level of stability of the social world around me. I trust the person that I cross when walking on the street not to assault me.

This is the minimal level of trust that a society should be able to arrange in order to perpetuate. I need to rely on some regularities of the social world in order to act. But the interest of the notion of trust in social sciences is that it takes into account not only social regularities but also commitments.

A recent explanation of trust that takes clearly in account expectations about other people's commitment – and not simply regularity - is found in Russell Hardin's analysis of trust as a *encapsulated interest*, that is, trust as belief that it is in the interest of the trusted to attend the truster's interests in the relevant matter [cf. Hardin 2002].

Thus, evidential accounts of social trust try to reduce it to justified expectations upon the objective probability of other people's commitments.

The key-aspect of evidential accounts that I would like to contrast with motivational accounts is that trust is viewed as a cognitive attitude, as knowledge or belief, for which we can find a rational justification in terms of the capacity we have to read and assess the commitments of others.

What about evidential accounts of intellectual trust? An evidential theory of intellectual trust assigns probabilities to our expectations on our interlocutors' truthfulness on a subject matter. And of course truthfulness is a matter of competence as well as of benevolence. But competence and benevolence are very different things. Competence seems to be a more objective trait than benevolence: I can trust you on your willingness to help me translating Herodotus even if I don't defer to your competence in Ancient Greek. Competence doesn't depend on your commitment to be trustworthy to me.

Most evidential accounts of intellectual trust explore the dimension of competence more than that of benevolence. The epistemological literature on assessing expertise insists on what are the cognitive strategies that we can adopt in order to assess the reliability of doctors, lawyers, witnesses, journalists, etc. Alvin Goldman argues that there exist “truth-revealing situations”⁶ in which a novice can test the competence of the expert even if she doesn’t know how the expert has come to collect her evidence. For example, the weather today is a truth-revealing situation of the expertise of the weather forecast that I have read yesterday on the New York Times. If the NYT weather forecast were systematically lower in predicting whether than the Yahoo weather forecast, I would have evidence to trust the latter more than the former even if I don’t have the slightest idea about how a weather forecast is produced. That’s our commonsense practice in order to calibrate our informants’ expertise even if we are novices in their domain of expertise. If my doctor’s therapy against my stomach ache is inefficacious, I am in a truth revealing situation to assess her competence. Of course, not every domain of expertise admits truth-revealing situations: A large portions of formal sciences such as mathematics of physics don’t. In these cases, there exist alternative strategies that allow us to assess the reliability of the overall social process that sustains the epistemic dependence of laymen to experts. Such strategies have been investigated by various authors, for example Philip Kitcher who names the overall project of describing the strategies of granting expertise to others as *The study of the organisation of cognitive labour*. As he points out: “Once we have recognized that individuals form beliefs by relying on information supplied by others, there are serious issues about the conditions that should be met if the community is to

⁶ Cf. A. Goldman "Epistemic Paternalism: Communication Control in Law and Society," *The Journal of Philosophy* 88 (1991): 113-131.

form a consensus on a particular issue – questions about the division of opinion and of cognitive effort within the community and issues about the proper attribution of authority” [Kitcher, 1994, 114]. For example, I can have methods to track your past records within a particular domain and grant you authority on the basis of your “earned reputation” in this domain⁷. Or I can grant you authority due to your better epistemic position: I call my sister in Milan and she tells me that it is raining there and I believe her because I am able to assess her better epistemic position about the weather in Milan. These accounts insist on the rational bases of our trust in other people’s epistemic authority and appeal to a conceptual framework similar to that of the evidential accounts of trust in social sciences by using the language of rational decision theory or of microeconomics⁸.

Evidential accounts of trust in authority illuminate the reasons why we reliably appeal to experts in specialized domains. But, as I said, trust in epistemic authority doesn’t seem to reduce only to assessment of expertise. Nor have we always the choice to trust or distrust. The examples that I give at the beginning of this paper show that it is not always a matter of deciding to defer to other people’s authority: It just happens that the very nature of some of our beliefs is deferential, and that’s not a phenomenon that seems to be captured by these accounts.

Motivational accounts of trust

Many authors in social science and moral philosophy claim that evidential accounts fail to provide an appropriate picture of trust, by appealing only to a set of rational

⁷ Kitcher [1992] defines this kind of authority: “earned authority”.

⁸ Cf. for an example of use of the economics framework A. Goldman and M. Shaked [1991] and P. Kitcher [1993] ch.8.

expectations on other people's motivations to commit to cooperation. Our commitment to trust is not only cognitive, that is, based on the degree of our beliefs about the future actions of the trusted. Trust involves also a motivational, non-representational dimension that may depend on our deep moral, emotional or cultural pre-commitments. Thus, in the paper I have mentioned above, Becker speaks of our trust as non-cognitive if it is a disposition to be trustful "independently of our beliefs or expectations about their trustworthiness" [p. 50]. In a book entitled *Authority* Richard Sennett defines trust in authority as an emotional commitment. And in a seminal paper, Annette Baier defines trust as the accepted vulnerability to another's possible but not expected ill will toward one, and explores the varieties of moral, emotional and cultural grounds on which we accept this vulnerability. On the same line, Otto Lagerspetz says: "trust is not the fact that one, after calculating the odds, feels no risk: It is feeling no risk without calculating the odds" [1998] These accounts try to capture the idea that in many circumstances our trust in others cannot be converted into subjective estimates of risk, because the margins of ignorance or uncertainty are too broad for such estimates to be possible. Also, as Baier points out, "trust can come with no beginnings, with gradual as well as sudden beginnings and with various degrees of self-consciousness, voluntariness and expressness" [Baier, 1986, p.240]. That is, the child who trusts her mother, the patient who trusts her doctor, the novice scientist who trusts the truth of the main results in her domain without having gone through the details of the proofs, have different degrees of control and thus of choice on their trustful attitude. As the anthropologist Maurice Bloch says in his explanation of the role of deference in rituals: "We are permanently floating in

a soup of deference” Most of the time we are not aware of the reason we have to trust. We simply do so⁹.

The moral philosophical literature on motivational trust tries to establish to what extent such trustful attitudes are morally justified. Baier’s conclusion is that they are insofar as there are minimal reasons to think that the trusted exerting her authority on us is caring for the goods we want her to care for. For example it is justified to trust our partner in the treatment of our child even if we don’t approve or understand her actions, if we have reasons to think that she cares for the child.

A more empirical literature in social psychology and economics tries to establish the effects of motivational trust on stabilizing cooperation and reliability in negotiations and in everyday life.

What about the epistemic implications of motivational accounts? Do they illuminate in any sense our trust in epistemic authority? At a first glance, motivational accounts seem better equipped to explain a broader spectrum of cases than evidential accounts do. A motivational dimension seems to be involved in the asymmetrical deferential relations of trust in epistemic authority that I’ve tried to suggest in my previous examples. Indeed, trust in epistemic authority doesn’t seem to be a matter of choice in the most straightforward examples: The child who trusts her mother when she tells her that she needs to breathe air to survive - even if she cannot see air and cannot figure out what is the role of oxygen in our survival - doesn’t have the choice to be skeptical, as well as the patient who is told by her doctor that she has contracted a potentially lethal disease. Also, we find ourselves committed to trust the intellectual authority of other people just

⁹ Bloch explains rituals as a collective moment of awareness of the deference to the tradition. Cf. M. Bloch (2004): “Rituals and Deference”, in H. Whitehouse and J. Laidlaw (eds.) *Rituals and Memory: Towards a Comparative Anthropology of Religion*, Altamira Press, London.

because we are part of the same linguistic and epistemic community, we share the same institutions and we acknowledge a “division of cognitive labour” in our community. But if we accept the principle that a certain amount of “default” trust - or spontaneous trust - is needed to sustain our cognitive life in a social environment, how do we avoid the risk of credulity that such a trustful disposition seems to imply? And if the motivational trust that sustains our social relations may be based on moral, cultural or emotional pre-commitments, what about the pre-commitments underlying the motivational trust that sustains our cognitive relations? Moral commitments to trust in other people’s intellectual authority typically ground the adherence to the most irrational beliefs. Religious beliefs or allegiance to a guru’s thoughts are often justified in terms of moral or emotional commitments. But that’s exactly the kind of beliefs that an epistemological account of trust should try to ban in order to avoid the risk of gullibility imputed to a default trustful attitude towards the words of others.

Reidian accounts of epistemic trust

Another way to argue for the role of motivational trust in knowledge acquisition is to see it as an innate disposition to accept what other people tell us. And indeed many authors have argued that a natural tendency to trust others is the only way to justify testimonial knowledge. The *locus classicus* of this position is Thomas Reid’s defense of trust in testimony: We are justified in believing what other people say because we, as humans have a natural disposition to speak the truth and a natural disposition to accept as true what other people tell us. Reid calls these two principles, “*that tally with each other*”¹⁰, the *Principle of Veracity* and the *Principle of Credulity*. But the match between these two

¹⁰ Cf. Reid [1764] *Inquiry into the Human Mind*, § 24.

principles, that Reid considers self-evident, is far from being clear. The principle of veracity is not well correlated to truth: It just affirms that people are disposed to say what they believe to be true, which does not mean that they say what it is actually true¹¹. Thus, an appeal to a natural trustful disposition doesn't suffice to justify our epistemic trust and protect it from credulity.

Reid affirms that if we deny any legitimacy, or at least, naturalness to our trust in others, the result would be skepticism. We believe “by instinct” what our parents and teachers say long before having the capacity to critically judge their competence. But that is just a way of acknowledging the pervasiveness of the lore of inheritance and upbringing to shape one owns' concepts and beliefs without explaining it. It is a fact that we are influenced by others, not only in infancy but in the acquisition of most of our beliefs. But acknowledging this fact is not sufficient explanation of why we are justified to comply with our trustful tendencies.

Modern defences of a Reidian epistemology¹² appeal to the existence of natural language as the material proof that the two principles (credulity and veracity) indeed tally with each other: Most statements in any public language are testimonial and most statements are true; if they were not it is difficult to imagine how a public language could have ever stabilized. [cf. Coady, 1992]. The very possibility of a common language presupposes a general truthful use of speech.

¹¹ See on this point K. Lehrer: “Testimony, Justification and Coherence”, in Matilal & Chakrabarti (eds.) pp. 51-67.

¹² For an overview of contemporary Reidian epistemology, see R. Foley [2001]

Tyler Burge relies on the “purely preservative character” of linguistic communication to argue that we have a priori justification to rely of what we understand other to be saying. Language, as memory, is a medium of content preservation¹³.

I have discussed these positions elsewhere¹⁴. Here let me just mention that, although these positions give us some hints of the ‘passive’, non-intentional trust that characterizes our role of cognizers in a social community, their appeal to some structural features of language is less convincing in solving the paradox of epistemic trust, that is, how it is compatible with intellectual autonomy. That is, what concerns us here is: “how intellectual autonomy is possible, given what we know about the power of one’s inheritance and surroundings to shape one’s concepts, opinions and even the way one reasons?” [Foley, 2001: 128]

Epistemic trust out of self-trust

A different line of defense of the legitimacy of trust in others has been recently pursued by Richard Foley in his book on *Intellectual Trust in Oneself and in Others* [Cambridge, 2001]. Foley derives it from the justification we have to trust ourselves. We grant a default authority in our intellectual faculties to provide us with reliable information about the world. This is our only way out of skepticism. But if we have this basic trust in our intellectual faculties, why should we withhold it from others? We acknowledge the influence that others had in shaping our thoughts and opinions in the past. If acknowledging this fact doesn’t prevent us to grant authority to ourselves, it should not prevent us to grant authority to others, given that our opinions wouldn’t be reliable today if theirs were not in the past. And even in cases of interaction of people from different

¹³ Cf. T. Burge: “Content Preservation”, *Philosophical Review*, 102, pp. 457-487.

¹⁴ Cf. G. Origgi [2004] “Is Trust an Epistemological Notion?”, *Episteme*, vol. 1, n.1, pp. 1-12.

cultures whose influence upon our thinking is poor or nonexistent, we can rely on the general fact that our cognitive mechanisms are largely similar to extend our self-trust to them [cf. Foley 2004, ch. 4]. This strategy of simulation of other minds leads Foley to a sort of “modest epistemic universalism” according to which “It is trust in myself that creates for me a presumption in favor of other people’s opinions, even if I know little about them”¹⁵ [cf. *ibidem* p.108].

I find Foley’s position attractive as it preserves intellectual autonomy and ends in justifying just the minimal trust necessary to sustain our epistemic life, avoiding the “deferential incontinence” and thus gullibility that is imputed to Reidian solutions. But Foley’s analysis lacks the motivational dimension that I think an explanation of epistemic trust should include in order to account for very heterogeneous cases such as deliberate deference to an intellectual authority, passive trust to the authority of our cultural heritage and default trust that we grant to others in spontaneous conversation. What his account misses is the idea that in many contexts trusting others doesn’t seem to be depend on what we *know* or *discover* about them, as for instance that they are similar to us. Rather, trusting others is a matter of commitment to their trustworthiness in the social as well as in the epistemic cases. One could go further, and suggest that we owe this kind of commitment even to self-trust, that is, that the authority on my own mental states does not depend on something that I discover about myself. Self-trust is the product of a responsible and deliberative commitment about the consequences of assuming some beliefs as my beliefs. Richard Moran defends this line in his recent book, *Authority and*

¹⁵ As Foley says, a stronger epistemic universalism would imply that other people’s opinions are necessarily *prima facie* credible. Cf. *ibidem*, p. 107.

*Estrangement*¹⁶. According to Moran, this act of commitment is *constitutive* of my self-knowledge. I would not expand further, but I think it shows how problematic is to ground our trust in authority in self-trust. How can we capture the motivational dimension of epistemic trust we need to have a full-fledged notion of trust in authority? As we have seen, we cannot follow moral/social accounts of trust and ground a motivational account in emotional or moral pre-commitments, because this would unavoidably lead to irrationality. Still, grounding it in some innate dispositions or deriving it from self-trust misses the whole point of understanding the nature of our commitment to trust in other people's authority.

In the last section, I will explore a different strategy, and consider one on the most straightforward contexts in which commitment, trust and knowledge bloom together, that is, human communication.

Conversation, trust and communication

One fundamental fact about the social transmission of knowledge that is surprisingly under-exploited in the epistemological literature on intellectual authority is that every social contagion of beliefs goes through a process of communication that ranges from street-level conversation to more institutionalized settings of information exchange. Our almost permanent immersion in talks and direct or indirect conversations is the major source of cognitive vulnerability to other people beliefs and reports, even when the exchange is not particularly focused on knowledge acquisition¹⁷. Communication is a

¹⁶ Cf. R. Moran [2001] *Authority and Estrangement*, Princeton University Press, especially ch. 2.

¹⁷ On the fortuitous character of lot of our knowledge, cf. R. Hardin: "If it Rained Knowledge", *Philosophy of the Social Sciences*, 33, pp. 3-24; and [2004] "Why Know?" manuscript. Cf. also Jennifer Lackey: "Knowledge is not necessarily transmitted via testimony, but testimony can itself generate knowledge" [Jennifer Lackey (1999) "Testimonial Knowledge and Transmission", *The Philosophical Quarterly*, 199, p. 490, vol. 49 n. 197]

voluntary act. Each time we speak we are intentionally seeking the attention of our interlocutors and thus presenting what we have to say as potentially relevant for them. Each time we listen, we intentionally engage in an interpretation of what has been said, and expand cognitive effort in order to make sense of what our interlocutor had in mind. In this last section of my paper, I will argue that it is the intentional, voluntary character of human communication that guarantees our intellectual autonomy even in those cases in which our epistemic position obliges us to defer to other people's authority. And the making and breaking of epistemic trust is related in many ways to our conversational practices.

There are many different styles of discourse that imply different degrees of reciprocal trust. Of course, the set of norms and assumptions that we tacitly accept when engaging in intellectual conversation¹⁸ are not the same we endorse in a party conversation where the common aim we tacitly share with our interlocutors is entertaining and social contact. Still, a basic reciprocal commitment, I will claim, has to take place in any genuine case of communication. And the cognitive dimension of this basic commitment has interesting consequences for our reciprocal trust.

Intentional analyses of communication have been a major contribution to philosophy of language and pragmatics in the last 40 years. We owe to Paul Grice¹⁹ the modern pragmatic analysis of linguistic interpretation as the reconstruction of the speaker's intentions. Simply decoding the linguistic meaning of the words conveyed in an act of communication is not enough to make sense of what the speaker wanted to tell us.

¹⁸ For an analysis of the mutually accepted norms that rule intellectual conversations, see P. Pettit and M. Smith [1996] "Freedom in Belief and Desire", *The Journal of Philosophy*, 93, 9, pp. 429-449.

¹⁹ Cf. P. Grice [1957] *Meaning*, ...

Successful communication involves cooperation among interlocutors, even when the ultimate aim of one of the parties is to deceive the other: Without at least a common aim to understand each other, communication would not be possible. Thus communication is a much richer and constructive activity than simply decoding a linguistic signal. According to Grice, we infer what the speaker says on the tacit assumption that she is conforming to the same set of rules and maxims that guide our cooperative effort to understand each other. Among these maxims, two of them are worth considering for the present purposes: One is a maxim of quality of the information conveyed: “Do not say what you believe to be false” that Grice considers as most important. This doesn’t mean that the participants in a conversation are actually truthful. But they act as they were telling the truth, that is, they conform to the maxim, otherwise the minimal common aim to understand each other would not be realized. So they need at least to pretend to be cooperative. On the hearer’s side, the presumption that the speaker is conforming to the maxim doesn’t imply that the hearer comes automatically to believe what the speaker says. She interprets the speaker on the presumption that the speaker is conforming to the maxims, and that leads her to infer what she meant, even if, later, she may be led to revise her presumption on the basis of what she knows already or what she has come to believe during the conversation.

The other maxim that I would like to consider is that of relevance. Contemporary pragmatic theories have developed a notion of relevance as the key notion that guides our interpretations. For example Sperber and Wilson’s pragmatic approach, known as *Relevance Theory*, says that each act of communication communicates a presumption of its own relevance. A relevant piece of information, in a given context, is one that

optimizes the balance between the cognitive effort I have to invest to process it and the benefits I have to entertain it in my mind. A potential communicator presents herself as having something to say that is relevant for us, otherwise we would not even engage in conversation. Communication is a very special case of behavior. It is always intentional and it requires, to be successful, to be recognized as intentional. I don't automatically give attention to every cognitive stimulus that is potentially relevant for me, but I cannot refrain to allocate at least a minimal attention to an overt act of communication that is addressed to me, because the very fact that it is addressed to me is a cue that it worth attention. The presumption of relevance that accompanies every act of intentional communication is what grounds our spontaneous trust in others. I trust a communicator who intentionally asks for my attention to convey something that is relevant for me, and adopt a stance of trust that will guide me to a relevant interpretation of what she has said (that is, an interpretation that satisfies my expectations given what she says and what she may assume we are sharing as common ground contextual information). In this rich and constructive process of building new representations and hypothesis on the presumption that they will be relevant for me, the speaker and the hearer are both responsible for the set of thoughts they generate in conversation, that is, what Sperber and Wilson call their "mutual cognitive environment". But the hearer doesn't automatically accept as true the whole set of common ground thoughts that have been activated in the conversation. She may decide to entertain them in her mind for the sake of conversation, and trust the speaker that this is relevant information for her. Our mutual cognitive environments, that is, the set of hypotheses and representations that we activate in our mind when we communicate in order to understand each other, do not overlap with the set of what we

actually believe. In conversation, our interior landscape enriches itself of new representations that have been created on the presumption of their relevance for us, a presumption we are justified to have because we have been intentionally addressed by our interlocutor. We trust our interlocutor in their willingness to share a mutually relevant cognitive environment, that is, to build a common ground that maximizes understanding and favors the emergence of new, relevant thoughts. But our previous knowledge and a more fine-grained check of the content communicated can end up in rejecting much of what has been said. Trusting in relevance of what other people say is the cognitive vulnerability that we accept in order to activate in our mind new thoughts and hypothesis that are shared with our interlocutors. There is never an automatic transfer of beliefs from one's head to another's. The "floating of other men's opinions in our brains"²⁰ is mediated by a process of interpretation that make us activate a number of hypothesis on the presumption, guided by the hearer, that they will be relevant for us. These online thoughts that serve the purposes of conversation are not accepted by default as new beliefs. They are worth considering given the trust we make in our interlocutors. There may be even worth repeating without further checking because of their relevant effects in certain conversational contexts. But they can be easily discharged if their probability is too low given what we know about the world or what we have come to know about the interlocutor. We trust our interlocutors to be relevant enough to be worth our attention. Our trust is both fundamental and fragile: It is fundamental because I need to trust in other people's willingness to be relevant for me in order to make sense of what they are saying. It is fragile because a further check can lead me to abandon most of the hypotheses I generated in conversation and withdraw credibility to my interlocutor.

²⁰ Cf. J. Locke, *An Essay Concerning Human Understanding*, ed. John W. Yolton, London 1961, 1, p. 58

Our mental life is populated by a bric-à-brac of drafty, sketchy semi-propositional representations that we need in order to sustain our interpretations of the thousands of discourses to which we are permanently exposed. We accept some of them as beliefs, we use others in our inferences and we throw a lot of them as pure noise. This doesn't make us gullible beings: We trust others to cooperate in generating relevant sets of representations, and we share with them the responsibility of these representations. Of course, our epistemic strategies vary in the course of our life. The trust of a child in the relevance of what her parents say may lead her to automatically believing the content of what is said, that is, understanding and believing may be simultaneous processes in early childhood²¹. As we grow up, we develop strategies of checking and filtering information. A presumption of trust in other people's willingness to deliver us relevant information is thus the minimal default trust we are justified in having towards testimony. This stance of trust ends up enough often to an epistemic improvement of our cognitive life²².

But our efforts in interpretation are not always rewarded. Trust in relevance guides our process of interpretation and may lead us to invest supplementary effort in trying to make sense of what our interlocutor is talking about. It is on the basis of our default trust, that we often invest too many resources with the only aim of making sense what the other person is talking about. Sometimes my supplementary efforts are rewarded, sometimes they end up in a too generous interpretation of what I was told. The overconfidence people sometimes have in the relevance of esoteric discourse depends on the direct

²¹ Interesting recent results in developmental psychology show that even young children are not gullible and have strategies for filtering information. See F. Clément, P. Harris, M. Koenig (2004) "The Ontogenesis of Trust", *Mind&Language*, vol. 19, 4, pp. 360-379.

²² D. Sperber and D. Wilson have explained the details of the correlation between relevance and truth in D. Sperber, D. Wilson (2002): "Truthfulness and Relevance" *Mind*,

proportionality between the effort people invest in interpreting others and the trust they have to receive relevant information. Trust in relevance may act as a bias that leads us to over-interpret or excessively rationalize what others say.

In a beautiful novel by Jerzy Kosinsky, *Being There*, adapted as a perhaps better known film with Peter Sellers, Chance Gardiner is a mentally impaired gardener who becomes an heir to the throne of a Wall Street tycoon, a presidential policy adviser and a media icon just by pronouncing few, enigmatic sentences about gardening.

As a result of a series of fortuitous accidents, Chance finds himself living in the house of Mr. Rand, a Wall Street tycoon and a close friend of the President of United States. In a dialogue with the President visiting Mr. Rand's house, when asked to comment about the bad season at Wall Street, Chance says: "In a garden, growth has its season. There are spring and summer, but there are also fall and winter. And then spring and summer again. As long as the roots are not severed all is well and will be well" (p. 54). Looking for a relevant interpretation, and trusting (in this case mistakenly) to Chance's willingness to be relevant, the President interprets it as an important statement about the fundamental symmetry between nature and society, and quotes him on television the day after.

We all have experiences of over-trust generated out of an over-investment in interpretation. And, conversely, an excessive investment in interpreting what a person says that reveals ill-founded may make our withdraw of trust more definitive.

Trust and comprehension are thus intimately related. An epistemology of trust should account for this relation. Our first epistemic objective in acquiring knowledge from others is to understand what they say and make sense of their thoughts within the context of ours. We are never passively infected by other people's beliefs: we take the

responsibility to interpret what they say and share with them a series of commitments on the quality of the exchange. The social dimension of knowledge is grounded in our cognitive activity as interpreters, an activity we always share with others.

References:

- Baier, A. [1986] "Trust and Antitrust", *Ethics*, vol. 96, no. 2, pp. 231-260.
- Becker, L.C. [1996] "Trust as Non-cognitive Security about Motives", *Ethics*, 107, 43-61.
- Blais, M. [1987] "Epistemic Tit for Tat", *Journal of Philosophy*, 7, pp. 363-75.
- Coady, C.A.J. [1992] *Testimony*, Clarendon Press, Oxford.
- Elgin, Catherine [1996] *Considered Judgment*, Princeton University Press
- Foley, Richard [2001] *Intellectual Trust in Oneself and Others*, Cambridge University Press.
- Gambetta, D. [1988] *Trust. Making and Breaking Cooperative Relations*, Basil Blackwell, Oxford, on line edition at: <http://www.sociology.ox.ac.uk/papers/trustbook.html>
- Goldman, A. [1991] "Epistemic Paternalism: Communication Control in Law and Society," *The Journal of Philosophy* 88: 113-131.
- Hardin, Russell [2002] *Trust*, Russell Sage Foundation, New York.
- Holton, R. [1994] "Deciding to Trust, Coming to Believe", *Australasian Journal of Philosophy*, 72, pp. 63-76.
- Kitcher, P. [1992] "Authority, Deference and the Role of Individual Reason" in E. McMullin (ed.) *The Social Dimension of Scientific Knowledge*, University of Notre Dame Press.
- Kitcher, P. [1994] "Contrasting conceptions of social epistemology" in F. Schmitt (ed.) *Socializing Epistemology*, Rowman & Littlefield, pp. 111-134.
- Lackey, Jennifer [1999] "Testimonial Knowledge and Transmission" *The Philosophical Quarterly*, vol. 49, n. 197, pp. 471-490
- Lagerspetz, O. [1998] *Trust. The Tacit Demand*, Kluwer.
- Matilal, B.M., Chakrabati, A. (1994) *Knowing from Words*, Kluwer Academic Publishers
- Moran, Richard [2001] *Authority and Estrangement*, Princeton University Press
- Raz, Joseph [1986] *The Morality of Freedom*, Clarendon Press, Oxford
- Raz, Joseph [1990] (ed.) *Authority*, New York University Press
- Recanati, Francois [1997] "Can We Believe What We Do not Understand?" *Mind and Language*, 12, 1,
- Sennett, R. [1980] *Authority*, W. W. Norton, London.
- Sperber, Dan [1997] "Intuitive and Reflexive Beliefs", *Mind and Language*, 12, 1, pp. 67-83.
- Sperber, Dan, Wilson, D. [1986/1995] *Relevance: Communication and Cognition*, Basil Blackwell, Oxford.

