TRUST:
Forms, foundations, functions, failures, and figures
Draft of the first three chapters,
January 2001

by

Bart Nooteboom
Erasmus University Rotterdam
b.nooteboom@fbk.eur.nl
CHAPTER 1: REASONS, AIMS AND TOOLS

This introductory chapter discusses the reasons, aims and foundations for this treatise on trust. Why is trust important? Why does it need to be discussed? What are the questions? How are they to be answered? What theories can we use tools of analysis?

1.1 Reasons and aims

Trust is of all times. In the practice of human relations, it has always drawn attention. Many of its intricacies have been recognised throughout history. Behavioural science, however, has had problems in tackling the concept. Social science neglected it until some thirty years ago. Since then, it has produced a kaleidoscope of insights, which now require some ordering for the sake of clarity and analytical grip. Economics still neglects trust. It is high time that a systematic and clear understanding of trust be incorporated in economic analysis.

1.1.1 The hide and seek of trust

There has always been interest in trust, if not in theory then in practice. Politics is about trust, and so is money. Trust is indispensible and pervasive. In all its diversity, the practice of human relations shows up in literature, and there the making and breaking of trust is a perennial theme. Consider Shakespeare. Think of his plays: ‘Hamlet, Macbeth, Othello, Richard III, The Merchant of Venice, Romeo and Juliet’. Even in his ‘lighter’ plays, we will probably find the theme of trust. ‘Much ado about nothing’, for example, is about trust and deception. Morton Deutsch, arguably the most important writer on the subject of trust, quoted Theognis in his Sententiae (666 BC): ‘He who mistrusts most should be trusted least’. And Samuel Johnson in The Rambler (1750): ‘It is happier to be sometimes cheated than not to trust’. And Nicolai Hartmannin in his Ethics (1930): ‘All human relationships ... are based on faith’ and ‘It is presumably a communal value; it is the most positive unifying force which welds together a variety of individual persons, with their separate interests, into a collective unit...’ (Deutsch 1973: 144-145). A nineteenth century Dutch politician recognised another truism about trust: ‘It comes on foot and leaves on horseback’. All these insights in trust are important and valid, as we shall see. Thus, there is no lack of insight in the importance and intricacies of trust, in the practical wisdom of human affairs.

Even Oliver Williamson in his earlier work on transaction costs (1975: 108) recognised that ‘trust is important and businessmen rely on it much more extensively than is commonly realised’, and he paid attention to what he called the ‘atmosphere’ in which transactions take place. This is striking in view of the fact that Williamson proceeded to neglect trust in his later analysis of transactions (1985), and upon later analysis (1993) explicitly rejected it as unwise and not fit for survival in markets. His argument was that if trust does not go beyond calculative self-interest it does not add anything to existing economic analysis, and if it does it necessarily yields blind trust, which is unwise and will not survive in markets, and is best limited to personal relations of friendship and family. Deutsch (1973: 144) expressed surprise at the lack of attention to trust in social science prior to his book: ‘Despite the obvious significance of trust and other related phenomena, an examination of leading empirical and theoretical works reveals that these phenomena have been largely ignored by the social scientist’. Subsequently, however, there has been a flood of literature on trust in sociology and in the business literature, as references presented in the course of this book will testify.

In modern economics, however, the neglect of trust has persisted, in spite of the fact that at the origin of modern economics, in the work of Adam Smith trust fit in naturally, in Smith’s (1759) ‘theory of moral sentiments’. I will return to that later. It is not difficult to explain this peculiar fact that in the history of economics the attention to trust did not increase but was sidetracked. That happened not only to trust but to social psychology more in general. In the development of economics after Adam Smith, economics narrowed from the ‘Enquiry into the nature and causes of the wealth of nations’ to a
calculus of rational choice. The machinery, or perhaps we should call it the toy, of modern economic analysis breaks down when we go beyond risk, in the economist’s interpretation as the variance of a distribution (‘density function’) of probabilities attached to alternative outcomes. Beyond risk there is radical uncertainty, in the sense of Knight (1921), which entails lack of knowledge concerning the set of alternative outcomes of a choice and of the alternative options for choice. One cannot rationally choose the best from a set of alternatives if that set is not well defined. Yet that kind of uncertainty arises in human interaction: the set of outcomes of choice and the set of options for choice are open. Some outcomes and options are not given prior to choice but are revealed or produced as a result of choice and action, so that we can see only in retrospect what we could have chosen if only we had known all the options and possible outcomes.

As I will argue at length in this book, the importance and nature of trust arise from the unpredictability, or radical uncertainty, of human behaviour. Trust is not the calculation of probabilities and outcomes but a wager concerning behavioral options that are unknown. A similar argument was proposed by Lazaric and Lorenz (1998). Radical uncertainty raises problems for the customary interpretation of trust as a subjective probability concerning someone’s actions. I will discuss this in Chapter 2. Trust is difficult to deal with in economic analysis, due to its preoccupation with formalised theory and with rational choice, and its neglect of psychological and epistemic obstacles to such choice. That currently provides a problem for economists who now recognise that trust can no longer be neglected, but cannot adequately come to grips with it within their theoretical framework.

This is reflected in Williamson’s more recent (1996: 261) ruling on trust: ‘If calculative relations are best described in calculative terms, then diffuse terms, of which trust is one, that have mixed meanings should be avoided when possible’. As this book will show, trust is indeed a diffuse concept. However, the point here is that under radical uncertainty, which is inevitable, relations can be calculative only to a limited extent. Furthermore, the elimination of concepts for their being diffuse is inherently conservative, because new theory inevitably starts out with diffuse meanings. It is only later in the development of new theory, and novelty in general, that meanings become clear and practice settles down in a ‘dominant design’ which is amenable to formalisation (Nooteboom 2000a). The proper response to diffuseness of a concept is not to excommunicate it but to try and reduce its diffuseness by developing the concept. That is the purpose of this book. In the game of human relations I would rather not hide trust but seek it.

1.1.2 The importance of trust

Trust can have extrinsic value, as a means to achieve social or economic goals. It can also have intrinsic value, as a dimension of relations that is valued for itself, as part of a broader notion of well-being or the quality of life. Many authors have pointed this out (Blau 1964, Arrow 1974, Jarillo 1988, Bradach & Eccles 1989, Powell 1990, Casson 1991, 1995, Helper 1990, Sako 1992, Gulati 1995, Berger, Noorderhaven & Nooteboom 1995, Chiles & MacMackin 1996, Nooteboom 1996). People may prefer, as an end in itself, to deal with each other on the basis of trust. Most economists tend to think of value in exchange as something that exists independently from the transaction. As formulated by Murakami & Rohlen (1992: 70): ‘The value of the relationship itself is typically ignored and the impersonality of the transaction is assumed’. In intrinsic utility, the exchange process itself matters, as does the economic surplus and its sharing that the exchange yields. This was also recognized by Buckley and Casson (1988).

When intrinsic, the value of trust can be hedonic or based on self-respect. Many people would prefer to have trust-based relations rather than relations based on suspicion and opportunism for individualistic, hedonic reasons. For most people it is more agreeable or pleasurable to have friendly relations than to have to deal with animosity. There is also an intrinsic motive of self-respect, based on adherence to internalised norms or values of decent or ethical conduct.

There is also a more socially oriented motive, in the will to be recognised, valued and respected by others. This may be served by accumulating riches, power or glamour, but also by being trustworthy.
and trusting, and thereby demonstrating adherence to established values, norms or habits of behaviour. This may merge with the earlier motive of self-respect, while analytically it can still be distinguished. Fukuyama (1995: 358) traced the urge for recognition to Hegel: a struggle for recognition, ‘that is, the desire of all human beings to have thier essence as free, moral beings recognised by other beings’. As indicated, this does not necessarily lead to trustworthiness and trust. Philosophers have also postulated other urges, such as the will to power. However, the point is that under proper institutional conditions and individual dispositions the urge towards recognition can lend value to trust and trustworthiness. When those conditions are lacking, however, it can lead to pathologies such as ethnic discrimination: recognition is derived not from benevolence but, on the contrary, from the condemnation and harassment of others (Fukuyama 1995: 361).

The extrinsic economic value of trust lies in the fact that it enables interaction between people and between organisations and can reduce transaction costs. In the analysis of this, I will employ elements from transaction cost economics, but in a wider theoretical framework, in recognition of a number of fundamental objections to that theory as developed by Williamson (1975, 1985, 1991, 1993, 1996). I will discuss transaction cost economics in the discussion of theoretical tools, later in this chapter.

To indicate the extrinsic economic importance of trust, here I simply posit that trust is one of several means, but an indispensable one, for conducting economic relations. The arguments will be developed later, in Chapter 4. The fact that trust is now gaining more attention, even in economics, can be explained by the fact that interaction between more or less autonomous agents (people, organisations) is now deemed more important for economic success than it was before. There are several reasons for this. One important and widely recognised current condition is that information and communication technology (ICT) enables interaction in new ways. A second, also widely recognised, current condition is that organisational disintegration into more or less autonomous units of activity yields a greater variety and flexibility in the configuration of activities than units that are rigidly integrated in an organisation, with strong and persistent linkages (Nooteboom 2000a). That variety and flexibility, in networks or ‘industrial districts’ of firms, are needed under current conditions of complexity and rapid change in technologies and markets. Such conditions obtain more generally at the upsurge of fundamental new technologies, at the beginning of long economic (Kondratieff) cycles. And indeed, at that stage we have witnessed the emergence of network-type organisational forms, and the ‘swarming’ of small new firms before. At a later stage this tends to be followed by more integration in larger and more stable organisational units.

In earlier work, I proposed another, deeper, cognitive reason for relations between a variety of more or less autonomous organisational units (Nooteboom 1992, 1993, 1999a, 2000a). For this I need a theory of knowledge, and I will summarise the theory I use in a later, section, in the discussion of theoretical tools. That theory is needed also because it has implications for the role of trust in relations between people and organizations, to be discussed in Chapter 4.

To say that trust can foster productive economic relations does not imply that it necessarily does so. Indeed, it can go too far and have negative effects, in making economic relations too exclusive and durable, and thereby creating rigidities. This will be discussed in Chapter 5, where I treat the failures of trust.

1.1.3 Need for coherence

In the past, particularly the past three decades, much has been written on trust, in social science and in the literature on management and organisation. Often, much of this is neglected, and wheels are re-invented. In their attempts to deal with trust, economists re-invent sociology, as they must to deal with it adequately, if they choose to ignore the existing accomplishments of sociology. In their attempts not to neglect rivalry, sociologists re-invent economics. My aim is to utilise insights that have already been offered, in a wide range of literature, although I cannot claim to master the entire literature on trust. In
particular, I will make much use of the work by Deutsch (1973). Of course, I will not agree with all he proposed, and I will use insights from others to correct what I see as errors or limitations in his work.

The literature offers a kaleidoscopic set of images: what one sees depends on the twist that one takes. My aim is to systematise this seemingly diffuse constellation of insights, and thereby make the notion of trust more precise and amenable to analytical grip. However, I do not wish to reduce the concept too much, and I wish to maintain its essential richness.

There is a widespread consensus that trust has to do with expectations of behaviour. For example, Sako (1992: 32) defined trust as follows:

‘.. a state of mind, an expectation held by one trading partner about another, that the other behaves or responds in a predictable and mutually expected manner’.

This leaves open many forms or types of trust. Most authors try to come to grips with the complexity of the notion of trust by offering distinctions between different kinds. Sometimes these distinctions are confusing or even yield category mistakes. Sako, for example, makes a distinction between the following three different kinds of trust: competence, contractual and goodwill trust. In my view, this confuses different categories. Competence contributes to one’s reliance on a partner. However, competence trust also refers to an object of trust, competence as something one can have trust in. Does one trust someone to be able to conform to expectations? In my view, this can be distinguished from intentional trust, which refers to someone’s intentions as the object of trust. Does one trust someone to intend to use his ability to conform to expectations? Contractual trust, on the other hand, belongs to the category of the basis for trust. It can apply to competence: one can safeguard competence by contractual means. It can also apply to intentions: to a greater or lesser extent one can make contractual safeguards against opportunism. I am not sure what Sako means by goodwill trust as her third type of trust. Does it refer to intentions as an object of trust, so that it becomes equal to intentional trust? Or does it refer to the basis for trust, such as perceived or assumed loyalty, which may be based on ethics, friendship, kinship or routinised behaviour? I propose that in categorising trust we should distinguish between its objects (what do we trust) and its basis (what is it based on).

Does the definition of trust depend on the basis of trust? There is a puzzle whether trust in the sense of expected conformance to expectations can be called trust if it is based on contractual obligations. The notion of trust in everyday language seems to suggest that trust goes beyond contractual obligation, and refers to expected cooperation even when it is not specified in a contract, and even when it may require a sacrifice from the partner. Thus, trust has been defined as the expectation that a partner will not engage in opportunistic behaviour, even in the face of countervailing short-term incentives (Bradach and Eccles 1984, Chiles and McMackin 1996). I will discuss this issue later, in Chapter 3, and I will make a distinction between ‘reliance’, which is a broad term including all bases of expectations, and ‘trust’, which is much stronger, and refers to expectations that are not based on safeguards in the form of contracts or self-interest.

Further distinctions offered in the literature are: cognition or knowledge based trust, experience based trust, process based trust, i.e. trust that results from interaction between partners (Macaulay 1963, Zucker 1986, Helper 1990, Larson 1992, Parkhe 1993, Gulati 1995, Zaheer and Venkatraman 1995), affect based trust (Mc Allister 1995), and faith. Knowledge can derive from own experience or from information obtained from others, as in a reputation mechanism. Clearly, experience and process based trust connect: experience arises from the process of interaction. However, interaction has wider consequences than offering experience. It affects perceived goals, norms, values and feelings of friendship, loyalty, rivalry, animosity and thereby can build trust or suspicion. Thus, it can also be connected with affect based trust. Faith suggests a belief which is not based so much on knowledge or rational argument as on conviction, which may be based on affect. In my view, the notions of knowledge, process and affect based trust belong to the category of the basis for trust.

Now, it seems to me that when we talk of the basis for trust, we should distinguish between the mental, cognitive or emotive processes underlying trust, the processes of interaction underlying
trust, and reasons or rational justification of trust.\(^1\) Here, I interpret cognition in a wide sense, including perception, interpretation and evaluation, and hence value judgements. It includes tacit next to codified knowledge. I will discuss these notions, and my theory of knowledge, in the paragraph on theoretical tools. The mental basis of trust, then, may be knowledge obtained from others (as in a reputation mechanism), or from one’s own experience. It also includes feelings of friendship or empathy, convictions, which include delusions, and routinised behaviour. One may neglect the possibility of things going wrong, due to blind faith. One may not be aware of the possibility of things going wrong, due to habit or routine. Trustworthiness is often taken for granted until violated. In routinised behaviour, certain actions in certain contexts never went wrong, and we lose awareness of the possibility (Parkhe 1993, Gulati 1995, Nooteboom, Berger & Noorderhaven 1997). Reasons, on the other hand, refer to a rational assessment of someone’s trustworthiness. This entails an evaluation of someone’s capabilities, goals, motives, inclinations, and opportunities, given conditions such as the observability of actions and accomplishments, competitive pressure, existence and efficiency of contracts, social conditions, personal ties, presence and efficiency of reputation mechanisms, etc.

Another issue that requires clarification is the relation between trust and risk. As indicated before, a fairly universal intuition about trust is that it is an expectation concerning the behaviour of others. Another fairly universal intuition is that it entails risk: the possibility that expectations will not be met, and that actual outcomes will be worse than expected, and possibly disastrous, i.e. yielding a negative outcome that is larger than the positive outcome hoped for (Deutsch 1962, 1973, Zand 1972). This has bred the widespread notion of trust as a subjective probability that a partner will achieve a certain performance (Gambetta 1988, Dasgupta 1988, Mayer, Davis & Schoorman 1995, Gulati 1995). There is trust if this probability is high, and hence perceived risk is low. However, that implies that trust is highest when the probability is unity, so that there is no risk left. But another intuition is that trust entails acceptance of risk, so that if there is no risk, we can no longer speak of trust. How is this puzzle to be resolved? I will turn to this in Chapter 2.

Another puzzle derives from a useful challenge thrown up by Williamson (1993). As noted before, he argued that the notion of trust adds nothing to existing economic analysis if it does not go beyond calculative self-interest. However, if it does, it inevitably leads to blind, unconditional trust, and that will not survive in markets. On the level of societies, Hill (1990) argued the opposite: trustful societies will enjoy lower transaction costs and will therefore have a competitive advantage and will survive in the long run. I think both are wrong, and I will present my arguments in Chapter 3. The issue thrown up by Williamson is an important one and requires an answer: how can trust be non-calculative without being blind and unconditional? This is clearly related to the basis of trust, indicated above.

1.1.4 Questions

In this book I try to give a systematic analysis of trust on the basis of the following categories. What is the relation between trust, risk and probability? What are the possible objects of trust: what can we have trust in? What is its foundation or basis: what are the underlying processes, and what are the reasons, or rational justifications, for trust? What are its functions: its role in the economy, and how does this work? What are its failures: its limits and possible negative effects? Towards the end of the book, I will consider the operationalisation and measurement of trust in empirical work, and in models for the simulation of processes in the making and breaking of trust. For reasons of alliteration I call this the figures of trust. This yields the following chapters.

- The forms of trust, in Chapter 2. Can we define trust? Is it a form of behaviour, or the motivation or disposition underlying behaviour? What is its relation to risk, and to subjective probabilities attached to

\(^1\) A similar distinction has been made in the philosophy of science, between the ‘context of discovery’, which deals with the way in which we develop theory, and the ‘context of justification’, which deals with logical and empirical arguments to justify a theory.
behaviour of others? Is it calculative or intuitive or habitual, or all of those? If it goes beyond calculation, or falls short of it, is it then necessarily blind, unconditional and irrational? What is its relation to information, knowledge and experience? What are the possible objects of trust: what can one have trust in? Why does that matter? What is the difference between trust and related concepts, of faith, reliance and confidence? What is the relation between trust in people and trust in organisations?

- The foundations of trust, in Chapter 3. What are the rational foundations of trust, or its reasons, i.e. what are the sources of trustworthiness? Is the only reliable foundation coercion or other forms of self-interest, or are there also reliable altruistic sources? How does one infer trustworthiness from observed behaviour? How, in what processes, is trust built up and destroyed?

- The functions of trust, in Chapter 4. This yields an elaboration of the economic value of trust, as a means for enabling productive relations and for reducing transaction costs. Chapter 4 summarises and extends earlier work on the place of trust in the governance of inter-firm relations (Nooteboom 1999a). The central question is what instruments there are for the governance of relations, and whether and how trust fits in. More specific questions are the following. Can trust yield a substitute for other, more costly instruments for the governance of relations, such as contracts? Or does trust form a necessary complement to contracts? Or is trust both a substitute and a complement for contract? How does trust relate to other instruments, such as the use of hostages, symmetry of dependence, reputation mechanisms, the use of third parties or network position?

- The failures of trust, in Chapter 5. Trust may be seen as part of ‘social capital’ (Powell 1990). But social capital may degenerate into social liability (Gabbay and Leenders 1999). When can trust be deceptive? To what extent can we trust trust? What are its limits? What is the pathology of trust; when is trust ill-founded? When does trust become counterproductive? What are the implications?

- The figures of trust, in Chapter 6. How can we measure it? How can we test its effects in empirical research? How can we model the process of making and breaking trust, under conditions of uncertainty? How can we deal with the limits of trust? Can trust survive under competition? Here, use will be made of the empirical literature, including work of my own (Nooteboom, Berger and Noorderhaven 1997), and of a recent application of the methodology of ‘Agent Based Computational Economics’ (ACE) for the simulation of the making and breaking of trust (Klos 2000, Klos and Nooteboom 2001).

- Conclusions and directions for further research will be given in the final Chapter 7.

1.2 Theoretical tools

This paragraph discusses some of the theoretical tools for the analysis of trust. I noted before that I would use elements from transaction cost economics, while recognising and taking into account the fundamental criticism that has been raised against that theory. Also, I indicated that we need a theory of knowledge, for several reasons. One is to mend the shortcomings of transaction cost economics. A second is that knowledge and learning have implications for the importance, the basis and the functioning of trust. The theory has implications for the theory of the firm, and for the relation between trust in people and trust in organisations.

1.2.1 Transaction cost economics

I will not summarise transaction cost economics, and assume that it is known. One objection to the theory, raised by many, is precisely that it neglects trust in its point of departure that in the ‘governance’ of relations we need to safeguard against opportunistic behaviour. I do not wish to
neglect the possibility of opportunism, but I do wish to recognise that next to the possibility of opportunism there is also the possibility of trustworthiness, and that neither should be neglected. Transaction cost economics does not claim that everyone is opportunistic, but that one cannot reliably assess the degree of opportunism so that the possibility of opportunism should be taken into account. There are three objections to this view. One is that it neglects reputation, which, if it works properly, yields information on the trustworthiness of an agent to strangers. A second objection is that it is possible to develop an informed opinion on trustworthiness, in the course of a relationship. How this is to be done will be treated in Chapter 3. A third objection is that opportunism or trustworthiness are themselves subject to change as a function of how a relationship develops, in what Zucker (1986) called 'process based trust'. This also will be discussed later, in Chapter 3.

A second objection to transaction cost economics, also recognised by many, is that it neglects innovation and learning. It does so by Williamson’s own admission (1985: 144-145). That neglect is unacceptable at a time when innovation and learning seem to form the very core of what is going on in economies, markets and organisations. To accommodate this, I need a theory of knowledge. That is not the focus of this book, but I will summarise the theory in the following section. It is important to take a dynamic view, including innovation and learning, because there trust is especially important, due to the importance of uncertainty in innovation, and the resulting limitations of contracts in the governance of relations.

A third objection is raised against the very idea, present in transaction cost thinking from the beginning, in the work of Coase (1937), that one could have transactions without costs, and that the carriers of transaction costs only form liabilities, while in fact they not only constrain but also enable transactions (Campbell and Harris 1993). Relations require social capital, and that carries costs. This point comes up most clearly when we do incorporate issues of knowledge and learning. If for a transaction, or more generally economic interaction with others, one needs to make oneself clear and transfer knowledge, concerning the product one has to offer, or the capabilities that one could contribute to the relation, then this carries a cost, but it also enables the relation. Without that cost the relation would not be fruitful. Building trust costs time and carries a cost, but enables relations that would not otherwise be possible, and may yield lower transaction costs than other means of governance. A similar oddity arises in the view one often encounters that institutions constrain behaviour, while in fact they also enable it. As Nelson and Sampat (1999) put it: it is odd to see a road across a swamp as a constraint.

Nevertheless, also enablers carry costs, so that it still makes sense to talk about the minimisation of transaction costs, or the costs of institutions, not by eliminating the facilitators of transactions but seeking ones that are effective and yet carry low costs. For an example, I might proceed with the earlier example of the transfer of knowledge. Transfer of knowledge is particularly difficult, time consuming and costly when the knowledge is tacit, i.e. not codified and documented in a formula, blueprint, algorithm or standard operating procedure. Thus, there is a premium on trying to document it, and thereby reduce transaction costs, while preserving the value of the knowledge. The question then is whether it is possible to codify tacit knowledge without loss, and if not, what the trade-off is between loss of knowledge and ease of transmission.

A fourth and perhaps most fundamental objection to transaction cost economics concerns the way it deals with the issue of radical uncertainty, indicated before. Williamson (1999) claims that he fully accepts bounded rationality: there is fundamental uncertainty concerning future contingencies. However, he claims, there is foresight: one can take such uncertainty into account, infer the hazards that follow from it and conduct governance accordingly (in a ‘discriminating alignment’) and ‘efficiently’, i.e. in an optimal fashion (to yield an ‘economising result’). We are not myopic, Williamson claims: we are not so stupid as not to take uncertainties into account when we design governance. And indeed, we can to some extent take risks and uncertainty into account. Firms can spread risks by participating in different markets and technologies, similar to the way that investors can spread risks in a portfolio of investments. Beyond that, to deal with real or radical uncertainty we can construct scenario’s of possible futures, prepare contingency plans for them, and identify the robustness of
strategies across different scenario’s. However, of course the question arises what the implications of bounded rationality are for the correct identification of relevant hazards. Doesn’t bounded rationality imply that we might be mistaken about them? Later in his 1999 article, Williamson (1999: 1103) admits that TCE ‘makes only limited contact with the subject of learning’, and indicates that we may be mistaken about hazards and may learn about them as events unfold (1999: 1104). And apart from hazards I add: how about new options? In spite of great imagination and ingenuity, the scenario’s we invented may not include what actually arises. And how about shifts in our preferences? Is it reasonable to assume learning without shifts in preferences? And if new insights in hazards arise, and new scenario’s, or new options or goals, are we then able to shift from the governance structure engaged upon to an adapted, optimal form? That would always be possible only if there is no path-dependence or lock-in in governance, and that is a very strong claim to make.

This is related to the issue of ‘efficient’, optimal outcomes. Williamson’s argument is that ‘dysfunctional consequences and other long run propensities will not be mindlessly repeated or ignored’ (Williamson 1999: 1105). But this begs a number of questions. It implies that dysfunctionality and long run propensities are stable, so that experience in the past is indicative of the future, and it implies that we know which are stable and which are not. How can we know that? And if we did know, how can one be sure that the firm survives to implement the lesson in time? Or is Williamson now falling back on the notion of selection: inefficient forms of organization will be selected out by ‘the market’? That is the usual assumption behind the economist’s assumption of efficient outcomes, going back to Alchian (1950). But if that is Williamson’s argument, he is deviating from the perspective of the firm strategist, who is talking about the survival of the firm.

As noted by Chiles and McMackin (1996), there are two perspectives in transaction cost economics. The first is the long-term evolutionary perspective, where objective transaction costs determine the survival of the fittest governance forms. The second is a short-term managerial choice perspective, where managers act on subjective costs which are based on varying perceptions and evaluations of risk. The latter explains why firms in similar circumstances may make different trade-offs in their make-or-buy decision.

If in the defense of efficient outcomes we fall back on the evolutionary perspective, I note that the selection argument was already shown to be weak by Winter (1964). In selection it is not the best possible but the best available in the population that survives. In the presence of economy of scale inefficient large firms may push out efficient small firms, and thus inefficiency may survive. Furthermore, the argument assumes efficient selection, but that cannot be taken for granted in view of possible monopolies, entry barriers and transaction costs.

However, in spite of all the criticism of transaction cost economics it still contains elements that are useful. As already indicated, it still makes sense to talk of transaction costs. Also useful is the notion that relations may require investments that are specific for a relationship, thereby create switching costs, and hence create dependence, with a resulting risk of ‘hold-up’, which requires attention in the governance of relations. While the theory neglects trust, it does yield insights in possible instruments of governance, next to contracts, such as a balance of mutual dependence, the use of hostages, and the use of third parties in ‘trilateral governance’ (none of these, however, was particularly new). Also still valid is the consideration that integration in an organisation may help to solve problems of dependence. Even though there are problems of asymmetric information, motivation, and control also under ‘hierarchy’, the opportunities for demanding information and imposing a settlement of conflict still exceed those one would have with regard to an independent outside organisation. Ultimately, such demands can only be enforced in front of a judge, while within a firm there is much more scope for it under the employment relationship.

1.2.2 Knowledge and cognitive distance

It is a truism to say that information is not the same as knowledge: to become knowledge, information needs to be interpreted in a cognitive framework. Here, I employ a theory of knowledge and language
derived from ‘symbolic interactionism’ in sociology (G.H. Mead), and the view, taken from cognitive psychology, that intelligence is internalized action (Piaget, Vygotsky). Rather than the dominant ‘computational representational’ view in cognitive science, this leads to the view of knowledge in terms of ‘situated action’. For a more detailed analysis, see Nooteboom (2000a). I take ‘cognition’ in a wide sense, including perception, interpretation and evaluation, hence including value judgement. In this view, emotions and intellect are connected.

The theory states that cognition takes place on the basis of cognitive categories that are developed in interaction with the physical and social environment. Those constitute our absorptive capacity (Cohen and Levinthal 1990). This precludes objective knowledge (or at least any certain knowledge whether or to what extent knowledge is objective). Our absorptive capacity is context- and path-dependent. However, this view does not necessarily entail the radical relativism or subjectivism that is exhibited by many postmodern authors. If we construct knowledge from interaction with our environment (of which it is reasonable to assume that it does exist), this entails that reality is at least a material cause of our knowledge: our knowledge is ‘embodied realism’ (Lakoff and Johnson 1999). Also, to the extent that we share the environment in which our knowledge develops, there will be similarities of cognition, which yields a basis for inter-subjective debate. Since we share the physical environment (including laws of physics) to a greater extent than we share our cultural environment, knowledge constructed from the former is likely to be more common than knowledge constructed from the latter.

The reverse of this coin is, of course, that to the extent that people have developed their knowledge in different environments, and have not been in communication with each other or each other’s environment, cognition will differ: there will be greater or lesser ‘cognitive distance’ (Nooteboom 1992, 1999a). I do not wish to imply that cognition is any simple, one-dimensional construct that allows for simple measurement of distance. In Nooteboom (2000a) I elaborate on cognitive distance in terms of overlap of, or mappings between, different sets of cognitive constructs.

Cognitive distance yields both an opportunity and a problem. The opportunity is that contact with others gives us an opportunity to escape from the myopia of our personal cognitive construction. A problem, however, is that the greater the distance, the more difficult it is to cross it, i.e. to understand the actions and expressions of a partner. Thus there is some optimal cognitive distance: large enough for partners to tell each other something new, and small enough for comprehension. Absorptive capacity is part of our ability to cross cognitive distance. The other part is communicative capacity, or the ability to help others understand what we do or say. Note that there is a difference between crossing cognitive distance and reducing it. For an elaboration see Nooteboom (2000a). Neither absorptive capacity nor communicative capacity is fixed. Depending on the type of knowledge involved they can be increased by R&D and experience in crossing cognitive distance. An implication for relations is that investment in absorptive or communicative capacity can be relation-specific and can thereby entail dependence on a relation in order to recoup the investment, along the lines argued in transaction cost economics.

We should also take into account that next to cognitive considerations in the narrow sense, there are also more emotion-laden considerations from social psychology. In particular, absorptive capacity can be limited due to cognitive dissonance: we may actively but subconsciously resist information that is in conflict with established and cherished views or convictions, particularly if it would require and admission of mistaken choices in the past. Past acts have to be justified to oneself and to others, even at the cost of distorting facts or construing artificial arguments. Deutsch (1973: 357) quoted Festinger as saying that:

‘Rats and people come to love things for which they have suffered ..(they) enhance the attractiveness of the choice that led to the suffering’.

For an example, Deutsch pointed to the American involvement in Viet-Nam. The heavier the casualties, the more the importance of being there had to be inflated. As shown by Deutsch, social-
psychological considerations are particularly important to explain the pathologies of trust: of maintaining trust or suspicion in spite of contrary evidence. This is further discussed in Chapter 4.

1.2.3 Implications for organisation and relations

The analysis leads to the notion of organization as a focusing device (Noooteboom 1992). An important function of organization is to create sufficient focus, i.e. alignment of cognitive categories, for people to achieve a common purpose. Arguably, this is more fundamental than the function of a firm to reduce transaction costs, as transaction cost economics proposes. Now, focusing, here again, yields a problem of myopia, by which organizations may fail to see or adequately interpret potential opportunities and threats to its existence. To compensate for myopia, organizations need outside partners for complementary cognition, or ‘external economy of cognitive scope’ (Noooteboom 1992). That yields the cognitive argument for inter-firm relations hinted at before.

The sharpness or narrowness of the focus depends on whether the organization, or part of it, needs to concentrate more on efficient exploitation (utilization of existing resources and competencies, including cognitive competencies) or on exploration (development of new competencies). The former requires a sharper focus, with more unity of perception and interpretation for the sake of efficient co-ordination, while the latter requires more diversity and volatility of linkages for the sake of finding Schumpeterian ‘novel combinations’.

The literature on organizational learning states that organizations need to do both: to combine exploitation and exploration. This not a trivial task. Typically, efficient exploitation requires division of labour, with clear and sharp linkages, on the basis of clear, unambiguous meanings and standards, while exploration entails crossing and shifting boundaries, linkages and meanings. How can they be combined? One can think of several solutions, and one can observe a variety of attempts. The innovation literature proposes that exploration yields novel forms, which converge on a dominant design, which is used as a basis for exploitation. Elsewhere (Noooteboom 2000a) I proposed that next the novelty is carried into new applications, in which it is challenged with failures, which motivate adaptations and yield opportunities and inspiration for experimenting with novel combinations. Here, exploration is combined with exploitation. Novel elements put the existing architecture under strain. That architecture constrains the utilization of the potential of novel elements, and pressure arises to engage in more radical architectural innovation (Henderson and Clark 1990), and here we are at the beginning of the cycle again.

Thus, up to a point exploitation and exploration can be combined. However, exploitation remains connected with a more or less integrated structure, in the sense of strict linkages between elements, or fixed standards, for some activities. Exploration requires more or less disintegration, with rupture of existing connections and standards.

One solution is to let an organization fluctuate between integration and disintegration, as needed for development to proceed, and we often observe that (ATT, Apple). That is not an option if the organization embraces a mix of activities in different stages of development. A solution can then be a network of organizations, or parts of organizations. That may be the underlying reason for the emergence of the ‘network economy’. This connects with the need, identified earlier, to combine cognitive distance, for the sake of exploration, with cognitive proximity, for the sake of exploitation. In a network one can have some units that concentrate on exploitation and some that concentrate on exploration, and activities are shifted between the units as they move through the cycle of development. Part of exploration then also lies in changing configurations of units. Not all units are used in all configurations. It can be an attractive position to act as orchestrator of activities in such a network. Of that we also see several examples (Benetton, Nike).

Absorptive capacity is not fixed. When the knowledge involved is codified, absorptive capacity can be increased and maintained by more formal, declarative forms of learning, such as R&D. Often, when firms outsource certain activities, they maintain R&D in that area in order to maintain absorptive capacity (Granstrand, Patel and Pavitt 1997). When knowledge is tacit, I propose that absorptive
capacity depends more on cumulative experience in absorbing tacit knowledge from a variety of sources. Then communicative capacity also matters more, to help partners to absorb tacit knowledge.

This cognitive theory of the firm and of inter-firm relations helps to incorporate innovation and learning in the analysis of inter-organisational relations, and the role of trust in them. This analysis will be taken up in Chapter 4, in an extensive discussion of the economic functions of trust. The notion of the firm as a focusing device is also taken up in the discussion of the relation between trust in people and trust in organisations, in Chapter 3.
CHAPTER 2: FORMS

This chapter discusses the definition of trust and different forms of trust, depending on what the object of trust is: what can we have trust in? Is trust a type of behaviour, or a disposition underlying such behaviour? How is trust related to risk? Can it be defined as an expectation? As a subjective probability? How is trust in people related to trust in organisations?

2.1 Definition of trust

2.1.1 Disposition or behaviour?

In some literature, trust is seen as a form of behaviour. This is found, in particular, in economics. Consider, for example, the well known book on ‘Economics, organization and management’ by Milgrom and Roberts (1992).

The index of the book indicates that trust is discussed on three pages out of 605. There, trust is treated in a game-theoretic framework as something that can be offered or withheld, and the partner in the game then has the options to honour trust by behaving fairly, or not to honour it. Trust here means behaviour that suspends the demand for contractual safeguards against relational risk. In a single-shot game there is a temptation for the second party not to honour such trusting behaviour, and to profit from opportunism. As a result the first party, anticipating that, will not offer trust. In a repeated game, however, reputation effects, combined with costs and imperfections of writing contracts, can yield the offering and the honouring of trust as an equilibrium outcome. The analysis is useful to demonstrate the working of a reputation mechanism. However, from the perspective of social psychology it seems a little odd to consider trust as something one can give or withhold. Isn’t it rather something that one has or not? This riddle can be resolved by distinguishing trust-based behaviour, as manifested by the suspension of contractual or other safeguards, from trust as the disposition to such behaviour. If one has trust, in that sense, this may be the basis for offering trusting behaviour, with the suspension of the demand for safeguards.

In social science also, trust has been defined in terms of behaviour. Consider, for example, the definition of trusting behaviour given by Deutsch (1962 quoted in Zand 1972: 230):

‘Trusting behaviour consists of actions that (1) increase one's vulnerability (2) to another whose behaviour is not under one's control (3) in a situation where the penalty one suffers if the other abuses that vulnerability is greater than the benefit one gains if the other does not abuse that vulnerability.’

There are also definitions of trust as the underlying disposition towards behaviour. Before, I quoted Sako (1992: 32), with her proposal that:

‘Trust is a state of mind, an expectation held by one trading partner about another, that the other behaves or responds in a predictable and mutually expected manner’.

Consistent with Sako’s definition we have the definition offered by Bradach and Eccles (1989: 104):

‘a type of expectation that alleviates the fear that one’s exchange partner will act opportunistically’.
However, here one might ask whether it is the expectation that alleviates fear or the reduction of fear that yields the expectation. If we define trust as an expectation or disposition, the question is on what that is based. That is the subject for Chapter 3.

2.1.2 Trust, risk and probability

Many authors have connected trust with risk, not in the (financial) economist’s sense of a variance of a probability distribution, but in the sense of what the financial economist would call ‘downside risk’: the possibility of loss (Luhmann 1988, Gulati 1995, Chiles and McMackin 1996, and many others), while this loss may be greater than the gain one may derive from the relation (Deutsch 1962, 1973, Zand 1972, who builds on Deutsch’s view).

Proceeding from that, trust has been identified with a subjective probability of an action that it is ‘beneficial or at least not detrimental’ (Gambetta 1988: 217). Similar views were adopted by Dasgupta (1988), Mayer, Davis & Schoorman (1995), Gulati (1995), and others. The advantage of that notion is that if trust is a probability, it can be subjected to probabilistic calculation, in a Bayesian procedure in which subjective probability is adapted in the light of new evidence. However, that raises a puzzle. The notion implies that there is high trust if the subjective probability of good results of the action is high, and hence perceived risk is low. That implies that trust is highest when the probability is unity, so that there is no risk left. However, another intuition is that trust entails a wager, an acceptance of risk, so that if there is no risk, we can no longer speak of trust. Pettit (1995) objected to this conclusion: we can have trust even if we are certain that objectives will be achieved.

Pettit gives the following example. Suppose that in an unknown town you wanted to get to the city centre by car. One approach would be to follow a bus with the centre indicated as its destination. Can you trust the bus driver? Does it matter whether or not you tell him that you are following him, and why? Pettit introduced this example to discuss the important concept of ‘trust-responsiveness’: how do people respond to the trust of others? I will return to that later.

In the present context, Pettit’s point is that you may be virtually certain that the driver will indeed take the bus to the centre, regardless of what you do, since that is his duty. Nevertheless, you can be said to trust him to do that. And indeed, if trust is based on certain considerations, then if they lead to certainty, you do not lose those considerations, and in that sense you do not lose your trust. While Pettit does not discuss this, the announcement to the bus driver may matter, as follows. If he is malevolent, for whatever reason, he may spurt through an orange light to leave you locked and lost at the red light. If he is benevolent, he may stop at the orange light. So, can you ever be 100% certain that the driver will behave as you hope he will? Nevertheless, Pettit’s argument is clear: it is odd to say that one no longer has trust when one is certain.

So, should we accept the view that trust is to be seen as a subjective probability concerning a positive outcome of another’s behaviour? Earlier, like many others I was inclined to sympathise with this view (cf Nooteboom 1999a, 1999c), but now I reject it, for several reasons. Perhaps the most important reason derives from the puzzle indicated above. Subjective probabilities can become unity, if the concept means anything, and that entails certainty, while there is a persistent intuition that there is always residual risk; that we can never be completely certain about another’s behaviour. One reason for this view is ethical, perhaps: such certainty would yield the possibility of perfect predictability of behaviour, which would eliminate freedom of action. A second reason is that certainty concerning behaviour would underestimate radical uncertainty, for both the trustor and the trustee, concerning future contingencies that may affect choices. This refers back to the discussion of radical uncertainty in Chapter 1. As has been argued before by Savage, the notion of probability, in any usual sense of satisfying the axioms attached to probabilities with a frequency interpretation, cannot apply under radical uncertainty. There, the set of all the things that could happen is not closed, and unpredictable
possibilities arise as we go along. If we talk of beliefs under radical uncertainty, we should drop the term ‘probability’. This resolves the puzzle. Perhaps we can talk of trust as an expectation, but not as a probability, because there will always be residual uncertainty about agency and unforeseeable contingencies, so that there is always a residual element of risk, in the sense of a non-quantifiable possibility that things go wrong. This yields an argument against blind or unconditional trust: one should be aware that there is always a possibility of things going wrong. To be 100% certain about someone’s future behaviour is to rob him of his freedom. No matter how loyal a friend is, one cannot expect him to remain loyal and not to betray you under any conditions, such as, say, torture.

For an illustration, let us return to Pettit’s example of the bus driver. Suppose I tell him that I will follow him to get to the centre of the city, and he pledges to take that into account and to help me, and I have offered him a substantial sum of money to lead me to the centre, to be paid upon arrival, or he is my uncle or a close friend, and I know that to keep his job he cannot afford not to go to the centre. I still cannot be sure. I may still lose him at a red light, or in the chaos of traffic (this situation happened to occur in Rome), because he simply could not help losing me. Perhaps he could only stop in time for a light about to jump to red, or avoid me getting jammed in, at the cost of causing a crash. And that he could not afford, in view of his job and his responsibility for his passengers.

A second reason to reject the notion of trust as a subjective probability is that it seems too rational and calculative, and thereby excludes trust based on instinct, possibly irrational convictions, and routinised behaviour (Nootboon 1999a, 1999c). This yields an argument against trust always being calculative. Apart from a positive expectation that things will not go wrong, one may simply not think of things going wrong, or not be aware of that possibility, or ignore it. Trust can thus be non-calcualtive. Even so, it can be rational, for example when it is based on routinised behaviour. Williamson (1993) suggested that non-calcualtive trust inevitably yields blind trust, but this is not so. Trust that is blind, i.e. unconditional, would indeed be very hazardous. However, routines may be subject to tolerance levels. As Herbert Simon has taught us, in view of bounded rationality routinised behaviour is rational. It allows us to focus our limited capability of attention and rational evaluation on new things that have priority, while for other activities we adhere to unreflective routines that have proven to work satisfactorily in the past. In the present context, we may trust someone simply because our relation has worked out satisfactorily in the past, or because we have never been duped in other relations. Of course, this expectation is more well-founded to the extent that this has happened under varying conditions, including ones that required a sacrifice from the partner. Later, in Chapter 3, using the work of Deutsch (1973) I will analyse in more detail how one can infer trustworthiness from observed behaviour.

Of course, the danger of routinised behaviour is that it is kept up under new conditions which make it fail. That is why we will often retain a ‘subsidiary awareness’ of risk, which is shifted into ‘focal awareness’ when phenomena exceed a certain tolerance level (Polanyi 1962). Emotion, called forth by such intolerable events, serves to break us out of the routine and trigger focal awareness, and put it on the agenda, so to speak (Simon 1983). This is part of the rationality of emotions: they catapult us from subsidiary into focal awareness. For other aspects of ‘emotions within reason’, see Frank (1988). The difference between routines and instinct is that the former are in this way susceptible to change, whereas instincts are hereditarily acquired and are ‘hard-wired’ into our cognitive and physiological make-up. This distinction is probably a bit too stark. Some routines may become so entrenched as to cause virtually irredeemable inertia.

Neglect of risk may also be based on naivety, or on unwillingness to face risks due to fear. It may be based on cognitive dissonance related to friendship or kinship: one does not want to contemplate the possibility of opportunism on the part of friends or family. Particularly in the latter case trust is to a greater or lesser extent based on affect. But also ethics based trust can be related to affect: emotions connected with the notion of justice or decency.
Summing up, there is nothing in the basic notion of trust that indicates that it must be justified or calculated. To allow for a non-calculative basis of trust, such as routinised behaviour, naivety or even sheer stupidity, the definition of trust should be expanded to include the unreflective neglect of possibilities for things to go wrong. This yields the following definition of trust:

‘Trust in things or people is the expectation that they will not fail us, or the neglect or lack of awareness of the possibility of failure’.

Note that here I allow for trust not only in people but also in ‘things’. By that I mean objects, institutions, organisations, and the like. I will expand on that later.

2.1.3 Reliance and trust

The definition of trust given so far, which includes the expectation that things or people will not fail us, leaves open a wide range of foundations for such trust, in the form of beliefs or tacit assumptions that a variety of possible causes of failure will not arise. Such beliefs may be based on contractual safeguards. This will be discussed more systematically in Chapter 3. However, here the point is the following. As noted in Chapter 1, this raises the question whether expected conformance to expectations on the basis of contractual obligations can properly be called trust. The notion of trust in everyday language seems to suggest that trust goes beyond contractual obligation, and refers to expected cooperation even when it is not specified in a contract, and even when it may require a sacrifice from the partner. Thus, trust has been defined as the expectation that a partner will not engage in opportunistic behaviour, even in the face of countervailing short-term incentives (Bradach and Eccles 1984, Chiles and McMackin 1996).

Later, in the analysis of the functioning of trust in the governance of relations, in Chapter 4, I will employ the notions, developed in earlier work (Nooteboom 1996, 1999a), of ‘opportunities, incentives and inclinations to opportunism’. Opportunities arise from openings for action left in contractual agreements and legal regulations, or imperfect information for monitoring compliance with them. Incentives refer to self-interested motives to utilise opportunities for opportunism. If an agent is himself highly dependent on the relationship, or is concerned with his reputation, or with the possible loss of hostages, then he may lack such incentive. Now, trustworthiness as intended in ordinary language can perhaps be defined as the renunciation of opportunism in spite of both opportunities and incentives, based on limited inclinations towards opportunism. That can be based on many things, such as friendship, empathy, ethics, or routinisation, to be discussed more systematically later, in Chapter 3. This leads to a narrower, stringer definition of trust, as follows:

‘The expectation that things or people will not fail us, or the neglect or lack of awareness of the possibility of failure, even if there are perceived opportunities and incentives for such failure’

Thus trust in this narrow or strong sense does not include coercion or self-interest as a basis. There is a source of possibly disastrous misunderstanding here. If in some exchange situation actors profess that they can be trusted, very different things can be meant. Such differences are likely to occur between different cultures and languages (cf. Sako 1998). One may mean trust in the wide sense, including all kinds of safeguards. One says one can be trusted because contractually one has no other option, or it is in one’s interest to stick to the intention of agreements. One may also mean that one will remain loyal and renounce opportunities for opportunism when they occur, due to gaps or loopholes in the contract, even when there are incentives to utilise them, unless temptations or pressures become excessive, and then one will announce the problems and try to solve them together. Misunderstanding about what is meant here can be disastrous.
Therefore, for the sake of clarity, and to prevent misunderstanding, perhaps we should reserve the term trust for the narrower, stronger notion, and use the term ‘reliance’ for the broader notion that encompasses all foundations of expectations, including contracts and self-interest. However, this might become somewhat laborious. So, in the following, I may still use the term trust for the broader notion of reliance, not to seem pedantic, but when it matters, I will pick up the distinction between trust in the broader sense of reliance and in the stronger sense of ‘proper trust’.

2.2 Objects of trust

What are the objects of trust, i.e. what can one have trust in?

2.2.1 Causes of failure

Previously, a distinction was made between trusting behaviour and the dispositions underlying it. A similar distinction arises when we consider the object of trust, or the thing we can have trust in. We can have trust in someone’s behaviour, in *behavioural trust*, or in the dispositions underlying that behaviour, in *dispositional trust*. The question then is in what disposition we have trust, as a basis for trust in the behaviour it yields. If trust is an expectation that things or people will not fail us, we can distinguish between different causes of failure. One is lack of *competence*. Another is lack of intention to utilise competence, i.e. lack of *dedication* and care. Another is *opportunism*, i.e. the disposition to take advantage of the partner at his expense. Yet another is *external conditions* that cause a mishap or accident. These yield different objects of trust. Thus we should make a distinction between *competence trust*, *intentional trust*, which includes dedication trust and benevolence or goodwill trust (trust that behaviour will not be opportunistic), and *circumstantial confidence* (confidence that external conditions will not negatively affect outcomes). The latter entails the often tacit assumption that no freak accidents will occur that affect outcomes against favourable conditions of competence, dedication and benevolence.

For an example of the difference between competence, dedication and goodwill trust, consider the following situation. For building roads through mountains there is a team for blowing holes in mountains with explosives. Members of the team have to rely on each other for their very survival. This entails that people know what they are doing (competence concerning the amounts and location of explosives, connection of wires, timing of the explosion, etc.), are loyal towards their colleagues and have no intentions of harming them (benevolence), and will pay attention meticulously to proper procedures to ensure safety (dedication). Circumstantial confidence here entails the confidence that no accidents will occur such as, for example, a colleague having a heart attack at some crucial moment. A more subtle point then concerns one’s reliance on a colleague to take proper remedial action if an accident does occur. This concerns his competence to improvise, and his dedication to do so, which may require courage, and his benevolence in willing to take risks and make a sacrifice for the benefit of his colleagues. Another example is trusting someone with one’s car. This may refer to driving competence, or to intentions to drive carefully and maintain the car well, or to not stealing it. Trust in competence will depend on the tests to which it will be put. One may trust someone’s competence in easy driving in good weather and relaxed traffic, but not in conditions of sleet, in heavy traffic or at night.

The distinction between these objects of trust is important, because their breach requires different responses. If lack of competence is the cause, one may respond by bolstering that competence, by support or training, or to subject it to less demanding conditions. If lack of dedication is the cause, one may think of further incentives, an appeal to sense of duty, or threat of replacement. If opportunism is the cause, one may think of penalties, retaliation, replacement or threats of any of these,
or an appeal to ethical behaviour or social duties of reciprocity. When confidence in external conditions is broken, one’s action will depend on whether a disturbance is seen as incidental or as systematic. When it is an incidental, freak accident, one may have to just accept the loss and perhaps share it among partners. When the disturbance is systematic, one may make an adjustment by taking out an insurance, or by diversifying the risk, or revising the terms or conditions of collaboration to make it more robust under such accidents. The problem of course is that if opportunism is the real reason, the perpetrator will not admit that if he can get away with it. He will claim force majeure, and if that cannot be made plausible, he will claim a shortfall of competence, with the promise to catch up, and he may ask for a little help. In view of this uncertainty, trust in truthfulness or honesty also enters the scene. I call this honesty trust. It is part of benevolence or goodwill trust, but deserves a special place. It is more important to the extent that causes of unmet expectations cannot be observed. This raises the question how we can infer dispositions from observed behaviour. That will be discussed in Chapter 3, where I will also give a further analysis of the role of honesty.

2.2.2 Variety of objects

In the previous definitions and discussions of trust, the implicit assumption was that trust concerns the behaviour of people, notably partners in exchange. I discussed honesty trust: the trust in someone’s truthfulness. This is a special case of trust in information, or informational trust: its factual truth, accuracy, completeness, impartiality, up-to-date-ness, accessibility, interpretation, relevance, authenticity, confidentiality, security. These are related to trust in the competencies and intentions of information providers. One can ask what the implications of INTERNET will be for the trustworthiness of information, in view of the low threshold for information provision and anonymity of providers. There is a need for certification, accreditation, licensing.

There can be trust in many other things. One can have trust in objects: one’s car, for example, with an expectation of performance. Here also, there are objects of trust that underly expected performance. In the case of the car: its engine, gear, brakes, etc. Objects are less interesting since they have no life or will of their own. Trust associated with the actions and motivations of people, i.e. behavioural trust, is more complicated, interesting and important. When the performance potential or quality of objects is difficult to judge, trust in objects may shift to trust in the provider of the object. In the case of the car this would be the car dealer and the garage.

One can have trust in laws of nature and their consequences, such as the weather. One can have institutional trust, i.e. trust in institutions such as the law, police, or government.

A definition of institutions, or part of such definition, is that they enable and constrain behaviour. I recall the discussion of the criticism of transaction cost economics, in Chapter 1, that it tends to see transaction costs in terms only of obstacles, rather than also enablers of exchange (Campbell and Harris 1993). It is odd to see a road across a swamp only, or even primarily, as a constraint, while it does of course constrain where you walk (Nelson and Sampat 1999). Another feature of institutions is that they are durable and in a sense inevitable, or not subject to choice: we are subjected to them. This, in turn, relates to the discussion, in social science, between structure and agency. It is true that institutions that form the structure in which agency takes place, and which both enables and constrains action, is man-made. However, that is a collective outcome of individual actions, and takes time. At any moment, institutions are given and more or less inevitable to individual action (Archer 1995).

This relates to the useful distinction that Luhmann (1988) made between confidence and trust. Confidence relates to bigger or wider systems or entities that we can hardly influence and that are more or less inevitable, such as God, the law, police, government, etc. In other words: institutions. When something goes wrong, we cannot reproach ourselves for relying on the institution, because we had little or no choice. If we have no confidence, we cannot exit. We can only sit back and hope for the best. Here is the reason that previously I used the term ‘confidence’ for trust in the external
conditions that affect outcomes of human relations. That is properly called confidence, as defined here, in that one cannot blame oneself for relying on the absence of freak accidents where it was reasonable to assume stable conditions. Regarding a judge, we would usually speak of confidence (or the lack of it), because we are not usually in a position to choose a judge or influence his judgement, and can only submit to what is imposed on us. If we can choose or bribe a judge, we might speak of trust. Regarding God also it would seem appropriate to talk of confidence. It would seem hubris to talk of trust, as if we could dodge his influence. Perhaps we could trust the gods of classical Greece, who were much like humans, and could be bribed, seduced or played off against each other. Summing up: we can talk of institutional trust, but mostly that will be a matter of confidence rather than trust, so that it would be better to speak of institutional confidence.

Institutions also include basic cognitive categories, in the wide sense of including perception, interpretation and evaluation. This is close, if not identical, to Veblen’s notion of institutions as ‘settled habits of thought that apply to the generality of people’. These satisfy the conditions of institutions that they enable (without them we cannot think), and constrain (perceiving in one way excludes perceptions in other ways; one cannot look in all directions at the same time) thought, and thereby behaviour. They are also more or less durable and inevitable, i.e. not subject to choice. This follows from the theory of knowledge discussed in Chapter 1. In interaction we construct our cognitive categories, on the basis of which we perceive, interpret and evaluate the world. Thus they are contingent upon the physical end social environment, including the associated institutions of society. We cannot freely choose fundamental categories. Thus we can only have confidence in them. However, on the basis of such categories we think. Such thought is subject to argument and other forms of learning, and there is room for choice in opening up to arguments, and making the effort to follow and understand them, and to learn the capabilities of doing so. There is more scope for choice here than there is concerning fundamental categories that support our thought, and other institutions.

We can speak of self-confidence, self-trust and self reliance. If we follow the previous definitions, self confidence would apply when we are inevitable to ourselves: we cannot change and can only hope for the best. This might apply to instincts, cognitive limitations or addiction. Cognitive limitations include the fundamental categories that underlie thought, as discussed before. Self reliance would apply if there is choice, and trust in our own behaviour is based on a variety of safeguards: duty, conviction, supervision, coercion. Self trust would apply if we trust our behaviour in the absence of coercion, and against temptations and the weakness of our will. For example, we are resolved to engage in a slimming program, and need not ask our partners to hide the fattening delicacies or block ourselves and our partners from buying them. We need not surrender our car keys at a party to keep ourselves from drunken driving. This relates to an old definition, in philosophy, of freedom not as lack of constraint but as voluntary restraint. Here I would say: self-trust. We are free if we have a choice to sin and choose not to do so.

This connects with the view, discussed before, that certainty concerning another’s behaviour, with a subjective probability equal to unity, would deny his freedom.

2.2.3 Institutions, organisations and culture

Can we have trust in organisations? Is that behavioural trust in individual people, or is it institutional trust, or institutional confidence?

Much has been written on the question whether an organisation is an institution. North (1990) defined institutions as the ‘rules of the game’. Originally, he emphasised the constraints that they impose, on opportunism, for example, thereby reducing transaction costs. In his view organisations are not institutions but players confronted with institutions. There is merit in this idea: it is useful to distinguish organisations from the institutional environment in which they operate. In transaction cost economics, by contrast, organisation is seen as an institution, as an alternative to ‘the market’ (whatever that is), on the grounds that it can help to reduce transaction costs. The implicit definition of institution then is anything that reduces transaction costs. We can reconcile these differences by
making a distinction between organisation as an ‘institutional arrangement’, and its ‘institutional environment’ (North and Thomas 1973). The latter conditions, i.e. enables and constrains the former.

This seems to correspond with the distinction between culture in sociological sense (habits, lifestyles) and culture in the anthropological sense (underlying norms and values). The former seems descriptive (how people behave) and the latter normative (how people are guided or constrained to behave). However, the two are intertwined: established practice serves as a norm, in routinized behaviour. The notion that institutions are rule - like, in the sense that they are established and that they govern conduct remains appropriate.

Earlier, I adopted the definition of institution as anything that enables and constrains behaviour, and is enduring and not subject to free choice. According to that definition, is an organisation an institution? It does enable and constrain behaviour. Earlier, in the discussion of the implications of theory of knowledge for the notion of an organisation, in Chapter 1, I characterised an organisation as a device for focusing cognition: perception, interpretation and evaluation. To function as such, the focus cannot be too variable and non-committal. In other words, it must have some stability and requires a certain engagement. Thus it does resemble an institution. However, the focus is not strictly inevitable and devoid of choice. One can decide to leave an organisation, depending on the alternatives one may have. Within an organisation one may be able to deviate from the established focus of cognition. In fact, one would hope that people within an organisation do that to a greater or lesser extent. If the focus of the organisation is too tight, this yields too much ‘group think’, with too little cognitive distance, which jeopardises the variety that is needed for exploration, i.e. for innovation and learning in the sense of discovery. So, the tightness of organisational focus is to be decided as a trade-off between on the one hand cognitive proximity needed for efficient exploitation and on the other hand cognitive distance for exploration. As discussed in Chapter 1, one solution here may be to utilise outside relations with other organisations, at a larger cognitive distance, to complement the narrowness of focus needed for efficient exploitation. So, here also I find that an organisation is like an institution but is less inescapable and more subject to choice and influence by individuals than other, ‘larger’ institutions.

Elsewhere (Nooteboom 2000) I proposed not to define institutions in terms of an inventory of institutional entities, but in terms of a functional definition: institutions enable and constrain behaviour, and are stable and engaging with respect to that behaviour. What the specific institutions are depends on the particular situation we are investigating. Legal systems are institutions for organisations, and organisations are institutions for the behaviour in them. This yields a hierarchy of institutions.

If institutions are ‘enabling constraints’, what precisely does that mean? I propose that it means that institutions affect the causality of action. To proceed, we need a theory of such causality. As in earlier work (Nooteboom 1999a, 2000), I employ a causality of action inspired by the multiple causality of Aristotle. Final causes (goals that people have), efficient cause (agency), material causes (inputs used), formal or exemplary causes (knowledge, skill, methods, models) and conditional causes that constrain or enable the operation of the other causes. The paradigmatic example is the carpenter (efficient cause) who uses wood (material cause) to make furniture, according to professional know - how (formal cause) or some guiding model (exemplary cause), in order to earn a living (final cause), subject to conditions of law and market structure (conditional cause). Final causality is a conjunction of man’s intentionality and external motivating conditions. I proposed that an organization is an institutional arrangement, as a focusing device to align perceptions, interpretations and evaluations. This is to say that it brings together actors (efficient cause), with certain common goals (final cause), and perceptions and knowledge how to do things (formal cause) and/or exemplars to imitate (exemplary cause).

As indicated, we can identify a hierarchy of institutions. For example:

‘.... the constellation of property rights; market entry and exit conditions, tax and social security affect the incentives to become an entrepreneur. When social security is low, people may be pushed into self - employment under adverse labour market conditions (unemployment, discrimination). This is the “refuge hypothesis” of entrepreneurship. Laws of bankruptcy affect exit conditions, and thereby affect the incentive for entry: if the venture
fails, with how little damage can one get out? There are also motivators on a deeper level of categories of perception, interpretation and evaluation, such as sense of responsibility, independence and ‘locus of control’. The latter concept is familiar in the small business literature. It indicates whether people attribute outcomes of actions to their efforts and initiative or to outside conditions. Note that the conditions interact. If there is little internal locus of control, people may not move into entrepreneurship even under adverse labour market conditions and low social security. Other institutions affect other conditions of entrepreneurship. Thus there may be lack of finance (seen here as a material cause) due to lack of venture capital institutions. Level of education and training affect how an entrepreneur and his firm will work, and affect transaction costs.’ (Nooteboom 2000a: 97-98).

In a hierarchy of institutions we can identify a ‘surface’ and a ‘deep’ structure. The first consists of specific institutional structures that enable and constrain activities of firms. The latter consists of underlying, fundamental categories of perception, interpretation and evaluation that are part of culture (in the anthropological sense). We find this notion of a deep structure also in Schein’s (1985) proposal that organizational culture has a deep structure of categories of thought concerning the relation between organization and its environment, reality, truth, the nature of people and their interaction. According to Schein (1985: 14, 88 - 110) the basic categories that constitute organizational culture concern the following:

- The relationship to the environment: does one dominate it or is one subjugated; is the primary focus technological, economic, political or socio - cultural. The domination/subjugation distinction is similar to the notion of ‘locus of control’: does control lie in oneself or outside?
- The nature of physical and social reality and truth: is one pragmatic or does one seek validation in a general philosophy, moral system or tradition; does one avoid uncertainty or does one have tolerance for ambiguity. How does one perceive time and space. For example: does one move towards the future or does that move towards one, or is time past duration (cf. Lakoff & Johnson’s 1980 analysis of ‘metaphors we live by’). Is truth universal or context - dependent; absolute or relative; subjective or objective.
- Nature of human nature: is it good, evil, or both, are people active or passive, perfectible or not, what are sources of motivation?
- Nature of human activity: is it oriented towards being or towards accomplishment, is it self - or other - directed.
- Nature of human relationships: views on power, influence, hierarchy; on intimacy, love, peer relationships; coercive, utilitarian, aimed at goal consensus; degree and source of authority. Here Schein also adds dimensions proposed by Parsons: dimensions of relationships, such as universalistic (equal rights) or particularistic, achievement oriented or based on ascription (family membership, class), oriented to self or collective.

This deep structure generates a surface structure of symbols, rituals, heroes, and role models that constitute a more proximate cause of organizational behaviour. Organizations develop their own specialized semiotic systems: language, symbols, metaphors, myths, and rituals, which institutionalize firm behaviour. These are similar for different firms to the extent that they share an institutional environment. However, within that environment they offer specializations, and those differ between organizations to the extent that they have accumulated different experiences, in different industries, technologies and markets

2.2.4 Organisational and personal trust

What is the relation between trust in people and trust in organizations? Can we have trust only in people or also in organisations? How does it differ? Is trust in organisations based on trust in people, or
vice versa, or both? For an illustration of the difference between trust in organisations and in people, consider the following example.

If we trust people to keep a secret, that is just a matter of their will, while for organizations the question is whether it can manage to get all people involved to keep the secret. Professionals, such as researchers, may want to brag about the outcomes of their research at conferences, while the organization wants to keep it a secret.

Clearly, one can have **behavioural trust** in organisations, in the expectation that in our dealings with them, they will not fail us. It is also clear that we can have **competence trust** in the collective ability of an organization to perform.

In organizations competencies are built up from competencies of people. For these organizational capabilities, March and Simon (1958) used the term ‘performance programs’ and Nelson and Winter (1982) used the term ‘routines’. We need to solve the multi-level problem: how do the levels of people and organizations connect? And beyond an organization, could we trust an entire industry? In earlier work (Nooteboom 2000a) I used the notion of a ‘script’ to deal with the multi-level problem. A script is an architecture of nodes that represent chunks of lower level activity. Scripts have been used to model cognition on the level of people (Shank and Abelson 1977). For example, a restaurant script consists of a typical sequence of typical activities (nodes): entering, seating, food selection, preparation, eating, paying, leaving. Such scripts can also be used to model organizational activities, as the restaurant example already shows. Scripts were used to model organizational processes before, by Gioia and Poole (1984). Individuals substitute activities into nodes of organizational scripts. Scripts can be used recursively on different levels. Organizations can be seen as substituting activities in higher level industry scripts, representing supply chains, for example. Activities substituted by people into organisational scripts can themselves be analysed as scripts. So, one can trust an organization in the sense of trusting its scripts, which entails trust in the abilities of people to contribute their parts. Trust in organisational scripts concerns the efficacy of those scripts, given their architectures, and the ability of the organization to co-ordinate the nodes, i.e. to implement and maintain a script. So, people are included according to the organizational roles they play. However, of course people also participate in many scripts outside their organizational roles. People are more than the organizational roles they play; they are also people ‘qua persona’ (Smith Ring and van de Ven 1994). Having interactions also outside the organisation, their cognitive categories will never become identical to those of other people in the organisation. This is a source of both error and innovation, and often the two cannot be distinguished until the outcomes have had a chance to work themselves out.

Can one also have **intentional trust** in organisations? Earlier, on the level of people, intentional trust included trust in dedication or care (attention not to make mistakes, effort), and trust in benevolence or goodwill (no opportunism). Can we have these in organisations as well? Can we attribute intentions to organizations? Organizations can certainly accept responsibility and liability. They can engage in contracts, renege on them, and be challenged before a court of law.

We can have trust in the competence of organisations to foster and control the competencies and intentions of the people in it, including their dedication and care, and their inclination towards benevolence rather than opportunism. This can be done on the basis of structures and procedures of control, with organisational scripts, as indicated before. As also noted before, organizations have cultural identities. On the surface of organizational culture we find symbols, myths, rituals, styles of behaviour, tall stories, heroes that set examples to be followed. Below that there are more fundamental shared categories of understanding. Organisational procedures (scripts) and culture can both help us to trust the people in the organisation. Intentions are shaped, even if unintentionally, by organizational culture. Intentions may be controlled in standard operating procedures, a code of ethics, and written and unwritten cultural codes of behaviour. These may be aimed to secure reputation and a corporate identity that are conducive to profit or survival. Thus our trust in people can be enhanced or limited by
our trust in the organisations they belong to. For example, our trust in a doctor may be derived from our trust in the hospital he is associated with.

However, there are malevolent organisations, whose procedures and culture are geared towards opportunism. Procedures and culture may be aimed, explicitly or implicitly, to incite people in the organisation to opportunism on behalf on the organisation. For the sake of reputation and corporate image this is likely to remain implicit and covert. Thus it makes sense to talk of the intentions of an organisation, in terms of the constraints and guidance of behaviour, generated by procedures and culture.

Note that it is people acting on behalf of the organisation who yield the results. For this they need to be motivated intrinsically or by monitoring plus rewards and/or punishments. People can be more or less trustworthy than their organizational role requires. Thus, to trust an organization, one must look at its interests and intentions as an organization, but also at the way in which those are embedded in culture and embodied and implemented in organizational roles, motivation and control. Conversely, organizational trust may be based on trust in the people one is involved with. That can be tricky. They may make promises that they cannot keep since those go beyond their organizational competence or influence. Trust in lower echelons may not be backed up by their bosses, and trust in bosses may not be supported by their subordinates or shareholders. Thus one has to look at both the people and their organizational roles.

Organizations can also be the subject of trust, i.e. not the trustee but the trustor. The trust that an organization has in others is built up from opinions, rumours, or facts accumulating inside or outside the company. It can be based on hearsay and gossip or on thorough investigation of performance. When conclusions are reached they may be formalized in decisions to stop dealing with no longer trusted partners, special procedures to be followed, or instructions to take special care. The further discussion of the foundations of organisational trust belongs in Chapter 3.

Summing up: the relation between organizational and personal trust lies in the assignment of people to roles, and the conditions and directions given for playing these roles, in the form of organizational structure, procedures and culture.

We can go beyond organizations to consider industries and institutions more widely. Industries are also said to have ‘industry recipes’ (Spender 1989): certain procedures that the industry takes for granted and that coordinate different activities in supply chains and the like. As indicated before, these also may be modelled by scripts. One can trust the ability of an industry to offer and supply products, protect the environment, deal with problems. One can trust intentions to be open on product quality, pollution, and safety. One may trust them not to engage in anti-competitive practices, such as cartels or the erection of entry barriers.
CHAPTER 3: FOUNDATIONS

In the previous chapter, trust was identified as a disposition towards trusting behaviour, i.e. behaviour that suspends the demand for safeguards. This entails an expectation that things or people will not fail us. Now we turn to the question: what are the foundations of trust; on what is such disposition based? Here, as suggested in Chapter 1, it is useful to distinguish between process and substance. On what mental or organisational process is trust based, and what substantive basis for trust does that yield, in terms of motives and reasons for trust? Here I first discuss the substance and then the process.

3.1 The basis for trust

I distinguish between rational reasons for trust, based on an inference of trustworthiness, and a more or less tacit assumption of trustworthiness caused by affect, routine or lack of awareness.

3.1.1 Reasons

Rational reasons for trust derive from inferred sources of trustworthiness. I focus on behavioural trust, i.e. trust with regard to people, and particularly benevolence trust: what will keep our partners from being opportunistic? That is the most difficult form of trust. An important question here was raised by Williamson (1993): is only calculated self-interest a reliable basis for trust, or can trust go beyond that, to rely on ethics and loyalty beyond coercion and self-interest? Or would that entail a foolish, blind, unconditional trust that will not survive in markets, as Williamson proposed?

Williams (1988) proposed a scheme for the determinants of cooperation, which is reproduced in Table 3.1. This rendering is adopted but also adapted from Williams: I have modified and added elements from other sources in the literature. The sources of cooperation may also be seen as sources of trustworthiness, and hence a rational basis for trust. Williams distinguished between ‘macro’ sources, which apply beyond any specific exchange relation, and relate to the ‘institutional environment’: ethics, god, the law. The ‘micro’ sources arise in specific relations. Williams also distinguished between self-interest and altruism as a basis for trust. Both taken together provide the basis for trust in the wide sense of reliance, as discussed in Chapter 2. The altruistic sources go beyond self-interested behaviour, on the basis of established, socially inculcated norms and values (macro), and in affect based trust (in kinship, friendship, empathy), to which I have added routine based trust (habituation). The latter provide the basis for trust in the narrower, stronger sense of ‘proper trust’, as opposed to the wider notion of reliance. The matrix may suggest a hard, clear separation between the cells, while this is not really there. As I will discuss later, there are spillovers or connections between them. For example, the observance of social norms of behaviour may be part of a pursuit of social recognition, which may be seen as part of self-interest. The intrinsic value of trust-based relations, which is here rendered as part of self-interest, may connect with both the social recognition just indicated, and the disposition to be loyal to friends, which is here rendered as part of affect-based altruistic motives.

The distinction between macro and micro sources is similar to the distinction between ‘universalistic’ and ‘particularistic’ or ‘generalised’ sources made by Deutsch (1973: 55), which goes back to the work of Parsons. Elements from Williams taxonomy have also been proposed by others. Social norms or embeddedness and moral obligations, including a sense of duty, go back to Durkheim’s (1933) notion of ‘organic solidarity’, and were also proposed by Bradach and Eccles (1984), Zucker (1986), Granovetter (1985), and Dore (1983), among others. Fukuyama (1995) employed the term of ‘spontaneous sociability’. Calculation of self-interest includes reputation (Weigelt & Camerer 1988) and the assessment of future benefits of present cooperativeness (‘shadow of the future’), as has been recognised by many (Telser 1980, Axelrod 1984, Frank 1993, Hill 1990, Heide and Meiner 1992, Parkhe 1993). A taxonomy that is similar in some respects to that of Williams was later proposed by Chiles and McMackin (1996). It included social norms, moral obligations, social embeddedness, sense
of duty, rational self-interest and reputation. I have added some of these elements to the original taxonomy proposed by Williams. I will now proceed to discuss each of the four cells in the matrix of Table 3.1.

Table 3.1: sources of cooperation

<table>
<thead>
<tr>
<th>macro</th>
<th>micro</th>
</tr>
</thead>
<tbody>
<tr>
<td>egotistic coercion by sanctions from some material advantage or self-interest:</td>
<td></td>
</tr>
<tr>
<td>authority (the law, god, Leviathan, dictator, patriarch, organisation)</td>
<td>shadow of the future, reputation, hostages, intrinsic value of trust</td>
</tr>
<tr>
<td>contractual obligation</td>
<td></td>
</tr>
<tr>
<td>altruistic ethics: values, social norms of bonds of friendship, kinship, proper conduct, moral obligation, empathy; routines, habituation sense of duty</td>
<td></td>
</tr>
</tbody>
</table>

source: adapted from Williams (1988).

Coercion may not be perfect. God of course sees all, and a Jewish god is unforgiving, but someone’s boss in an organisation may not. Thus trust in people on the basis of organisational safeguards cannot be certain. It depends on the perceived or inferred effectiveness of organisational structure, procedure and culture. In economics, coercion takes the form of contractual obligations. These require appropriate and reliable institutions for law and enforcement. Even under the best of institutional conditions, legal ordering cannot be closed, including all relevant future contingencies, to carry their implications to the present (‘presentiation’) and cover them in the contract (Macneil 1980). This is related to the previous discussion of radical uncertainty, in Chapters 1 and 2. Also, contracts carry costs, and they entail economies of scale (Nooteboom 1993), so that they are relatively costly for smaller firms or smaller transactions. For this reason, Williamson (1985) proposed that for infrequent transactions, one may need to resort to ‘trilateral governance’, where limited, simple contracts are complemented with arbitration by some third party.

Material self-interest includes the fact that one is so dependent on a relationship that one will not hazard opportunistische behaviour, for fear of retaliation. As already indicated, it thus includes expectation of future rewards from cooperative conduct in the present (the ‘shadow of the future’). It also includes the preservation of reputation (Weigelt & Camerer 1988). Note that efficient reputation mechanisms cannot be taken for granted (Hill 1990, Lazaric and Lorenz 1998b). They require that malevolent or incompetent behaviour is reliably detected and communicated, and that the culprit has an interest in future relations with potential partners from the community where his reputation is known. The culprit must not engage in hit-and-run; he must not have the option to exit from the community when his malevolence or incompetence is detected.

Self-interest may also include the need to protect ‘hostages’ (Williamson 1985). The essence of a hostage is that it has value for the hostage giver but not for the holder, so that the latter would not hesitate to sacrifice the hostage in order to punish the hostage giver in case he reneges on his promise to collaborate, and this would hurt the hostage giver. In business, hostages often take the form of information or knowledge that is sensitive, in the sense that it could cause great damage when leaked to competitors. It can also take the form of cross-participation, or the borrowing of staff, with the threat of buying them out. One problem with hostages is that the holder may refuse to return the hostage at the agreed moment in spite of the fact that the hostage giver has satisfied his obligations. One reason for this may be that the holder wants to maintain his hold over the giver. Another is that after a time the hostage acquires value for the holder. Conversely, the hostage may die or lose its value to both sides.

25
A classic and mundane illustration of a hostage is the shoe that one leaves with the cobbler. He will not demand payment in advance, since he holds the shoe as a hostage. One will not demand proof of having surrendered the shoe, since it is of no value to anyone else, so that there is no chance that the cobbler will appropriate or sell it.

Another illustration is the following. A number of years ago, the Dutch passenger car manufacturer DAF could not survive, and was rescued in a joint venture between Volvo, Mitsubishi and the Dutch government, orchestrated by the then minister of economic affairs Andriessen. The concern of the Dutch government was that one of the foreign partners would take out the design capability from DAF, and leave only a hollow or ‘screwdriver’ factory. Asked about this on TV, the minister grinned smugly and told of the shoe-box he had sitting on his desk, containing the design for the next generation Volvo cars, developed in Holland. This suggested that if Volvo reneged on the agreement and took out the design capability, these designs would be in jeopardy. Of course, meanwhile those cars have appeared on the roads, and the hostage has outlived its usefulness. Ford has now taken over Volvo, and again the question is what will happen to the Dutch design capability.

Values and norms of social behaviour, in an ethic or custom of conduct, are part of culture in the anthropological sense (see the discussion of culture in Chapter 2). This applies to countries, but as discussed in Chapter 2 it also applies to organisations, with their specialised cultures in an institutional environment. These values, norms or customs tend to be absorbed into tacit knowledge by socialisation. Yet they belong here, in the discussion of the rational foundations of trust, since one can rationally assess the non-rational, ethical sources of behaviour. Of course, one can never be sure ex ante to what extent a stranger without reputation has actually internalised such norms and values. Here, Williamson was right in saying that under such conditions of behavioral uncertainty one must take the possibility of opportunism into account. However, that does not automatically imply that one should go for contractual safeguards, if they are feasible and effective. That also may have various negative effects, which will be discussed later.

The disposition to observe norms and to follow rules of social behaviour can be motivated by a more or less autonomous, internalised sense of duty, or by the desire for social recognition (Fukuyama 1995), as noted in Chapter 1. As quoted by Fukuyama (1995: 359), Adam Smith in his ‘Theory of moral sentiments’ already recognised that social recognition more generally is an important driver of behaviour: ‘It is the vanity, not the ease or the pleasure, which interests us’ (Adam Smith 1982: 50). This includes the role of wealth as a means to collect not only material goods but also social status or recognition. This desire for social recognition could perhaps be construed as being part of self-interest. Thus, the different boxes in figure 3.1 may not be as distinct as the table suggests: there are spillovers or connections between them.

Economists, in particular, but also rational choice sociologists, are prone to push everything into self-interest, to preserve their theoretical perspective of self-interested, individual, rational choice. However, the point is that even then a conflict remains between different dimensions of self-interest. People may be torn between self-interest in terms of what Adam Smith called ‘ease or pleasure’ and what he called the ‘vanity’ of social recognition. Here, such recognition derives from a certain amount of altruism, with a greater or lesser sacrifice in terms of money. Even if altruism is pursued for the sake of social recognition, it is still altruism, in the sense that material sacrifices have to be made for it. Therefore, I prefer to keep it in the category of altruism, in contrast with what I call ‘material advantage or interest’.

Intrinsic value of trusting behaviour, liked for its own sake, and preferred to the aggravation of suspicion, is included in self-interest, but it may be linked to the desire for social recognition of loyalty that I included in the box for values and norms of social behaviour.

The micro bonds of friendship, kinship or empathy, in table 3.1, yield altruistic sources of trustworthiness that are attached to specific relations ex ante (kinship, friendship), or are built up or broken down in the course of relations (friendship, empathy). Thus, they require a discussion of
underlying processes, to which I will turn later, in this chapter. As pointed out by Williamson (1993), here there is a risk of blindness: such personal bonds may prevent us from recognising and facing hazards of opportunism. Therefore, Williamson argued, they are unwise and will not survive in markets. They should be reserved for non-economic relations of family and friendship. And indeed, problems of trust and betrayal in family firms and partnerships between friends are well documented. Several massively popular television series have been based on that theme. Thus, the important question arises how trustworthy such trust is. Can it be genuine without going so far as to be blind and unconditional? Can we establish the limits of such trust without thereby falling back into calculative self-interest? I will turn to these questions in the discussion of the failures and limits of trust, in Chapter 5.

In the box on micro/altruistic sources I have included routinised behaviour, which may also be called ‘habituation’. This is part of my answer to Williamson’s claim that behaviour that is not based on calculative self-interest behaviour is blind and cannot survive in markets. Routinised behaviour by definition is non-reflective, and thus cannot be called calculative. Nevertheless, as Herbert Simon has taught us long ago, routinised behaviour is rational in view of bounded rationality, since it allows us to focus our limited capacity for attention and rational evaluation on matters that are new and have priority. Routines are rational also in the sense that they are based on proven success in past behaviour. Their lack of awareness creates the problem that they may no longer be adequate when conditions change. However, when results or perceived events exceed certain tolerance levels, routines may be summoned back from subsidiary into focal awareness (Polanyi 1962), to be subjected to rational scrutiny. On the other hand, they may have subsided so deeply into tacit knowledge as to become inextricable, like instinct. However, this is likely to occur only after such a long time, in a succession of generations, that the change that makes routines fail must be very drastic and unusual.

Williams argued that none of these sources by itself suffices and that, in cooperation, some mix will always be operative, while no universally best mix, regardless of specific conditions, can be specified. Often, trust based on friend- or kinship will not suffice as a basis for cooperation. It may not be sufficiently robust to extremes of temptation. Conversely, material self-interest and coercion are seldom sufficient as a basis for cooperation: one needs trust on the basis of non-egotistic sources to the extent that one cannot fully control the partner’s conduct by threat and reward (cf. Deutsch, 1962, quoted in Zand, 1972); to make this fragile basis for cooperation more resilient (Ring and van de Ven 1992).

3.1.2 Organisations

Chapter 1 indicated that organisations also can be the object of trust, in both their competence and their intentions. We can trust an organization to behave responsibly, regarding its stakeholders and the environment. Of course an organisation itself does not have an intention, but it has interests and can try to regulate the intentions of its workers to serve those interests. One’s trust in an individual may be based on one’s trust in the organisation he belongs to. Trust in an organisation can be based on trust in the people in it. It can be affected by corporate communication, which aims to project a certain image. But ultimately the proof lies in the performance of its people. Particularly important for the perceptions that underlie trust in an organisation are the public conduct of the firm’s leadership and roles that connect the firm with customers or outside partners. These are the ‘boundary spanners and gatekeepers’ such as purchasers, marketers, negotiators, staff exchanged or combined in joint teams with partner firms.

In both the foundation of organisational trust on the trust in people and the foundation of personal trust on the trust in an organisation, we need to carefully take into account the position and role of those people in the organisation. Are their competencies and intentions supported and backed up by the organisation? Are the interests and the culture of the organisation properly endorsed and implemented by the people?

Earlier, in Chapter 1, I proposed the notion of firm as a focusing device, to align perceptions and interpretations and evaluations sufficiently to achieve a common goal. How narrow or sharp and
how tight, in the sense of controlled, should that focus be? The wider and looser the focus, the more tenuous the relation between organisational and personal trust. In Chapter 1 I indicated a trade-off between a sharp and tight focus, for the sake of efficient exploitation, and a wider and looser focus for the sake of exploration. One way to solve the problem of combining exploitation and exploration is to have a relatively narrow focus in one’s own organisation, complemented with outside partners at a cognitive distance that is sufficiently large to yield novelty of insight and sufficiently small to yield understanding. Here we see that another reason for a sufficiently sharp and tight scope is to establish a coupling between personal and organisational trust.

If the basis for one’s trust in an employee is friendship or empathy that has built up in the course of the relationship with him, is this in line with the interest of his firm, and with the norms of conduct that form part of the firm’s culture? If not, to what extent can he back up his promises nevertheless, and is there a risk that he will be forced to renege on his promises when found out? Conversely, if one’s trust in the organisation is based on the fact that its interests demands loyalty to an agreement, can one rely on the person implementing them? Suppose there is personal animosity between you and him. Could he sabotage the agreement and get away with it?

Lorange and Roos (1992) recounted the case of an alliance between two producers in garden equipment in Sweden and Norway (‘Norpartner and Swedpartner’). This was based entirely on a personal relationship between the CEO’s of the two companies. A third party, which competed with both, in Sweden and Norway, took over the owner of Swedpartner and ended its alliance with Norpartner, who meanwhile had become so dependent on the alliance that he went bankrupt. In this way the third party eliminated two competitors in one stroke. This case illustrates the point that organisational interests need not be identical to individual interests, and trust in a personal relationship may not be a reliable basis for organisational trust. In this case, one should also have secured a basis for relying on owners, by building in their interests or somehow blocking the possibility of a damaging takeover.

Trust in organizations may be based on its survival conditions. For example, to survive it must guard its reputation for quality, its brand names. Such safeguards depend on industry conditions, such as the effectiveness and importance of reputation mechanisms, and then intensity of price competition. Clearly, a firm that is under great pressure to make all the profit it can get and to cut costs where it can, as a matter of sheer survival, can afford less goodwill or benevolence, at the sacrifice of profit, than a firm that is not under such pressure. This yields differences especially between industries.

In the early nineties, I gave a talk on the governance of inter-firm relations, with an emphasis on buyer-supplier relations, to an audience from the car industry, in the Netherlands. There, I presented the results of my efforts to combine considerations of trust with the recognition of hazards of opportunism from transaction costs economics, in a wider theory of governance. The audience scolded me for being so naive as to take trust seriously. Trust, in their view, is naive and not fit for survival in markets (rather like the position taken by Williamson 1993). A few days later I gave a similar talk to the company that explores and produces natural gas in the Netherlands. They took an opposite view, and scolded me for my cynical talk of opportunism. Then I realised what lay behind this marked difference. At the time, car sales had slumped, which greatly intensified price competition, and the crisis caused manufacturers to renege on the promises of durable relations to suppliers that had previously been given. The natural gas firm, by contrast, had a tight, government-backed monopoly in their field. They could afford to make life easy for their suppliers and themselves.

Summing up, in dealing with organisations one has to consider the basis for reliance both on the people one is dealing with and on the organisation. What mix of foundations of trust is there on the different levels, of people and organisation, and how consistent are they, i.e. to what extent do they
support or compensate for each other? Structure of ownership and control, organisational culture and procedures for guiding, supporting and controlling people in their organisational roles have a crucial mediating role here, and become part of the basis for organizational trust.

3.1.3 Varieties of capitalism

Contractual sources of trustworthiness (or more accurately: reliability) require appropriate and effective institutions: appropriate laws of property and trade, reliable and competent lawyers, non-corrupt police and judiciary, etc. These vary considerably between countries. Clearly, values and norms of social conduct also vary greatly between countries, and may change in time, as has been emphasised by Fukuyama (1995), among many others. Fukuyama used the term ‘spontaneous sociability’, but this is not really spontaneous, in the sense that it requires, i.e. is enabled or constrained, muted or amplified, blocked or transmitted by a variety of ‘intermediate communities distinct from the family or those deliberately established by governments’ (Fukuyama 1995: 27).

Similarly, the importance and the functioning of kinship, friendship and routinisation also vary greatly across cultures. These differences between countries have also been demonstrated by other researchers (e.g. Zucker 1986, Hill 1990, Sako 1992, 1994, 1998, Hollingsworth and Boyer 1997, Whitley 1999).

Zucker (1986) argued that in the US increasingly trust based on family and (local, ethnic or religious) community and produced in ongoing relations has eroded. Communitarianism was replaced by individualism, and the vacuum had to be filled by means of trust (or more accurately: reliance) based on legal and associated institutions. A similar argument was echoed by Fukuyama (1995), who stated that in the US there has been a development from a prevailing sense of trust and community, based on norms of social behaviour and duty, towards a more individualistic and legalistic society. In other words: a decline of spontaneous sociability, although pockets of it still exist, in some communities. In terms of table 3.1, this entails an upward shift from altruistic to egotistic sources of trustworthiness. He proposed that ‘.. the French capacity for spontaneous sociability was effectively destroyed in the sixteenth and seventeenth centuries by a victorious centralising monarchy’ (1995: 28). His analysis is not always correct, however. He characterised Chinese societies (in Taiwan, Hong Kong and mainland China) as ‘familialistic’, because ‘the essence of Chinese Confucianism is the elevation of family bonds above all other social loyalties’ (1995: 29). However, here he overestimates the spread and depth of Confucianism across the large diversity of cultures and conditions in China, and neglects the importance of the more widespread social ties produced in army networks.

These differences accord with the view that there are ‘varieties of capitalism’, in different ‘business systems’, due to institutional differences between countries (Hollingsworth & Boyer 1997, Whitley 1999, Nooteboom 1999cd). Hill (1990) argued that an institutional basis for trust yields lower transaction costs, and thereby produces a competitive advantage among nations, so that in time, in an evolutionary selection process, trust-based societies will prevail. This argument is the opposite of Williamson’s, who argued that in markets trust will not survive. Against Hill’s argument one can raise the objection that it seems to neglect the possible effects of the invasion of a trust-based society by opportunists. Will trust be robust under such invasion, or will the basis for trust unravel, as opportunists exploit trusters? To preclude that, one may have to erect barriers to entry, but this may provoke penalties imposed by the global economic community. And if trust can be maintained only at the cost of being barred from world markets, the trade-off is not so clear (Nooteboom 1999c).

A private communication from Barbara Krug in 2000.
In the previous section, several complications were identified in the operation of different sources of trustworthiness. Complete contingent contracts are generally impossible, depending on the uncertainty of relevant future contingencies. For small or infrequent transactions, contracts may be too costly, and one may need to revert to the use of a third party to yield arbitration, in ‘trilateral governance’. A reputation mechanism often requires a filter to sift out slander, and an amplifier and broadcaster of justified complaints. This may be a task of a third party. Hostages may not be returned to givers, in spite of their good behaviour, and it is a classic solution to have hostages kept by independent third parties to offer guarantees to both sides: a guarantee of sacrifice if the giver reneges, but also a guarantee to the giver that the hostage will be returned (de Laat 1996). These are just three roles for third parties out of a larger number of such roles (Nooteboom 1999b). I will return to this later. The point here is that such third parties form an important element in the institutional structure that supports, or fails to support, trust. The roles identified here substantiate the roles of the ‘intermediate communities’ noted by Fukuyama. The roles may be played by trade associations, professional bodies, banks, ‘interlocking directorates’ of firms, etc. Here also, there is a wide variation between countries.

3.1.4 Affect and routine based trust

Next to rational reasons for trust, based on an inference of trustworthiness from its possible sources, as discussed above, feelings of trust, and trusting behaviour, can be based on affect: empathy, friendship, love. Perhaps here we should not talk of reasons but of causes of trust. This distinction may suggest a greater cleavage between rationality and emotion than is valid. Like many others, I believe that rationality and emotions are intertwined (Polanyi 1962, Merleau-Ponty 1964, Damasio 1995, Hendriks-Jansen 1996, Lakoff and Johnson 1999). As I indicated earlier, in the discussion of theory of knowledge, in Chapter 1, I include in cognition not only perception and interpretation but also evaluation, i.e. value judgements. Not only value judgements but also interpretations and even perceptions are emotion-laden.

Lakoff and Johnson (1999) offer an evolutionary argument to take into account how cognition is ‘embodied in the flesh’. Contrary to the opinion of Descartes, cognition is not separate from the body but is embodied. In evolution, it has been tacked on to bodily processes thrown up by previous evolution. It is rooted in proprioceptive perceptions and mechanisms, reflexes, emotions (see also Damasio 1995). This bodily embedding of thought was argued before by Merleau-Ponty (1964). It has implications for how cognition works. Thus, Lakoff and Johnson build on their previous work (Lakoff and Johnson 1980), in which they proposed that to conceptualize abstract notions, such as happiness, goodness, etc., we can only fall back on more primary bodily perceptions, so that even the most basic categories we use are in the nature of metaphors rather than literal descriptions. For example, goodness, happiness, health, achievement, control are ‘up’: performance going up, people going up in society, having high standing, feeling up to something, the upper classes, etc. This is due to the primary experience of standing up when healthy and lying down when ill or dead.

Earlier, I also followed Herbert Simon in identifying the role of emotions in reason, to shift routinised behaviour from subsidiary to focal awareness. My position is also consistent with the notion of cognitive dissonance. As Deutsch (1973: 159) put it:

‘A person’s perceptions of another will be determined not only by the information he receives from his direct experiences or from what others tell him, but also by his need to absorb this information in such a way as to prevent disruption of existing perceptions, cognitions, or evaluations to which he is strongly committed’

Nevertheless, we can distinguish more or less rational inference of trustworthiness from less reflective causes of trust, based on affect of friendship or kinship or on routinised behaviour. Here the
basis is not so much analysis and argument as assumptions based more on faith or routine. As indicated before, and to be discussed further in Chapter 5, this yields dangers of blind trust.

Here also, the need arises to go more deeply into the processes underlying ‘trust production’, including processes of routinisation, habituation, perception, absorption, cognitive dissonance, negation, blindness, etc.

3.1.5 The pervasiveness and inevitability of trust

This chapter discusses the basis of trust, but it must be noted that we must trust even without a basis in rational evaluation, affect or proven routines. This is due to radical uncertainty of behaviour and the incompleteness of language, or ‘inscrutability of reference’, as Quine (1960) called it.

Firstly, there is the argument of radical uncertainty discussed earlier, in Chapters 1 and 2. Probabilistic calculation of possible actions ignores the possibility of unforeseeable contingencies. Furthermore, from an ethical perspective it robs people of their essential freedom, which includes the option not to satisfy our expectations.

Secondly, language and communication cannot work without trust that goes beyond reason and affect. If we demand safeguards, these have to be phrased in some code or language, and if we insist on defining all the terms of the safeguard unambiguously, we would fall into an infinite regress of defining the terms of definitions. Knowledge and meaning are embedded in tacit knowledge that cannot be completely codified without any residual. At some point we have to stop and take terms for granted, and that is where trust necessarily starts. As indicated before, thought and language are based on metaphors that are difficult to subject to rational analysis and justification. The question is not whether we trust but where we lay its limits. This is indeed trust in the strong, narrow sense defined before, going beyond contractual safeguards and safeguards based on self-interest. It goes even beyond affect from kinship, friendship or love, and beyond routines that may be recalled into focal awareness. It involves tacit, unreflective assumptions that are taken for granted. This point is an extension, perhaps, of the argument of routinisation discussed before. Here, however, we are dealing not with routines that we developed on the basis of our personal experience, but routines that are part of our social inheritance, or perhaps even our genetic make-up, that go so deep that they cannot be called into focal awareness, for the purpose of revision, and in that sense are more like instincts. Perhaps, here we should not speak of trust but of confidence, as defined before. Such fundamental, ineradicable, virtually instinctive categories are inevitable, and cannot be chosen or rejected.

The argument comes close, I think, to Macneil’s argument (Macneil 1980 and much earlier work), that contract is inevitably embedded in society:

‘.. contract between totally isolated, utility maximising individuals is not contract, but war; contract without language is impossible; and contract without social structure and stability is -quite literally- rationally unthinkable, just as man outside society is rationally unthinkable (1980: 1)’.

We have to take for granted, at some point, that people are competent to understand what is said, and will adhere to basic social and linguistic conventions that, as we delve more deeply, become more tacit and difficult to codify. At some level, even the dishonest have to mean what they say.

When Saddam Hussein broke a fundamental ethical convention by using his own citizens as a shield against attack, he still had communicate this to the enemy, and adhere to what is meant by the term ‘citizen’, for his threat to have the intended effect.

Summing up, trust is pervasive and inevitable because future contingencies and motives are never completely known, and language cannot yield certainty of meaning, so that contracts and self-interest always leave a gap of uncertainty. This does not depend on states of the world, or people, firms, industries or nations: to a greater or lesser extent it is always there. At that point, we must
surrender to trust or die from inaction. At that point trust is blind, in the sense that it is based on tacit assumption, not rational evaluation, let alone calculation, or even affect of love or friendship, or revisable routines. However, this blindness does not mean that trust is unconditional. There is a limit, where trust ends and conditions begin. This functioning of trust, in relation to conditions, will be discussed in Chapter 4.

3.2 The process

In the previous paragraph, a number of sources of trustworthiness were identified. The question next is how we can infer the existence and reliability of those sources. A question also is how other, more affect-based sources of trust arise and develop. At several points, I identified the need to go into the processes of trust.

3.2.1 Modes of trust production

How does one decide the extent to which different sources of trustworthiness are in operation, or are feasible? What is the basis for inferring the disposition of people to obey social norms, and their personal loyalty? We cannot directly observe intentions or capabilities, but only certain personal characteristics and people’s actions, and we can listen to what they say. To use the terminology of Zucker (1986): what are the modes of ‘trust production’? To some extent this was discussed in the previous paragraph, where I indicated the role of institutions and third parties, but further analysis is needed.

Zucker made a distinction between different modes of trust production that is complementary to the different sources in Table 3.1, and adds to my previous analysis. They are given in Tables 3.2 and 3.3. For the basis of trust a distinction was made between personal characteristics, institutions and the process in which relations develop. In Table 3.2 again I focus on intentional trust, in particular benevolence trust (trust in the absence or limitation of opportunism) and adapt Zucker’s scheme accordingly. In Table 3.3 I indicate the result when we focus on competence trust. In that table there is no relation to Table 3.1, since that focuses on intentional, not competence trust.

### Table 3.2: modes of intentional trust production

<table>
<thead>
<tr>
<th>Basis</th>
<th>Examples</th>
<th>Connection with table 3.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristics</td>
<td>membership of family, community, culture, religion</td>
<td>mostly the altruistic sources of social norms and kinship</td>
</tr>
<tr>
<td>Institutions</td>
<td>rules, ethics, professional standards</td>
<td>the macro altruistic source of social norms</td>
</tr>
<tr>
<td>Process</td>
<td>loyalty, commitment, habituation</td>
<td>the micro altruistic source of friendship, habituation</td>
</tr>
</tbody>
</table>

source: adapted from Zucker (1986)

### Table 3.3: modes of competence trust production

<table>
<thead>
<tr>
<th>Basis</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristics</td>
<td>membership of professional associations, educational achievements</td>
</tr>
</tbody>
</table>
Institutions technical/professional standards, benchmarking
based trust

Process mutual adapatation, learning by doing, routinization
based trust

source: adapted from Zucker (1986)

As indicated before, Zucker (1986) argued that in the US increasingly characteristics based trust related to family and (local, ethnic or religious) community and process based trust in ongoing relations have eroded. Communitarianism was replaced by individualism, and the vacuum had to be filled by means of institutional based trust. Note that in my previous analysis, ‘institution based trust’ also includes the institutional basis for contracts as a source of reliance.

In the attempt to create conditions for process trust to develop, detailed contracts can be destructive: they demonstrate mistrust which engenders mistrust, so that mistrust becomes a self-fulfilling prophecy (Macauley 1963, Zand 1972).

In what sense can trust be ‘produced”? Institutions in the form of values, norms and standards of conduct can be part of the macro sources of cooperation: the institutional environment. But they may also be built up in specific relations, on the micro level, as part of institutional arrangements. And the two interact: institutional arrangements build upon the institutional environment and in due course may shift it (Sydow 1996). Some institutions, such as technical standards or systems of certification, can be developed on the basis of some rational design, although this may take quite some time, since they tend to affect established interests and may require a political process. One can select a partner on the basis of his characteristics, such as being a member of a family or community. But one cannot simply buy into characteristics based trust: one can marry into a family and one can become a member of some communities, but entry selection can be strict and it can take considerable time. Process trust by definition has to grow. Such trust can be facilitated, by creating favorable conditions, but it cannot be purchased and installed by the pound. To ‘produce’ process trust is as paradoxical as to order spontaneity: if it worked it would not be real. Process trust is as much the outcome of a relation as the basis for it. Sydow (1996) approached this from the perspective of Giddens’ structuration theory: to the extent that process trust is already available it provides the basis for a relationship, it is reproduced in the relation, if it goes well, and may be further deepened to provide the basis for further extending the relationship. Evidently, we need to investigate processes of building and breaking trust in more depth.

As discussed, to some extent sources of trustworthiness can be inferred from the availability and effectiveness of institutions (such as laws and judiciary, professional standards), and personal characteristics of various kinds (such as membership of family or of professional, cultural or religious associations, or educational achievements). However, much is derived as ‘process based trust’ in interactions in specific relations. The question now is how that works.

Dedication shows itself best when there are opportunities for slack. Benevolence show itself best under opportunities for opportunism and temptations or pressures to utilise them. It is especially important to assess whether benevolence is intrinsic or extrinsic. In the first case loyalty is an internal goal, while in the second case it is a means to pursue self-interest. The first does not require monitoring, and the second does. The first is based on ethics and conscience, based on norms and values, or on emotions of friendship or kinship, or on the enjoyment of trusting relations. The latter is based on quid pro quo, lack of opportunity for opportunism, dependence or fear of reputational loss (Deutsch 1973). Connected with this, Deutsch also recognized the notion of ‘focus’, with three possibilities: focus on results for the other, on warrantable effort (is one seen to be doing one's reasonable best?), or on doing as one is told. Is one genuinely trying to cooperate or is one intent only on legitimizing one's actions?
So, when a supportive action by X is observed, how does one judge whether this is based on trustworthiness? Table 3.4 gives a sequence of questions that can be an aid to such assessment (adapted from Deutsch 1973):

Table 3.4: attribution of trustworthiness

<table>
<thead>
<tr>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. was the outcome intended by X, or was it an unintended result of his action?</td>
</tr>
<tr>
<td>2. did the action entail significant risk to X?</td>
</tr>
<tr>
<td>4. was X aware of the risk, and was it not neglected out of impulsiveness?</td>
</tr>
<tr>
<td>5. did X attach a positive value to this risk, out of masochism, sensation, (self) image?</td>
</tr>
<tr>
<td>6. did X have a choice, or was the action dictated by compulsion or conformity?</td>
</tr>
<tr>
<td>7. was it out of confidence in the system rather than a positive evaluation of the situation?</td>
</tr>
<tr>
<td>8. was it out of enlightened self-interest?</td>
</tr>
<tr>
<td>9. was it out of enjoyment of trust relations?</td>
</tr>
<tr>
<td>10. was it out of ethics, friend- or kinship, habituation?</td>
</tr>
</tbody>
</table>

Deutsch noted that one’s power can have an adverse effect on one’s trust. This can be seen from the above: if one is very powerful, there is more suspicion that people subjected to one's power are trustworthy only because they have no choice. In case of absolute power, the hypothesis that this is the case can never be rejected. Thus power can breed suspicion, and absolute power can yield rampant paranoia. Here, one recognises stories about Stalin. The problem here is that mistrust tends to feed upon itself even more than trust does. (Mis)trust by X tends to engender (mis)trust on the part of Y, which justifies and deepens X's (mis)trust. While trust can be falsified because it leads to reliance on others which can be disappointed, mistrust cannot, because it blocks trusting action that might disprove it.

3.2.2 The dynamics of trust

In Chapter 2, trust, or more widely: reliance, included an expectation that things or people will not fail us. However, the disappointment of such expectation does not necessarily yield a breach of trust. Remember that trust can have as many objects as there are causes of things going wrong. As discussed in Chapter 2, in human relations the object of trust can be a partner’s competence, his dedication, his benevolence (absence or limitation of opportunism), or external conditions. One may observe that external conditions rather than a partner’s behaviour was the cause of disappointing outcomes. It is possible that this cannot be seen, but it is claimed by the partner as an explanation. Will the trustor accept the explanation, or will he see it as an excuse for something else: for lack of dedication, or competence, or worse, a sign of opportunism? Will he, perhaps, turn a blind eye to the mishap, negate evidence of incompetence or opportunism out of cognitive dissonance or blind trust? Or will he give the partner the benefit of the doubt, but remain alert to future mishaps and their possible causes? Or will he jump to a conclusion of opportunism? Which is the case will depend on a number of things: the sources of previous trust or suspicion, the history of the relationship, the psychological make-up and the personal experience of the trustor, external social and economic conditions (e.g. intensity of price competition).

Note that in the light of the theory of knowledge discussed in Chapter 1, with the virtuous role of ‘cognitive distance’ for the purpose of learning, misunderstanding and conflict of opinion are often creative, in shifting existing views and jointly generating novel insights, in a learning process. This reinforces the idea that the dissatisfaction of expectations does not necessarily break down trust. On the contrary, the joint solution of conflict can enhance and deepen trust, in several ways. One way is
that it yields learning, as just indicated, which confirms the value of the relation and thereby increases mutual commitment. Another way is that the fact that problems are solved in itself reduces perceived risk in the relation. The conflict yielded a test of the strength of mutual benevolence and the dedication to ‘work things out’, in a mutual ‘give and take’. The fact that the relation survived the test strengthens trust in the strength and resilience of benevolence and dedication. This is how the process of ongoing and successful relations, with solutions of conflicts, can deepen trust.

The positive effect of the solution of conflicts carries force especially because the reverse is so often observed, that under adverse conditions a relation breaks down in mutual recrimination and suspicion. This can easily arise especially when the stakes are high at the beginning of a relation between strangers. As indicated, when things do not go all right, this may be due to accident, lack of dedication, lack of competence or opportunism. If in fact the cause is opportunism, this will not be admitted. Knowing this, one may suspect opportunism even when it is denied, or for the suspicious especially when it is strongly denied. In this delicate stage of a beginning relation with high stakes of dependence, a third party may play a useful role in eliminating such incipient misunderstandings and attributions of fault before they become so large, and evoke such hostile reactions from the unjustly accused partner, that they escalate beyond repair.

Pettit (1995) introduced the notion of ‘trust-responsiveness’. It is related to the desire for social recognition discussed earlier. When offered trust, people may reciprocate ‘due to the love of regard or standing in one’s own eyes and in the eyes of others’ (Pettit 1995: 203). To the extent that this is true, trust is reciprocated and there is a possibility of an upward spiral of trust. However, the converse also seems to apply, where mistrust engenders mistrust, and then there may be a downward spiral of suspicion.

In view of this potential self-reinforcing dynamic of trust and suspicion, honesty and trust in honesty are crucial. Honesty here is openness: giving appropriate and truthful information. Dishonesty is the withholding or distortion of appropriate information. Honesty and trust reinforce each other, as suspicion and dishonesty do. When there is no intentional trust, one is afraid to be honest, lest the partner will mis-use information for opportunistic purposes or to lower the level of dedication. If there is lack of trust in dedication, one may think praise will cause slack. If there is lack of trust in benevolence, one may think that information on one’s needs, opportunities and their limits, or competencies will be used opportunistically in power play, such as bargaining, further reduction of one’s options, or treating sensitive information as a hostage (blackmail). There is also a more subtle reason for dishonesty that is benevolent. One may withhold criticism out of fear of (further) reducing a partner’s self-confidence. When the other side perceives that he is receiving neither trust nor information, he is likely to reciprocate with dishonesty. One elementary lesson is the following. When a disaster is foreseen, one is tempted to keep it secret. This should be resisted. Here is a chance to win trust by announcing the problem before it becomes manifest, asking for help and engaging in a joint effort to redress or mitigate the disaster.

A special problem here lies in the situation where collaboration has to develop between partners who are unequal in their dependence on each other (Klein Woolthuis 1999). The most dependent partner may be suspicious because of the one-sided risk he runs, whereby he starts the relation on the basis of mistrust or apprehension, and is on the look-out for signs of opportunistic exploitation of his dependence. This is a special case of a more general phenomenon that lack of self-confidence engenders mistrust, which breeds mistrust. Earlier, I identified a benevolent reason for dishonesty, not to reduce a partner’s self-confidence or expose his weakness. Here, it becomes clear that this can be in one’s self-interest, not to make the partner more apprehensive, defensive and suspicious.

Deutsch suggested that there is circular causation between characteristics of participants and the results of interaction. He offered his ‘crude law of social relations’:

‘The characteristic processes and effects elicited by a given type of social relationship (cooperative or competitive) tend also to elicit that type of social relationship’.
If this is true, then one must be very careful how to start a relationship, because it may be difficult to get out of the initial mode of interaction. This yields a lesson for the development of collaboration, including the preparation of an alliance or merger. Often, in the initial situation of bargaining, games are played in the manoeuvring for position that set the relation going in a mode of rivalry, which may then be difficult to turn around to cooperation, in the stage of implementation. The bargainers from the top of the firms throw the problem in the lap of the poor implementation manager, who is confronted with a huge obstacle that can jeopardise the relation before it has properly started. Therefore, those responsible for implementation should be included in the bargaining process, or bargainers should also be made responsible for implementation. Perhaps, there is another role for a third party here: to guide negotiations prior to collaboration.

There is another pitfall. If two strangers need to rely on each other, with large dependence and much at stake, and there is no basis for trust from earlier experience, reputation, advice from a third party, kinship, or apparent values and norms of behaviour, then one may be tempted to make a safeguard in the form of an extensive contract to limit ‘opportunities for opportunism’. There are several problems with this, which will be discussed more systematically in Chapter 4. In the present context, the problem is that the relation starts in an atmosphere of mistrust and rivalry, which may be difficult to turn around into one that allows for the building and deepening of trust.

The lesson from this is that a good way to start a relation between such strangers is to take small steps that are likely to yield positive results soon, with limited risk, so an extensive contract against opportunism is not needed, and a basis for trust may soon be built. Process based trust moves through several stages. Generally, for people who are not acquainted ex ante, at the beginning of the process calculativeness will tend to prevail: the focus is on the generation of material value. The greater this value or its urgency, the more risk one may be willing to take. Partners are given the benefit of the doubt, within certain margins of risk left by incomplete contracts, accepted on the basis of some heuristic that takes into account prevailing ethics, customs, experience, urgency and the intrinsic value attached to trust. As the relation begins to generate value, trust in both competence and intention may deepen due to perception and understanding of the sources of both. Next, bonds of friendship and identification with each others’ interests may develop to deepen trust further, and routinization may weaken the awareness of risk.

To say, as Deutsch did, that the type of relationship tends to reinforce itself of course does not mean that this is inevitable. A mistrustful, rivalrous relationship may develop into a trustful, cooperative one, and vice versa. The latter is more probable than the former, as is expressed in the saying that ‘trust comes on foot and departs on horseback’. The preceding analysis clarifies this. Consider the attribution of trustworthiness, in table 3.4. This requires that all conditions specified in the table be met. Favourable actions have to be intended, and go beyond self-interest, and entail a risk or sacrifice, and the partner must have seen that, and he must have had other options of action, before one can attribute trustworthiness, going beyond coercion and self-interest. For complete or incipient mistrust, or doubts about trustworthiness, or at least lack of an increase of trust, only one of these tests has to be failed.

Zand (1972) proposed a cycle in which trust engenders openness, yielding information, which provides a basis for the exertion and acceptance of mutual influence, which yields the willingness to demand less and accept more control from the partner, which further engenders trust. In other words, the provision of information based on trust promotes the responsiveness to trust identified by Pettit, which may already be latent but requires a trigger of information that has the dual function of demonstrating trust and reducing the risk of trust reciprocation. This can set a positive dynamic of trust going. Thus, Zand (1972: 238) included openness in his conceptualisation of trust as:

‘... behaviour that conveys appropriate information, permits mutuality of influence, encourages self-control, and avoids abuse of the vulnerability of others’.
Here, we see a positive relation between trust and information: A trusts B and therefore gives information (even if B could use that to the detriment of A), which makes B trust A and give information in return. Mari Sako (1998) suggested that there is a paradox in the relation between trust and information. In her empirical work she indeed observes a positive relation: trust goes together with information exchange. On the other hand, Sako says, trust is widely seen as a substitute for control: if there is trust there is less need for control. Such control requires information, for monitoring. Hence, theoretically, if there is more trust there is less control and hence less information exchange.

There is a misunderstanding here. There is no paradox. To the extent that trust is indeed a substitute for control, and the extent to which this is the case will be discussed in Chapter 4, it means that if A trusts B he will demand less formal control over him. However, as Zand indicated, he may at the same time accept more control from B. Furthermore, his trust in B, and limited demand of formal control over B, is likely to be the result of the fact that B is open and gives appropriate information. Thus all the following propositions can be true at the same time: trust goes together with much information exchange, formal control requires information, and trust can be a substitute for formal control.

Part of the issue lies in the ambiguity of the notion of control. When trust is a substitute for control, this refers to formal, contractual control with an imposition of procedures for monitoring. The information provided voluntarily in trust also yields a form of control, but this is in important ways quite different from formal control. First, in trust the information is voluntary. It is not based on demand but on reciprocity. One does not need to demand appropriate information on potential or imminent problems and their true causes because that is already offered. Second, it tends to be less codified and explicit, and more tacit and implicit. Third, not being specified as to its conditions for delivery, it is more flexible, geared to unanticipated contingencies, and therefore it is more robust under uncertainty. That is a great advantage, because in formal control such contingencies would generally be impossible to specify completely. Fourth, and most importantly, by assumption the information given in trust is truthful, while in formal control truth would have to be tested. In view of the multiple possible causes of disappointing outcomes, and the possibility to claim accidents as an excuse for failure or a mask for opportunism, this is a great benefit.

The analysis is reminiscent of what Hirschman (1970) called ‘voice’, as opposed to ‘exit’. In exit, one walks out when dissatisfied, avoiding argument. I.e. one quits from one’s job, fires people, sells shares or part of a firm. In voice, the first response is not to walk out, but to seek amends. One reports one’s dissatisfaction, asks for an explanation, and asks for and offers help to ‘work things out’. The importance of this in inter-firm relations has been indicated and demonstrated in empirical work, in particular by Susan Helper (1990). I applied the notion in an analysis of differences in systems of corporate control (Nooteboom 1999e). I propose that the process of voice is as described and analysed above.

Summing up, a positive dynamic of deepening trust hinges on honesty, in a commitment to voice.
References


Durkheim, Emile, 1933, The division of labour in society, New York: Macmillan.


Frank, R.H., 1988, Emotions within reason, the strategic role of the emotions, New York: W.W. Norton.

Fukuyama, 1995, Trust, the social virtues and the creation of prosperity, New York: Free Press.


Klein Woolthuis, R., 1999, Sleeping with the ennemy: trust, dependence and contracts in inter-organisational relationships, Doctoral dissertation, Twente University, P.O. Box 217, 7500 AE Enschede, the Netherlands.

Knight, F., 1921, Risk, uncertainty and profit, Boston: Houghton Mifflin.

Klos, T.B., 2000, Agent-based computational transaction cost economics, doctoral dissertation, Groningen University, the Netherlands.


----- 1999a, Inter-firm alliances, analysis and design, London: Routledge.


----- 2000a, Learning and innovation in organizations and economies, Oxford: Oxford University Press.


Sako, M., 1992, Prices, quality, and trust.; Inter-firm relations in Britain and Japan, Cambridge: Cambridge University Press.

----- 1998, The information requirements of trust in supplier relations: Evidence from Japan, Europe and 
the United States, in: Lazarić and E. Lorenz (eds.), Trust and economic learning, Cheltenham UK: 
Edward Elgar: 23-47.


Erlbaum.


Smith, Adam, 1982 (originally published in 1759), The theory of moral sentiments, Indianapolis: Liberty 
Classics.

Smith Ring, P. and A. van de Ven, 1992, Structuring cooperative relationships between organizations, 

----- 1994, Developmental processes of cooperative interorganizational relationships, Academy of 
Management Review, 19/1, 90 - 118.


Sydow, J., 1996, Understanding the constitution of inter-organizational trust, paper SASE conference, 
University of Geneva, 12-14 July.


Weigelt, K., and C. Camerer, 1988 Reputation and corporate strategy: A review of recent theory and 

Whitley, R., 1999, Divergent capitalisms: The social structuring and change of business systems, 
Oxford: Oxford University Press.

Williams, B., 1988 Formal structures and social reality, in: D. Gambetta (ed.): Trust; making and 


----- 1985, The economic institutions of capitalism; Firms markets, relational contracting, New York: 
The Free Press.

----- 1991, Comparative economic organization: The analysis of discrete structural alternatives’. 

486.

