



MORE HEDGEHOG THAN FOX THE COMMON THREAD IN THE STUDY OF CRIMINALS, TAXI DRIVERS AND SUICIDE BOMBERS

© *Diego Gambetta*

**Unpublished paper, presented at the Nuffield Sociology Seminar,
Oxford, October 2005**

There is a line among the fragments of the Greek poet Archilochus which says: 'The fox knows many things, but the hedgehog knows one big thing'. (...) taken figuratively, the words (...) mark one of the deepest differences which divide writers and thinkers, and, it may be, human beings in general. For there exists a great chasm between those, on one side, who relate everything to a single central vision, (...) – a single, universal, organising principle in terms of which alone all that they are and say has significance – and, on the other side, those who pursue many ends, often unrelated and even contradictory (...) their thought is scattered or diffused, moving on many levels, (...) without (...) seeking to fit them into (...) any one unchanging, all-embracing, (...) at times fanatical, unitary inner vision. The first kind of intellectual and artistic personality belongs to the hedgehogs, the second to the foxes...

Berlin, Sir Isaiah (1953), *The Hedgehog and the Fox*, New York, Simon & Schuster

I will start by telling something about the style of research that I prefer even though manifestos and rule books are not my cup of tea. I do not like preaching. There are plenty of books that tell you how you should do research whose authors shy away from doing any of their own. The best way to persuade others as to the value of what one is doing is by the examples of one's own research. We now still read with profit Durkheim's *Suicide* but not so much his *Rules of sociological method*. If you glance over old copies of the main journals there is nothing as boring and obsolete as reading about methodological diatribes.

The value of one's research shares company with the devil – both are found in the accuracy of the details of the explanations we provide. Sociology suffers from what the British philosopher Alfred Whitehead defined as the main problem of young sciences:

"It is characteristic of a young science to be as ambitious about its goals as it is superficial about the details?"

We do not often read Comte or Spence or Pareto or Parsons, their grand schemes are lost in ether, but we, at least some of us, read Tocqueville – an indefatigable precursor of a mechanism-based approach.

Far too much sociology is sloppy about details. We should try harder to check the solidity of our reasoning using maths and game theory. And we should not forget that writing with analytical precision is the other main way we have to know whether we are thinking clearly and let others know that we are. A sloppy writer may be a great thinker but we shall never know.

I detest jargon, which has done so much to ruin, often deservedly, the reputation of sociology. I do not mean technical words, or new words that identify something for which we have no synonyms in our vocabulary. I mean all embracing loose metaphors – such as liquid modernity or risk society or space of flows. Jargon is the make up donned by bad theorising to veil its hollowness. When I encounter words such as *discourse*, *modernity* or *structuration* I stop reading.

I am also generally uncomfortable with explanatory concepts that appear too distant from clear individual behaviour and attributes – such as culture, social capital or habitus. These concepts have an elusive relation with actions, beliefs, desires and opportunities. If the behavioural features one attaches to them are precise, then one should check if the vocabulary offers a better way to express them. You say for instance that social capital is about trust? Then call it trust. Even trust is too vague. Better still to make a distinction between trusting and trustworthiness, for one identifies a belief the other an action.

Generally, I believe that one's explanations should have micro-foundations, if only of a stylised kind. A book that influenced me very much – this was 1979 – was *Micromotives and macrobehavior* by Thomas Schelling. And a force that brought me to become a social scientist was my dissatisfaction with Marxism, the dominant approach of the 60s and 70s. It was hopelessly naïve when it came to making its macro views match with a plausible notion of individual actions. When I discovered the existence of analytical Marxists, who became known as the *No bullshit Marxist group*, which included scholars like Jon Elster, Adam Pzreworski, Jerry Cohen, I felt the world was a saner place. The analytical revisitation of Marxism was a major step forward for the social sciences.

Not only that. I believe there is a lot of room for working on the micro-foundations themselves even before thinking about how to use them. Fehr's work on *homo economicus* and *homo reciprocans* is one such example. Jon Elster's work on emotions is another. A third example is the posthumous book by my friend and colleague Michael Bacharach *Beyond Individual choice: teams and frames in game theory* (Princeton UP 2006). In this highly original book Michael aims right at the foundations of game theory subverting its narrow and individualistic rationality assumptions. He rigorously shows how we can incorporate into the theory psychological mechanisms such as frames, social identities, and salience, which crucially shape human behaviour. This is a kind of foundational theory we badly need. These are some of the mechanisms that underlie general notions such as social interaction and context effects.

Simple as it is, the above list of desiderata is enough to screen out about 9/10th of what is produced today under sociology beleaguered name. But a lot of good is left. My tastes are catholic and I find lots of interesting things in lots of different types of research, whether problem or theory driven. Even in some driven by data. And even though I think that the compulsive unthinking application of one method to everything is not to be encouraged, I

think there is a lot to say for research that tries out new methods of both data gathering and analysis.

Believe it or not, I myself have started a project that is method driven – a survey-based experiment with John Ermish, an economist at Essex. We plan to run a newly designed trust game with about 300 subjects taken from a British Household Panel Survey (BHPS) related panel on a representative cross-section of the UK population. So far there have been only two cases of the use of this method. It will allow us to measure people's trustworthiness in a simple situation and analyse it through all the features they have and that we know through the survey. The difficulty of measuring trustworthiness is the single biggest obstacle that has hampered trust research. If you ask "are you trustworthy" in a survey, of course people say yes – words are cheap. And the typical experiment that allows such measure behaviourally has the problem of being carried out on students... who are not exactly people.

As a consumer of social sciences therefore I am distinctly a fox. In fact, if I may add a last point to the grumpy old man list, I dislike sectarianism – that consists of thinking that only what you do is worth doing. The failure to understand that the evolution of our discipline positively depends on having a varied ecology of research is a serious failure indeed. I have always found much to learn not only from my colleagues who do different things, but also from neighbouring disciplines. In fact, a large share of my work since the late 70s has been carried out in collaboration with economists. Sometimes I feel that the best sociological work is carried out under the aegis of other disciplines. Not a nice feeling.

In my career I investigated a wide variety of topics, more than I care to list here. Topic-wise too I am a fox. I get bored easily or rather excited by new things. Even before my doctorate I started out studying the links between the Italian political system and the class structure. Italian politics, with its spectacular vagaries, has remained a compulsive interest of mine, and I have, sporadically and against my best judgement done more work on it, especially on electoral reform. And before going to Cambridge in 1979 I published a second book in Italian about the youth labour supply. Both these books made heavy use of statistical works and the latter was based on a large survey that I conducted when I was a civil servant.

This was the same survey I later used for my doctorate. It was about individual school choices, and became the basis of the first book in English that I published in 1986, *Were they pushed or did they jump*. As far as I know this was one of the first research in sociology to use logistic regressions, which were then used by transportation research and one had to write the computer program by oneself. This was a theory-driven study explicitly aimed at testing explanatory mechanisms by using schooling decisions, trying to work out to what extent educational decisions were driven by rational considerations or by constraints whether normative, financial or cognitive. (Ten year later I collaborated in a study that replicated *Were they pushed or did they jump?* L.Abburrà, D.Gambetta and R.Miceli 1996 *Le Scelte Scolastiche Individuali*, Torino, Rosenberg & Sellier.)

With that book on education choices I took an intellectual step that has remained very stable throughout the rest of my work, and that is an interest in micro explanatory mechanisms. The contours of a hedgehog were beginning to emerge.

After my PhD my substantive interests changed however. I began a long stage of research whose offspring are still with me today. In this I have been partly problem-driven. But I have not shied away when the problems led to a theory-driven follow up and then back again to new problems tailored for the theory. The interaction between theory and the empirical world is not only found in each of my pieces of research, but has also inspired my journey from one piece to the next. I did not have an overall plan, just one thing led to another.

So here, trying to avoid any ex-post rationalisation, I will describe the trajectory that led me first from studying trust to studying the mafia; then how the mafia led to my interest in mimicry and signalling theory, and how this has now come full circle leading me back to studying trust among taxi drivers and an assortment of criminal behaviours. Finally, I will tell you something about how certain forms of self-harm, including the ultimate one, namely suicide missions, can be illuminated by signalling theory. Signalling theory is the common thread. How did I get there?

At the beginning of it all there was an empirical question. In the mid-80s I became absorbed by arguably the most enduring puzzle in development economics: why has the South of Italy manifested such a persistent inability to develop both socially and economically? In the South three unfortunate states of affairs have coexisted for a long time: people do not often cooperate when it would be beneficial for them to do so; they often compete in harmful ways and consider violence never too remote an option for settling their controversies; finally, they refrain from engaging in that kind of competition from which they could all gain. While there are other parts of the world cursed by a similar situation, the puzzle becomes a lot more intense once we consider that the South of Italy is part of a country that has otherwise grown rapidly and successfully after the Second World War to become one of the world's richest nations. The close proximity with the rest of Italy should have offered plenty of opportunities for the South to develop. Yet, despite the emergence of a few sparse niches of growth, these opportunities were not taken up.

It was while searching in the literature for an answer that was convincing at the micro level and not finding one that I landed on the notion of *trust* (Gambetta 1988). My idea was simple: since there was no evidence that southerners were less rational than anyone else, how could they fail to promote their common good? What was the element that made them so persistently bad at cooperating? So I hypothesised that the lack of trust in the South might be at the source of its tenacious development difficulties. The literature on trust, which has since become a very popular topic, was then miniscule and trust was not a scholarly concept. To make rapid progress a collective effort was required.

This is why I ran a seminar in 1985-6 inviting different people from the social sciences, but also from philosophy and animal behaviour. This line of research proved fruitful in several ways, and the edited volume that came out of that seminar has had some influence on subsequent studies, for example on Robert Putnam's *Making Democracy Work* (Laitin 1995: 173). It left however many unanswered questions and unsatisfactory answers, which have kept occupying me till today.

The study of trust in that context also had an unexpected effect on my research. It 'forced' me into a perilous deductive chute: for if it is the case that trust and trustworthiness are scarce in the South of Italy, which they are, and hamper economic development, which they

almost certainly do, the next question is: how has a minimum of social order and productive activity managed to persist? What is it that obviates the absence of trust preventing an even deeper social dissolution?

At the chute exit, as it were, I found myself in Palermo in 1986-7, doing fieldwork on the Sicilian mafia. This is still the most demanding piece of research I ever did. I did not have a family then.

My intuition was that the mafia, at its core, is an institution that exploits and thrives on the absence of trust, by providing extra-legal protection, largely in the form of enforcing contracts, settling disputes and deterring competition. While not the only response, it epitomises the South of Italy responses to the lack of trust.

It does not follow that mafia activities are socially beneficial. Quite the opposite, for the manner in which the mafia discharges its services, while giving a sense of security to those under its wings, ends up reinforcing distrust for all. It does so directly by providing an incentive for people to behave in an untrustworthy manner when they are under mafia protection (except towards those who are also under mafia protection); it drives out all intrinsic motivations for good behaviour in business and any incentive to find self-managed and thus cheaper ways to cooperate. The mafia is the opium of distrusting people.

The theory of the mafia that I developed goes against common interpretations and media commonplaces on organised crime, and it has not really percolated into the wider world, and only partially in that of law enforcement. Still, it has had some academic influence; several scholars have now shown how the overall model developed for the Sicilian mafia can be successfully applied to other mafia-like organisations in other parts of the world, Hong Kong, Russia, Japan, Colombia.

The book on the mafia also had an impact among economists. Oriana Bandiera, an LSE economist, has recently (2003) produced an elegant empirical test of the part of the theory that deals with the origins of the mafia. And Avinash Dixit, a theoretical economist at Princeton, has started to systematise the general field in a recent book called *Lawlessness and economics*. This field is in essence a set of models of extra-legal protection, of which the mafia model is one. Dixit challenges a simplifying assumption which economics has made since its inception – that protection and property rights enforcement are matters smoothly and effectively performed by legitimate states under the rule of law. This assumption works well enough for developed countries but it leaves the theory ill-equipped when it comes to dealing with the rest of the planet. The awakening of economic theory to the problem of modelling extra-legal protection is a step forward of considerable importance, which may turn out to be of help in shaping policies in developing countries as well as in countries beset by unstable political institutions or civil war.

The study of the mafia left me with a number of unforeseen findings and related questions. Why do they use nicknames, why do they imitate movies, why do they flaunt their incompetence in every field except in the use of violence? And I duly tackle each of them in the book on which I am working, called *Crimes and signs: cracking the codes of the underworld*.

More than anything else Mafiosi left me with one finding that stimulated the direction of my research since 1993, when the mafia book came out. I discovered that they were very

worried about impostors – people who claim to be Mafiosi and are not, and people who claim to be protected by Mafiosi and are not. (I leave you to imagine how they deal with either of these types if they catch them.) The same worry has been found in other mafia-like groups, by Federico Varese, Peter Hill and by Yiu Chu Kong.

Under secrecy constraints, the opportunities to pass off as someone you are not are greater for it is harder to be discovered; and the beauty of it all is that when intimidation is the key means of production if you can persuade someone that you are a real Mafioso, there is nothing else you need to do. Once established, reputation in this world does all the work and creates an incentive to parasite on it.

It is by pondering over these findings that I became intensely interested in signalling, both honest and fraudulent. After all, an impostor displays the right signals without having the corresponding qualities. And I took a long theoretical detour, in an attempt to find the general principles that could be applied to make sense of the strategies of both impostors and of their victims, whether in the mafia or anywhere else.

I ought to add, that personal experience intruded in sensitising me to the importance of signals. I also found out, less surprisingly, that Mafiosi are worried about other impostors too, even more dangerous for them – namely undercover police and informants, and, as you can guess, nosy academics. When I arrived in Palermo I heard the story of a Canadian researcher who had just been rapidly leaving town. Someone had stolen his dirty laundry from his car boot. A few days later he found it cleaned and ironed back in the boot of his car with a note that simply said ‘buon viaggio’. When in Palermo I myself masked my research as being about trust, but towards the end of my stay my presence there must have become more obtrusive than I would have wished. I too received interesting subtle messages that encouraged a rapid exit. I thus realised that violence was not the only resource of these men.

Signalling theory

In studying signalling theory I joined forces with the late Michael Bacharach, who had the finest mind of anyone I ever encountered, and from whom I learnt a vast amount. We worked together for several years. Signalling theory has been developed in game-theoretic form in both economics and in biology since the 1970s. It was the only candidate we knew suitable to tackle the very general question lying behind the problem of impostors.

The question is: under what conditions can truth be perfectly transmitted when signallers have an interest in deceptive communications? More precisely, under what conditions can a signal be rationally believed by the receiver, when the signaller has an interest in merely pretending that something is true, when he has in other words an interest in mimicking truthfulness in some way?

Signals are the stuff of *purposive* communication. They are any observable features of an agent, which are intentionally displayed for the purpose of making the receiver believe that a certain state of affairs obtains. The state of affairs could be anything, but it usually refers to things that are not easily observable such as trustworthiness, toughness, generosity, wealth, innocence and even identity. Things that the receiver wants to know about, and that the signaller whether truthfully or otherwise would like the receiver to believe.

The ‘features’ that make up a signal too could be anything: they include his verbal claims but also parts or aspects of his body, pieces of behaviour by him, and his appurtenances. The first steps to apply the theory are to identify what exactly are the signals and to which audience they are addressed.

The main condition that the theory posits is that among the possible signals there is at least one, which is cheap enough to emit, relatively to the benefit, for signallers who have the quality, but costly enough to emit, relatively to the benefit, for the would-be mimics who do not. If the cost relationships are such that all and only those with the true quality can afford to emit that signal, the equilibrium in which they do so is called ‘separating’ or ‘sorting’. In such an equilibrium signals are unambiguous, and the receiver is perfectly informed. No poisoner seeks to demonstrate his honesty by drinking from the poisoned chalice. When signals have such perfect discriminating property mimicry cannot occur for no mimic can afford it. (This condition is referred to in biology as ‘the handicap principle’ and in economics as the costly-to-fake principle, see Frank 1988.)

The trouble is that in real life most signals are only semi-sorting. There is hardly such a thing as a fully mimic-proof signal. Virtually everybody who boards a plane, for instance, gives a signal, most of the time unthinkingly, that he is not intent to cause it to crash. And most of the time this is true. But as we know only too well through recent events some terrorists may be prepared to do just that and can afford to mimic a normal passenger by boarding. Some people, in other words, can drink from the poisoned chalice. In a semi-sorting equilibrium there is a signal which is emitted by all those who are truthful about something but not only by them. A certain proportion of mimics can just about afford to emit it too. Here, observing that signal is not conclusive evidence; it makes it more likely that the signaller is genuine but does not imply that he is.

Theoretical work

The theory of signals is very abstract, and quite a lot of intermediate work is needed to apply it properly. One thing it lacks is semiotics structure. And we worked to extend it to cover different forms of costly signals – such as *cues*, things that do not cost anything to display to those who truly have them such as your gender or skin colour or face, while they can be very costly to display to those who do not have them. Or *marks*, things acquired as by-products of the life that each individual lives. As people grow up in given cultural settings, they absorb social features such as language, accent, mannerisms, or undergo common experiences that cannot be acquired in any other way and are therefore mimic-proof. Ethnic signatures are often of this kind. They come at no extra cost to those who had the experiences, while being very costly for anyone else to copy. The British tested potential German spies by asking them about well known cricket results.

By contrast, *symbolic* or conventional signals can be very cheap to mimic. Art Spiegelman reports that, during the German occupation of Poland, his father used to travel to town by tram. Trams had two cars: “One was only Germans and officials. The second, it was only the Poles. He always went straight to the official car” where a simple salute, ‘Heil Hitler’, was enough not to call attention, whereas “in the Polish car they could smell if a Polish Jew came in”. It was harder for a Jew to mimic the nuanced multiple signs of a Polish gentile than the fewer superficial signs of a pro-Nazi (Spiegelman 1991: 142).

Nazi signs are not cues or marks but symbols. These are configurations of characters or gestures however physically realised, exemplified by names, logos and Nazi salutes. What makes them open to mimicry is that among the physical realisations there are usually some which are very cheap for anyone to produce. The efficient production cost of a Nazi salute is zero. Symbolic signs are attractive for signallers because they are cheap, but since they violate the cost differential condition their evidential value is weak. The expansion of the scope for ultra-cheap transmission via electronic means of symbol-strings is indeed a major cause of the growth of mimicry in our time.

Next, we showed how one can extend signalling theory to cover *identity* as one of the non-directly observable qualities we signal. This step was essential to deal with impostors of the kind the Mafiosi are worried about. A great deal of human signalling takes place indirectly by signalling one's identity both as a specific individual or as a member of a group or category. After we encounter an individual or group member and experience dealing with them, we form an idea as to whether this person or group have or lack certain qualities that interest us. Identity signalling enables the signaller to exploit a reputation. If persons or group members are re-identifiable by some signature, the next time we meet them we infer the presence or absence of a given quality by establishing the identity. Without re-identification reputations would be still born.

The re-identification of a signature, however, can itself be problematic. This is because, frequently, the fact that someone is the bearer of a certain reputation is itself an unobservable property of that person. For example, Armani has a reputation for selling well-designed clothes, but to exploit this reputation a seller must convince customers that he is Armani. Islamic Jihad has a reputation for carrying out its threats against hostages, but to exploit this reputation a group of kidnappers must convince governments that they belong to Jihad. Much of human deceptive mimicry does precisely that: it exploits signalling via identity. The last step we took is linking signalling theory to trusting decisions, which we recast as a response to trustees' signalling acts.

In virtually all trust situations which occur naturally, the truster sees or otherwise observes the trustee before deciding. She therefore can, and should, use these observations as evidence for the trustee's having, or lacking, trustworthy-making qualities. Since the truster will be proceeding in this way, and given the payoff structure, there is a motive for an opportunistic trustee to 'mimic' - to emit signs of trustworthy qualities when he lacks them. This complicates the truster's problem: she must judge whether apparent signs of trustworthiness are themselves to be trusted. What we did was to show how one can transform the problem of trust into a 'secondary problem of trust, of trust in signs'. This is important because once you see trust in this way, you can apply signalling theory to it and you are on much more solid grounds. This work is just too rich to do it any justice here. Read these if you wish to know more.

Applications

There are almost no applications of signalling theory in sociology. Yet, I believe that it is a most promising theory, which offers a tremendous potential for research, and can generate far reaching testable predictions. Some of its potential lies in revisiting works whose core ideas are of great value but may be made more rigorous and tested – Veblen and Goffman.

I put the theory and its extensions to work in several directions, and it is this theory that makes me more of a hedgehog, as it runs through quite a share of my work. Here, to conclude, I will give you a sense of the applications by discussing the key questions that inspired each piece of research.

a. *Streetwise: how taxi drivers establish their customers' trustworthiness* (co-author Heather Hamill, New York: Russell Sage Foundation, 2005)

Do taxi drivers in dangerous cities pay more attention to hard-to-fake signals when deciding whether to trust a passenger?

Making sense of suicide missions (edited volume, OUP, 2006)

Three questions discussed in this book are relevant to signalling theory.

How can one spot a suicide bomber before detonation? Suicide bombers try to mix with the crowd, veil their true identity – they have been known to pretend to be pregnant or Orthodox Jews....

Is dying in suicide missions just a cost to achieve what cannot be achieved by violent but non-suicidal means or does it have some additional instrumental value in terms of communication? Does it scare enemies, make observers believe that your cause is important, persuade sympathisers to join?

And how can someone who wants to die for a cause convince others he is not just killing himself because he is unbalanced?

b. *Codes of the underworld* (Princeton University Press, 2009)

No one appreciates the ever-present possibility of deception more than a criminal does, whether his intent is deceiving or avoiding being deceived. There is hardly any domain in which information is at once so valued and so imperfect as in the underworld. The world of criminal communications is replete with instances which can be understood through the theory of signals.

Identity signals – How to test criminal credentials...

The protection of conventional signals – How to make cheap signals hard to mimic ...

Why prisoners fight – Those who can credibly signal 'toughness' fight less...

The value of self-harm – Are self-harmers – in armies, prisons, or gangs – nuts or is there a signalling logic to their madness?

A larger research lesson hides in my interest in matters criminal there is. As two biologists wrote

“At the beginning of every new science information is densest at the boundaries. Darwin, for example went to the Galapagos where every species is on the margin.”¹

Criminals operate at the boundaries; they are my Galapagos.

===

So, fox or hedgehog? Once you think you have a good theory you can run around like a fox searching for applications without fear of getting lost – the fox rides on the hedgehog. Applying theory is deemed a second rate operation as if one was feeding dough into a pasta making machine, but in my view is the most difficult operation of all, in which all the skills of an imaginative sociologist are required. The risk is, as it were, that once you have a hammer everything looks like a nail. But the advantages are what I really find exciting about my work, that is, to discover the hidden logic of social phenomena. It does not matter how one begins a piece of research whether theory driven or puzzle driven. In the latter case what is exciting is to explain something complex or seemingly strange with a simple model, to show through theory that facts, often apparently small facts about social life -- mafia ‘quaint’ style, taxi drivers daily work, engineers among Islamic radicals – have a logic and more than anecdotal value. In the former case the excitement comes from discovering, guided by theory, that behavioural phenomena that are either disregarded or explained away by psychological traits such as self-harm can be otherwise explained.

Above all, life as a mere hedgehog can be boring, and as a pure fox can feel shallow. In the essay I cited at the beginning, Isaiah Berlin went on to write: *“like all over-simple classifications of this type, the dichotomy becomes, if pressed, artificial, scholastic, and ultimately absurd.”*

¹ From a letter to the *New Yorker* by Steve Coombes and Dave Robinson, September 29, 1997.