Framing, attribution and scripts in the development of trust

Paper for the symposium on 'Risk, trust and civility', Victoria College, University of Toronto, 6-8 May 2005.

First draft, april 2005

by

Bart Nooteboom
Tilburg University, the Netherlands
b.nooteboom@uvt.nl /www.bartnooteboom.nl

Summary

This paper makes an attempt to improve understanding of how trust is built up and broken down in interaction between people. In particular, it investigates two dualities of conduct reported in the trust literature. One duality is that of acting in self-interest as well as in the interest of others. The second duality is that of acting deliberatively and automatically, in routinised or impulsive behaviour. For this purpose, the paper employs several notions from social psychology. One is that of mental frames that initiate and guide actions, in response to relational signals from the behaviour of others. Some of these cater to self-interest, in 'hedonic' behaviour or 'guarding one's resources', and others cater to the interest of others, in 'acting appropriately'. A second notion taken from social psychology is that of decision heuristics. These help to understand how people may attribute mental frames to others and select their own mental frames. An attempt is made to further elucidate how this may work on the basis of the notion of scripts adopted from cognitive science.

Introduction

This paper analyses 'the trust process': How, in the interaction between people, does trust build up and break down? Here, trust is interpreted as a four-place predicate: a trustor (1) trusts a trustee (2; an individual, organization or institution), in some respect of behaviour (3; competence, intentions), depending on circumstances (4) (Nooteboom 2002). The fourth point recognizes that trustworthiness has its limits, and depends on the trustee's resistance to temptations and pressures towards opportunism, which depend on situational contingencies of opportunity and threat (cf. Pettit 1995). The paper focuses on intentional trust, i.e. the assumption or expectation that the trustee will not act opportunistically. In other words, not to complicate the analysis too much, it does not include trust in competence, truthfulness, or availability of resources. Concerning intentional trust, this paper aims to incorporate two dualities recognized in the trust literature.

One duality is that trustworthiness may be based on self-interest, but also on benevolence, solidarity or loyalty. Transaction cost economics appears to deny the latter possibility (Williamson 1993), but it is widely recognized elsewhere (in sociology and the management literature). This is related to two definitions of trust. According to one definition, trust entails dependence on possibly harmful actions of the trustee, with the expectation that, for whatever reason, such harm will not be done. The reasons for this may include control, in which the trustee refrains from opportunism either because he has no opportunity for it, due to contractual or hierarchical constraints, or no incentives for it, since he is dependent on the trustor or wishes to protect his reputation. For this general notion, which includes safeguards on the basis of control, Nooteboom (2002) proposed not to use the term 'trust' but the more general term of 'reliance'. Reasons for trustworthiness may also include motives that go beyond (narrow) self-interest, such as the wish to behave appropriately, according to social or moral norms or values, or feelings of friendship, solidarity or identification with the trustor (Nooteboom 2002). The latter is what people mostly mean by the term 'trust'.

The second duality adopted from the trust literature is that of rational and automatic response. In the little that Georg Simmel wrote on trust, he recognized that while trust may be based on knowledge, it also entails a leap of faith beyond knowledge (Simmel 1950). Related to this, Pagden (1988) proposed that trust entails an information paradox: it is based on information and entails lack of information. Trust may be based on information in the form of evidence from behaviour, or reputation. However, one cannot be completely certain about a trustee's future behaviour. If one were certain, it would be odd to still speak of trust. Esser (2005) also recognized rational deliberation and automatic response as two modes of 'information processing'. However, the non-deliberative or automatic mode seems to split into two different forms: unemotional routine and emotion-laden impulse, out of faith, friendship, fear, in a leap of faith or a plunge of fear.

The duality of deliberative and automatic response has also been recognized in some of the economic literature, e.g. by Elster (1989) and Herbert Simon. According to Elster, unreflective, automatic, impulsive, emotion-laden response can make threats of reprisal credible that would not be credible under rational deliberation, in view of the high costs and risks involved. However, while impulsive response might be called rational, in that

sense, it would be odd to say that one makes a rational choice of a non-rational response. According to Simon, bounded rationality makes it rational to employ behavioural routines, so that scarce capacity can be saved for rational evaluation of unfamiliar situations. Here again, it is odd to see routinised behaviour as a rational choice, and again the question arises how it is selected. While originally Simon argued according to the economic logic of decreasing returns of intellectual attention, he later argued more in terms of the triggering of pre-established mental frames (Esser 2005: 88). Simon (1983) recognized that one may need emotions, such as fear, to break out of routinised behaviour where that turns out to be inappropriate. In sum, emotions may generate impulsive behaviour and they may trigger a break of routinised behaviour. A question then is whether the latter automatically triggers an automatic response, or whether an emotionally triggered break with routine can lead on to a rational deliberation of response. For that, the emotion would have to be somehow neutralized, controlled, supplemented, or transformed for the sake of deliberation.

In the build-up and breakdown of trust this is of particular importance in view of the indeterminacy of causation. If a relationship has been going well for some time, trust and trustworthiness may be taken for granted, in routinized behaviour. A jolt of fear from exceptional events may be needed to break out of the routine, but in view of the causal ambiguity of what went wrong, one may need to give the trustee the benefit of the doubt, allowing for mishaps or lack of competence, rather than jumping to the conclusion of opportunism. When does this happen and when does it not?

In the trust literature, it has been proposed that as a relationship develops, at some point reliance (whether it is based on control or trust) is based on cognition, i.e. on knowledge concerning the intentions and capabilities of a trustee. Subsequently, actors may develop 'empathy', i.e. understanding of how a partner feels and thinks, and next partners may develop 'identification', i.e. they see their fortunes as connected and they start to feel and think alike (McAllister 1995, Lewicki and Bunker 1996). As noted by Luhmann (1980: 32, quoted in Dietz & Nieswand 2005), when people start to cooperate, they get the chance to adopt each other's perspectives. In empathy trust may be associated with feelings of solidarity and in identification with feelings of friendship. In going from knowledge based trust to empathy and identification based trust, behaviour becomes less deliberative and more automatic.

The question then is how a relation starts, prior to the development of knowledge, empathy or identification. According to some, at the beginning vulnerability can only be covered by control (Lewicki and Bunker 1996), while according to others a relation may better be built-up by small, incremental steps that do not require much control (McAllister 1995). Starting with control may generate and perpetuate mutual distrust that may be very difficult to turn around (Deutsch 1973). Such approaches in terms of a fixed sequence of stages seem too rigid, and they do not explain how attitudes towards relations and partners arise and change.

The purpose of this paper is to further clarify the trust process, in terms of how people think and judge, making and changing interpretations and choices of action. For this, it employs the notion of mental 'framing', adopted from sociology and social psychology (Lindenberg 2000, 2003, Esser 2005). To deal with the duality of self-interest and solidarity, following Lindenberg this paper assumes two self-interested frames, one oriented towards hedonic satisfaction of urges, and one oriented towards survival, by

'guarding one's resources', and a frame of 'acting appropriately', i.e. according to social and moral norms, in order to be accepted and respected within a community. These frames may support or oppose each other, and while at any moment one frame is 'salient', in determining behaviour, conditions may trigger a switch to an alternative frame. To analyse how, more precisely, people make interpretations and select frames, use is made of decision heuristics recognized in social psychology (Bazerman 1998).

In sum, the paper aims to increase understanding of the trust process in terms of mental frames and switches between them, on the basis of relational signals and decision heuristics. Earlier, in the analysis of trust framing and relational signaling were used, among others, by Wittek (1999), Lindenberg (2000, 2003), Esser (2005), and Six (2004). Decision heuristics were used by Nooteboom (2002). Here, the two are combined: how do decision heuristics help to explain frame selection?

The paper proceeds as follows. First, on the basis of Nooteboom (2002) the paper summarizes a few basic principles for disambiguating the slippery notion of trust, which is still beset by misunderstandings and confusion. Second, the paper sets out and discusses mental framing, relational signaling, frame attribution and frame selection. Third, it reviews relevant decision heuristics. Fourth, it gives an elaboration in terms of scripts. Finally, it combines all for an analysis of the development of trust.

The basics of trust: who, what, why and when?

In spite of the large size of the trust literature, particularly in sociology, social psychology and management and organization, much ambiguity and confusion remains. Nooteboom (2002) tried to disambiguate the notion and to resolve misunderstandings, and some of the basic points are summarized here. Basic questions concern the subject of trust (who trusts?, the trustor), the object of trust (who is trusted?, the trustee), the aspect of behaviour in which someone may be trusted (what?), the foundations of trust and trustworthiness (why?), and the limits of trust (when and where?). Here and there, the term 'trust' is used loosely, also when it denotes 'reliance' rather than trust, to remain in tune with customary parlance. However, the distinction between trust and reliance will be made when needed.

The subject of trust (or reliance), or trustor, is a person, but perhaps an organization can also be seen as a trustor, as when someone says 'We don't trust X'. In the latter case, people have deliberated or gossiped among each other about how trustworthy a trustee is. This may lead to a certain reputation of the trustee among those people.

The object of trust, or trustee, may be a thing or, in behavioural trust, a person, an organization or an institution. For collaboration between organizations one needs to trust the persons involved, as well as the organization they work for. One may trust persons on a personal basis and an organization on the basis of its reputation, for example. When trusting people one needs to ensure that in their commitments they are supported by their organization, which depends on their position and role in the organization (Smith Ring and Van de Ven 1994). When trusting an organization, one needs to ensure that organizational trustworthiness is reliably enacted by the people one deals with, on the basis of organizational culture, values, norms or mechanisms of control.

To know what aspect of behaviour one may trust, one can ask the question 'What causes may there be of things going wrong?' Things can go wrong for lack of

competence, lack of resources, lack of commitment, opportunism, or accidents beyond anyone's control. Thus one may have trust in competence, means/resources, intentions (commitment, benevolence or lack of opportunism), and robustness under outside contingencies. These distinctions are important, because one will want to respond differently to limits of competence than to opportunism. Of course, when something goes wrong, one may not know which cause is at play. Opportunists will claim mishaps. When one feels very vulnerable to opportunism one may infer opportunism where in fact mere mishaps occur. This causal ambiguity of disappointed expectations or hopes will play a role in the later analysis of frame selection. The present paper focuses on intentional trust.

The distinction between different aspects of behaviour, in particular between competence and intention, is often neglected, in surveys, so that in answering survey questions some respondents may be referring to competence and others to intentions, thus confusing the survey.

Trust may have psychological causes or rational reasons. The core aim of this paper is to sort out how that works. For rational intentional trust it is important to know what reasons people may have to act in a trustworthy fashion, i.e. to act to the best of their competence to conform to the letter and spirit of agreements. Nooteboom (2002) proposed the scheme of reasons for intentional trustworthiness, or, more precisely, reliability, that is specified in Table 1.

Table 1 about here

Reasons for reliability are split in two ways. One split is between micro, particularistic, relation-specific reasons, possibly as part of the relation's organization, and macro, universalistic, institutional reasons, outside the relationship. The second split is between extrinsic self-interested motives, by which a trustor can try to control a trustee, and more intrinsic motives, whereby trustee may feel little inclination towards opportunism, from a motivation to act loyally or 'appropriately'. This distinction between self-interested motives, which yield a basis for control, and motives to act appropriately, which go beyond control, plays an important role in the trust literature. A much debated question is whether they are substitutes or complements. Does one need less contract when there is more trust? Or does one always need both, since both are limited? Klein Woolthuis et. al. (2005) showed that trust and contract are both substitutes and complements. In the present paper, the distinction between self-interested and other-directed motives, and between trust and control, is reflected in different mental frames.

According to Table 1, control on the basis of self-interest has two forms. One is to limit opportunities for opportunism, in constraint of action, by legal enforcement (macro) or hierarchical control (micro). The second is to use incentives, on the basis of reputation (macro) or the trustee's own material interest in the relation, on the basis of trustor's value to him, or costs of switching to a different relationship, or a risk of losing a hostage (micro). The role of hostages is adopted from transaction cost economics (TCE). Hostages may take the form of strategically sensitive information (that the trustor may threaten to divulge to trustee's rivals), minority shareholding, or staff seconded by the trustee to the trustor (who may be poached if the trustee misbehaves). Other-directed

reasons include institutions in the form of values and norms of decent conduct, identification with a community (macro), and routinised conduct, empathy, or identification within a relationship (micro). Empathy may carry an affect of solidarity, and identification tends to carry an affect of friendship or comradeship.

As noted before, a terminological question is whether the term 'trust' includes all the reasons, including control and self-interest, or only those that go beyond that, in other-directed motives. This paper adopts the second convention. Most people will not call behaviour trustworthy when it is based on enforcement or material self-interest. Hence the term 'reliability', which includes all motives, in contrast with trust, which goes beyond self-interest. In trust, one expects people to conform to expectations even if they have both the opportunity and incentives for opportunism (cf. Bradach and Eccles 1984, Chiles and McMackin 1996).

Finally, the question is how far trust can go: where are its limits? It can go beyond self-interest, in moral or ethical behaviour and personal loyalty, but it does have its limits, since people may succumb to temptation and may give way to self-interest under pressures of survival. Such limits of trustworthiness depend on character, organizational culture, and on conditions of survival and competition. The harsher competition is, the more difficult it is for firms to generate the slack that may be needed to make sacrifices for partners (Pettit 1995). Such conditions will form part of the mechanism of frame selection.

Mental frames, selection and attribution

According to Esser, a mental frame is an 'situation defining orientation' that consists of '.. two simultaneously occurring selections: the selection of a mental model of the situation on the one hand and that of the mode of information processing in the further selection of action'(Esser 2005: 95, present author's translation from the German). For mental frames, Lindenberg (2003) recognized three: 'acting appropriately' (AA), also called the 'solidarity frame' (Wittek 1999), 'guarding one's resources' (GR), to ensure survival, and a 'hedonic frame'(H), where one gives in to temptations for gratifying the senses.

These three frames are adopted in the present paper because they align closely with the distinction, in the trust literature, between 'benevolence' and 'opportunism', with the latter including both pressures of survival, which seems close to 'guarding one's resources', and vulnerability to temptation when it presents itself, which seems close to the 'hedonic frame'. Similar ambiguity of motives has been recognized in the economic literature in the form of 'preference reversal'. People may be honestly committed to loyalty and yet succumb to temptation when it nears. In formal modeling, this can be reproduced on the basis of hyperbolic rather than exponential discounting of future benefits, and there is independent psychological evidence that discounting does take that form. Lindenberg's three frames also align with the sources of reliability specified in Table 1.

If frames serve to both 'define a situation' (Esser) and to guide actions (Lindenberg), how are these two combined? As noted by Luhmann (1984: 157, quoted by Dietz and Nieswand 2005), in interaction people start building expectations of each others' expectations, on the basis of observed actions. According to the notion of relational

signaling (Lindenberg 2000, 2003, Wittek 1999) the actions that a trustee undertakes, triggered by a mental frame, in deliberation or automatic response, constitute relational signals that are observed and interpreted by the trustor.

Here, the following proposal is made. The trustee selects a frame, which generates actions that function as signals to the trustor, who on the basis of these signals attribute a salient frame to the trustee and selects a frame for his own response, which generates actions taken as signals by the trustee, who attributes a frame to the trustor, and selects his own frame. This yields a cycle of selection and attribution, in ongoing interaction, as illustrated in Figure 1. Note while a trustor (trustee) may select the same frame as the one attributed to the trustee (trustor), in what amounts to a 'tit-for tat response', this is not necessarily the case. One may persevere in acting benevolently in the face of opportunism, and one may opportunistically exploit the benevolent. Along this cycle, in deliberative response people may try to anticipate effects of actions, their signaling and the response in attribution, selection and action. This models Luhmann's notion of the formation of expectations of expectations. The question arises to what extent relational signals may be false, i.e. be seen to signal one frame while in fact another frame is salient.

Figure 1 about here

One question next is whether or not people have the same repertoires of frames and associated actions that enact the frame, and the same interpretation of actions they observe. According to Esser (2005: 96) frames are '.. collectively diffused cultural patterns, anchored in the thought of actors, collective representations of typical situations'. This is doubtful in so far as it suggests that different people have *identical* 'shared cultural models'. Categorizations, interpretations, views of the world and meanings can be similar and yet different between people. This point is important and can benefit from a brief excursion into the theory of meaning.

In ordinary language, typically no definition of the meaning of a word can be given in terms of necessary and sufficient characteristics. How, for example would one define 'chair'? Legs are not necessary, since there are legless pouch chairs. Chairs have seats, but so do benches, couches, and bicycles. Wittgenstein, in his later work of the 'philosophical investigations' (1976), offered the idea of 'typical cases' that represent a norm, and we deal with borderline cases by reference to the norm. Different occurrences, in different contexts, do not always share common features, let alone necessary and sufficient features, but sometimes at best only 'family resemblances'. Proximate members of a family may have shared characteristics, but distant members often do not. Members of a class form a chain, with common characteristics at each link, but no characteristic shared by all members of the class. X is in the same class as Y not because they have common characteristics but because they both share characteristics, but different ones, with a third member Z. Others have subsequently proposed similar ideas. Well known, in particular, is Rosch's (1977) idea of 'prototype', which represents an exemplar of a class that connects others in the class. Class membership is decided on the basis of resemblance to a salient case, or a typical case, which serves as a prototype.

Thus, to identify and categorize objects we use criteria from experience with similar contexts of action in the past, until we run into experience that violates the criteria, and then we change them. Johnson-Laird (1983: 189) employed Minsky's (1975) notion of 'default values' to elucidate how conventional criteria of meaning might work. Characteristics are assumed unless there is evidence to the contrary. They are assumed on the basis of established practice, i.e. on the basis of what it is possible to think, until contested by new practice, which shifts what we can think.

What, then, is the origin of mental frames? Are they instinctive, in a repertoire inherited from evolution, as part of a genotype, or do they develop ontogenetically, in the course of one's life? The view is adopted here that mental models are constructed on the basis of interaction, but individually, by assimilating perceptions in existing mental structures that have emerged from previous experience and are thereby dependent on individual mental endowments and life trajectories. Assimilation may yield a shift or accommodation of mental models (Nooteboom 2000). However, between people initial endowments are similar, to some extent, due to shared evolution. They may develop in similar ways in shared environments of action.

The evolutionary argument for a mental frame of opportunism is clear. A penchant for opportunism, to cheat whenever that can be pulled off without too much damage, yields a greater share of resources and enhances the ability to pass on the trait to offspring. However, there is also an evolutionary argument for reciprocity and sacrifice, in the evolution of man in hunter-gatherer societies. In gathering edible plants, roots, nuts, etc., and even more in hunting, there is a large variance of yields. This, together with problems of durable storage, entails an evolutionary advantage of the willingness to surrender part of one's yield to others in need, in the expectation to receive from them when they are successful (Cosmides and Tooby 1992: 212). This is enhanced by the ability to assess such willingness of others, in a cheater detection mechanism (cf. de Vos & Wielers 2003) and the ability to signal a credible threat to sanction cheating. This requires selection of group- rather than individual characteristics, which once was considered to be non-viable but more recently is considered viable under certain conditions (Ridley 1996).

In sum, mental models are to some extent similar across people but also to some extent idiosyncratic. In other words, there is 'cognitive distance' between people to the extent that they have developed their cognition along different life trajectories (Nooteboom 1992). While mental frames are stable relative to daily action, they are also subject to development, as a function of experience. Frames may be confirmed and they may adapt, particularly when they are subject to switching. An important implication will be, in the later analysis, that disappointment of expectations may yield learning, i.e. an adaptation of frames, and this may lend a positive aspect to disappointment and the breach of trust. As noted by Luhmann (1980), in their interaction people may learn and widen their horizons.

The following questions remain:

1. How, more precisely, do frame selection and attribution take place
2. How does frame selection lead to action?
3. What determines automatic or deliberative response (in selection and attribution)

For an attempt at answers, use will be made of decision heuristics recognized in social psychology. These are summarized in the following paragraph.

Decision heuristics

Decision heuristics help people to interpret their world and to take action, and hence they should help in understanding the attribution and selection of frames. For a survey see e.g. Bazerman, 1998, and for further elaboration e.g. Kahneman, Slovic & Tversky 1982. Here, they are classified as pertaining to attribution and selection respectively.

Heuristics in attribution:

- *Representativenes*: the likelihood of an event is assessed by its similarity to stereotypes of similar occurrences. This is related to the role of defaults and prototypes discussed earlier. We recognize something according to the likeness of some focal features to those of a prototype, which may be a stereotype, and on the basis of that we attribute other features from the stereotype that are not in fact present. This can easily yield prejudice.
- *Availability heuristic*: people assess the probability and likely causes of an event by the degree to which instances of it are 'readily available' in memory, i.e. are vivid, laden with emotion, recent and recognizable. Less available events and causes are neglected. This contributes to impulsive behaviour.
- *Confidence and sense of control.* People in an underdog position, or with limited self-confidence, may be overly suspicious of opportunism and power play (the 'Calimero syndrome'). However, when subject to extreme and unavoidable power they may indulge in a false sense of control, to reduce anxiety, by *mis-attributing blame* to oneself, and *condoning malefactors.* With excess self-confidence, one may overestimate one's ability to survive and underestimate opportunism.

Heuristics in selection:

- *Anchoring and adjustment*. Judgment is based on some initial or base value ('anchor') from previous experience or social comparison, plus incremental adjustment from that value. People have been shown to stay close even to random anchors that bear no systematic relation to the issue at hand. First impressions can influence the development of a relation for a long time (Deutsch 1973).
- According to the *Endowment effect*, people often demand more money to sell what they have than they would be prepared to pay to get it.
- Related to this, according to *prospect theory* people are not risk-neutral, but can be risk-taking when a decision is framed in terms of *loss*, and risk-averse when it is framed in terms of *gain*. They are prone to take more radical action to prevent loss than to achieve gain.
- *Escalation of commitment.* Often, in violation of rational behaviour, sunk costs, such as sacrifices made in a relationship, are not seen as bygones that should be ignored in an assessment of future costs and benefits. They are seen as sacrifices that would be seen as in vain if one pulls out after having incurred them. Thus,

failures may lead to deeper entrenchment in past decisions, rather than the decision to cut one's losses. It is associated with *cognitive dissonance*: cutting one's losses and pulling out would entail an admission of failure, of having made a bad decision in the past.

- In *enacting a just world* one feels that by acting benevolently one confirms a benevolent world, which is reassuring.

While the heuristics can yield errors and substantive irrationality, they may be procedurally or adaptively rational. In view of uncertainty, bounded rationality and the need to take rapid action under crisis they may well be adaptive, contributing to survival.

Concerning the *representativeness* heuristic, recall the role of prototypes in language and categorization that was discussed earlier. We categorize by comparing observed characteristics with those of a prototype, and even when there is only a partial and superficial fit we may also attribute unobserved characteristics that belong to the prototype. The mechanism of attributing unobserved characteristics upon recognition of observed ones enables fast pattern recognition that can be conducive to survival, but it does yield prejudice.

How much and how fast we jump to conclusions depends on the heuristic of *availability*: we often pay attention only when there are emotion-laden triggers. To the extent that we experience stress and a need for urgent action we will make attributions more impulsively according to prototypes that may be very inaccurate stereotypes. In the discussion of routines it was noted that an emotional appreciation of a suspicious event may be needed to jolt behaviour out of its routine. If we did not apply such filters our consciousness would likely be overloaded. The analysis of trust building indicated the importance of empathy and identification. This is clearly related to the availability heuristic: behaviour that one can identify with is more 'available'. This affects both one's own trustworthiness, in the willingness to make sacrifices for others, and one's trust, in the tolerance of behaviour that deviates from expectations. One will more easily help someone when one can identify with his need. One can more easily forgive someone's breach of trust or reliance when one can identify with the lack of competence or the motive that caused it. One can more easily accept the blame for oneself. Since one can identify with him, one may sympathise with his action, seeing perhaps that his action was in fact a just response to one's own previous actions.

Concerning *anchoring and adjustment*, under uncertainty and the need for speedy response, to avoid 'paralysis by analysis', cognition does need such an anchor, as a reference point or default for action. Then, taking the most recent value of a variable, or a value observed in behaviour of people in similar conditions, with whom one can empathize or identify, may well be rational. The notion of a default entails that one adapts past guidelines for behaviour on the basis of new evidence. Incremental adjustment can be inadequate, but so can fast adjustment. Studies of learning and adjustment have shown that hasty and large departures from existing practices can yield chaotic behaviour (Lounamaa & March 1987). Heiner (1983) argued that behaviour becomes more, not less predictable under increased volatility of conditions, since then people fall back more on their existing default action. Anchoring is clearly related to the automatic, as opposed to deliberative, behaviour discussed earlier.

The *endowment effect* and *loss/gain frame* entail, among other things, that in a relation people will accept a greater risk of conflict when they stand to incur a loss, such as a break of the relation, than when they stand to obtain a benefit, e.g. in an new, alternative relation. This may contribute to loyalty and stable relations, as follows. Relations typically end when one of the partners encounters a more attractive alternative, while the other partner wants to continue the relation. The first partner is confronted with a gain frame, the second with a loss frame. This may cause the second partner to engage in more aggressive, risky behaviour, to maintain the relation, than the first partner, who may be more willing to forego his profit and run less risk of a harmful separation procedure. One wonders what the adaptive rationale of this difference between a gain- and a loss-frame is, if any. Perhaps it lies precisely in the effect just mentioned: it reduces defection and thereby stabilizes relationships.

Concerning *escalation of commitment* also, one cannot say that this mechanism is always bad, because it yields perseverance in the face of setbacks, which can be a good thing, and is in fact a trait of many a successful innovating entrepreneur. Research shows that when someone not involved in the initial commitment decision is faced with the decision whether or not to cut losses, or when the threat of an admission of failure in earlier decisions is removed, the rational decision to pull out is made. This phenomenon can be connected with the effect of a loss frame versus a gain frame. The person, or group, that made the initial decision experiences a loss frame, with the inclination to accept further risk in order to prevent acceptance of the loss. The decision maker who enters fresh experiences a gain frame, to make a decision that will offer profit in the future, regardless of past sunk costs, and will be less inclined to accept the high risk of continuing losses from sticking to past decisions. However, while this heuristic may have its adaptive rationality, it does also lead to the continuation of a relationship where it is not beneficial.

For a *sense of control*, if it is perceived to be impossible or very difficult to influence or avoid someone's else's damaging behaviour, one may attribute blame to oneself. By doing that, one relieves the stress of feeling emprisoned in the power of others. For people with little self-confidence or a low self-image, this is a move of desperation, and self-blame fits with the preconception one had of oneself. For people with self-confidence, self-blame may yield a sense of control: if the cause lies with oneself, one can more easily deal with it. Of course, that may be an illusion, due to overconfidence in oneself. By enacting justice, even anonymously, one confirms its existence by contributing to it, and thereby maintains a sense of security. However, when the sacrifice for another would be too high to accept, in the view of self-interest, then to avoid a self-perception of callousness one may convince oneself that his hardship is his own fault.

Scripts and attribution

A frame is linked with action repertoires that enact the frame, i.e. how to act appropriately, or to guard resources, or to satisfy hedonic urges. It may be useful to conceptualise this in terms of scripts, i.e. architectures or graphs of nodes that represent component activities, with the connections or 'edges' between the nodes representing temporal, causal or logical sequence. Shank and Abelson (1977, 1995) argued that much of our action-oriented cognition can be modeled as scripts, with the restaurant script as a paradigm example, with its successive nodes of entering, seating, ordering, eating,

paying, and leaving. Each node of a script has repertoires of actions associated with the node, which each, in turn, can be represented as sub-scripts. For example, there are alternative modes of paying in the payment node of a restaurant script. A restaurant may be seen as a node in larger superscripts in the supply of customers and goods.

The notion of scripts helps, among other things, to clarify different levels and forms of learning and innovation (Nooteboom 2000):

- changes in the subscript associated with an existing action associated with a node (e.g. how to use a credit card for payment)
- novel subscripts in a node (e.g. introduction of a chip card as a new mode of payment)
- novel nodes, with a re-arrangement of new and old action repertoires across nodes (e.g. adding a floor show in a restaurant script)
- novel architectures of existing nodes (Henderson and Clark 1990) (e.g. moving from service to a self-service restaurant)
- novel architectures of old and new nodes (in a new kind of retaurant, e.g. where customers participate in the cooking).

Take the innovation of going from a service to a self-service restaurant. As illustrated in Figure 1, this entails a change of the order of nodes that largely remain the same. In self-service the sequence is: entering, selecting food, paying, seating, eating and leaving. However, under such change of order, the repertoires of subscripts associated with a node do not remain identical. While paying one has to somehow balance a tray with food.

Figure 1 about here

In the present paper, scripts are used to model, for a more detailed analysis, the working of decision heuristics in the selection and attribution of frames. Mental scripts constitute our absorptive capacity (Cohen and Levinthal 1990): we can make sense of actions that we can fit into mental scripts, and we are likely to be blind to actions that we cannot fit anywhere. Attribution on the basis of relational signaling may now be seen as the attribution of a script, triggered by the observed fit of observed actions into one or more nodes of that script. Such a script is next seen as representing, or belonging to, the repertoire of actions associated with a mental frame. This is how observed conduct is interpreted as signaling what the trustee's salient frame is.

Note, however, that scripts are never complete in regulating action and interpretation.

In the restaurant script, what happens if a dog enters? There is no prescribed behaviour. But in some restaurants it will be allowed if the dog is accompanied and lies under the table. Some people may sneak food to it. In the US some restaurants provide 'doggy bags' to take home remaining scraps of food. Such eventualities are not provided for, but neither do they have to be taken as excluded. What happens if a goat walks in? That will probably be forbidden. What is not prescribed is left to the discretion of the management, and will vary with who is in charge. (Nooteboom 2000: 127)

Everybody will have his personal mental experiences and hence associations with any script, such as having taken a dog to a restaurant, which is shared with few if any other people. This elaborates the earlier claim that mental structures are never identical between different people.

Attribution of a script and mental frame can be quite false, in several ways. Firstly, one can only fit actions that one is familiar with. One is likely to be blind to unfamiliar actions that do not fit anywhere. Secondly, one can attribute an action only to a node for which the action is recognized as belonging to its repertoire. For example, people not being familiar with using smart cards for paying may find it difficult to interpret it as a payment: perhaps they think the action with the card is an identity check, for security reasons. Third, attribution can be more or less hasty or 'gappy', with attribution of an entire script triggered by the attribution of only one of its nodes, on the basis of only one of the actions that may belong to that node. This models the possible prejudice of the representativeness heuristic. In particular, in the present context of the build-up and break-down of trust, a given action, such as failure to comply with expectations, may be interpreted as evidence of opportunism, which may then yield the attribution of the AA or the H frame to the trustee, while in fact the failure was due to lack of competence, or some accident.

Earlier, it was noted that frame selection can be automatic or deliberative. That applies also to frame attribution. In automatic attribution, there is risk of the third error indicated above: a script and its attendant frame are attributed on the basis of scant evidence, i.e. the fit of observed action into only a single node from which an entire script and corresponding frame is inferred. In deliberative attribution, a trustee is given the benefit (or liability) of doubt, treating the attribution of a script as a hypothesis rather than a foregone conclusion, with a more systematic scrutiny of evidence for other nodes that pertain to the script, and their architecture, either building up to the entire script, or yielding evidence that the action is also consistent with a different script.

Now, automatic attribution can be of two kinds, routinised, in *anchoring and adjustment,* or on emotionally driven impulse, i.e. on the basis of *availability.* Routinised attribution, in particular, is driven by symbolic behaviour, where certain actions are seen as salient and as being representative of a certain script and mental frame. This is the function also of rituals: to confirm anchors and re-confirm the salience of a given frame. Impulsive attribution is triggered, in particular, by a rush of fear from the suspicion from impending *loss* of resources. As indicated by prospect theory, loss can generate impulsive and emotion-laden response. The adaptive rationality of this is that it serves to trigger speedy action in order to avert impending disaster. The attribution of opportunism is then the most *available* one, most likely yielding attribution of a GR or H frame to the trustee. However, that is not necessarily the case, as will discussed later. Emotional impulse may also trigger compassion or pity, which may trigger an AA frame even under threat of survival.

In sum, scripts may serve to further clarify heuristics of representativeness, anchoring, automatic and deliberative response, loss vs. gain and availability, in the attribution of a frame to a trustee.

Selection, attribution and scripts

The salience, and hence stability, of a frame, and the likelihood of switching to a subsidiary frame, depends on whether it is supported by those other frames. For example, acting appropriately, in a trustworthy fashion, is most stable when it also builds resources and satisfies hedonic drives. One will switch to a frame of self-interest, when temptation or pressure exceed one's ability to resist. Conversely, one will switch from a self-interested to an other-directed frame when threat or temptation subside and loyalty assumes more prominence. The decision heuristics, and their elaboration in terms of scripts, may be used to understand how this happens.

Attribution of a self-interested frame (H, GR) to the trustee seems likely to trigger the defensive selection of a similar frame by the trustor, particularly when the attribution is based on availability by fear of loss, in what amounts to a 'tit for tat' strategy. However, that is not necessarily the case, even when the attribution is automatic rather than deliberative. People may control a shock of fear of loss and stick to an other-directed frame (AA), in several ways. Firstly, such a response may be deliberative, in the realization that a misinterpretation may be at play, with a mis-attribution of opportunism where in fact a mishap or lack of competence may be the cause of failure. However, this may be a psychologically difficult feat to achieve, and one may need the sobering caution from a third party or go-between[1].

The trustor may respond with a different frame from the one he attributed to the trustee, and both attribution and selection may be automatic, in the two ways of routinised or impulsive response, or deliberative. Three frames for attribution and selection (AA, GR, H), in three modes (routinised, impulsive, deliberative) yield 81 logically possible action-response combinations, as illustrated in Table 2.

Table 2 about here

Deliberative attribution entails rational inference of scripts and frames, and deliberative selection typically entails game-theoretic analysis of projected response in attribution to chosen actions. Here, the connection between action scripts and mental frames may be confounded in 'interest seeking with guile': one may choose actions that belong to scripts that enact an AA frame, while in fact one's salient frame is GR.

An illustration of deliberative attribution of AA is given in Deutsch's scheme for inferring trustworthiness, reproduced in Table 3.

Table 3 about here

Impulsive attribution combined with impulsive frame selection will tend to yield instable relations, while routinised attribution in combination with routinised selection, if attributed and selected frames are the same (lie on the diagonal of the table) is likely to

[1] See Nooteboom (2002) for an analysis of roles that go-betweens can play in the building and maintenance of trust.

result more in stable relations. For example when one routinely interprets observed actions as opportunistic and routinely selects an opportunistic frame. If the trustor routinely attributes benevolence, signaling an AA frame, then a routine selection of an opportunistic frame is not likely to arise, since the trustee is not likely to maintain his position as a 'sucker'. Routinised mistrust, with a routine attribution of opportunism, is not likely to result in a routinised selection of AA, but may yield either routinised or impulsive selection of GR.

The heuristics of *anchoring and adjustment, escalation of commitment, cognitive dissonance* and a *false sense of control* all contribute, in different ways, to the stabilization of frames and hence of relations. So, in a more indirect fashion, does the difference between *loss and gain,* as noted before.

Between agents there is cognitive distance to the extent that they have different bundles of mental frames and corresponding action scripts, and different inclinations towards automatic or deliberative response. The analysis demonstrates the importance of empathy, for correct attribution, on the basis of knowledge of the trustee's idiosyncracies of conduct and thought, and his strengths and weaknesses, in competence, loyalty, and resistance to temptation and pressures of survival.

In addition to refinement of decision heuristics, scripts may help to specify further how frame selection and attribution may work. The proposal here is to see frames as superscipts that trigger subscripts of action that enact the frame, and to see attribution as the attempt to fit observed actions into a familiar script, and the script then triggers a frame as a superscript, in which the observed action fits as a subscript. This is illustrated in Figure 3, which yields an elaboration of Figure 1.

Figure 3 about here

Foe example, one may try to interpret an action as enacting the frame of acting appropriately. For example, the trustee's openness about a mistake is seen as fitting into the set of actions that belong to acting appropriately. In deliberate attribution one carefully tests assumptions concerning the attribution of scripts, considering whether other actions confirm that script, and whether the action may also fit alternative scripts. In routine attribution one attributes without much consideration, according to past anchors, and in impulsive attribution one tries to fit actions into scripts that are 'available' on the basis of fear or other emotion.

The notion of 'salience' is now doubled. The trustor acts from a salient frame of his that was in place before, and operates as a default, until frame attribution to the trustee triggers a switch to a subsidiary frame. Similarly, one tries to attribute scripts that enact a frame that is salient concerning the trustee's behaviour, and operates as a default, until there is a switch, as the result of deliberative or automatic attribution of scripts.

Table 2 is based on the assumption of a stable set of symmetric mental frames, while in fact people may differ not only in details of action scripts and their mental representations and associations, but also in repertoires of scripts and mental frames, particularly in cross-cultural interaction.

From interaction, including the disappointment of expectations, one may learn and innovate in several ways. One may discover new variations upon existing scripts, a new allocation of scripts across mental frames, novel scripts or even novel mental frames. This learning may serve for a better attribution of frames to trustees, and for an extension of one's own repertoires of action and mental frames. Here, even the breach of trust may be positive, and may be experienced as such.

Conclusions

This paper has combined three streams of thought: mental framing, with relational signaling, and decision heuristics from social psychology, and the notion of scripts from cognitive science. Decision heuristics clarify how attribution and selection of frames may take place, and scripts further specify how this may work. In particular, the analysis serves to clarify the following phenomena.

According to prospect theory, in line with the principle of availability, a threat of loss is more likely to yield an impulsive attribution of opportunism, and an impulsive selection of a GR frame, than an opportunity of gain, which allows more time and attention for a deliberative attribution and selection. This explains the familiar idea, in the trust literature, that 'trust comes on foot and leaves on horseback'. It also explains the 'Calimero effect', where under asymmetric dependence and power the weaker party is more prone to suspicion of opportunism and defensive behaviour, because he is more in a loss situation than the more powerful side is.

The analysis also reconstructs the importance of openness, in the generation of trust (Zand 1972). Openness is needed to avoid misattribution of frames and inappropriate frame selection. Misattribution arises, in particular, when mishaps or mistakes are misinterpreted as evidence of opportunism, particularly by a trustee who feels insecure and weak relative to the trustee, which yields high availability of a frame of loss and defensive guarding of resources.

This analysis also demonstrates the fragility of unbalanced relationships, with unilateral power, because of one-sided sensitivity to loss, so that ceteris paribus balanced relationships are preferable to unbalanced ones.[2]

While the analysis serves to interpret or reconstruct stylized facts from the trust literature, it is conceptual/theoretical rather than empirical. In further research the analysis should be developed into testable implications and confronted with tests. Those may take the form of statistical surveys, case studies or experiments. The analysis also seems to provide a conceptual basis for agent-based simulation of attribution and frame selection, in ongoing interactions of adaptive agents.

In particular, a next step in research could be the application of the conceptual framework for a more systematic analysis of damage control. After identifying how things can go wrong, in attribution and selection, what measures could we find to prevent damage and to redress damage and repair relationships? As indicated above, this will yield the importance of openness, among other things. This analysis may yield hypotheses concerning trust building actions that could be tested empirically.

[2] This is an issue in the alliance literature. For a survey, see Nooteboom (2004).

Figure 1 Cycle of frame selection and attribution

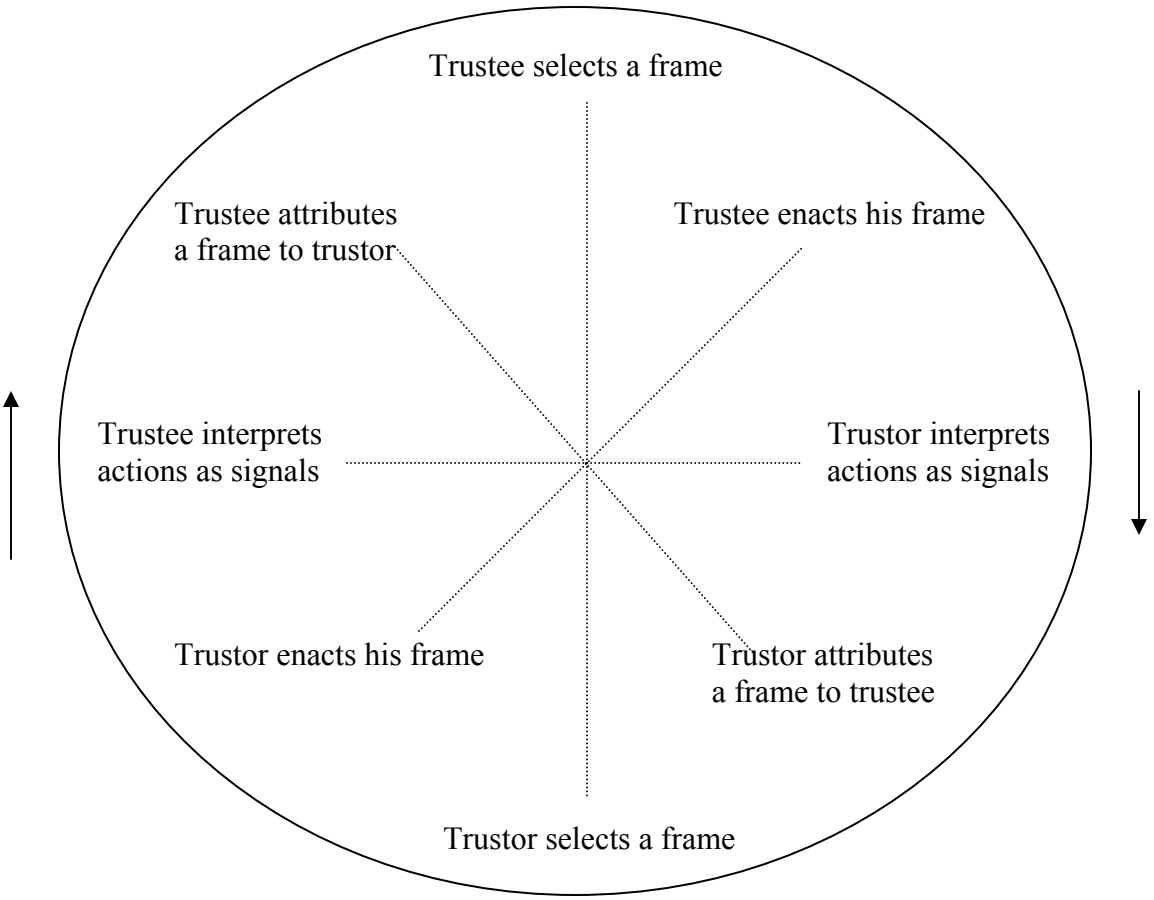

Figure 2 Service and self-service scripts

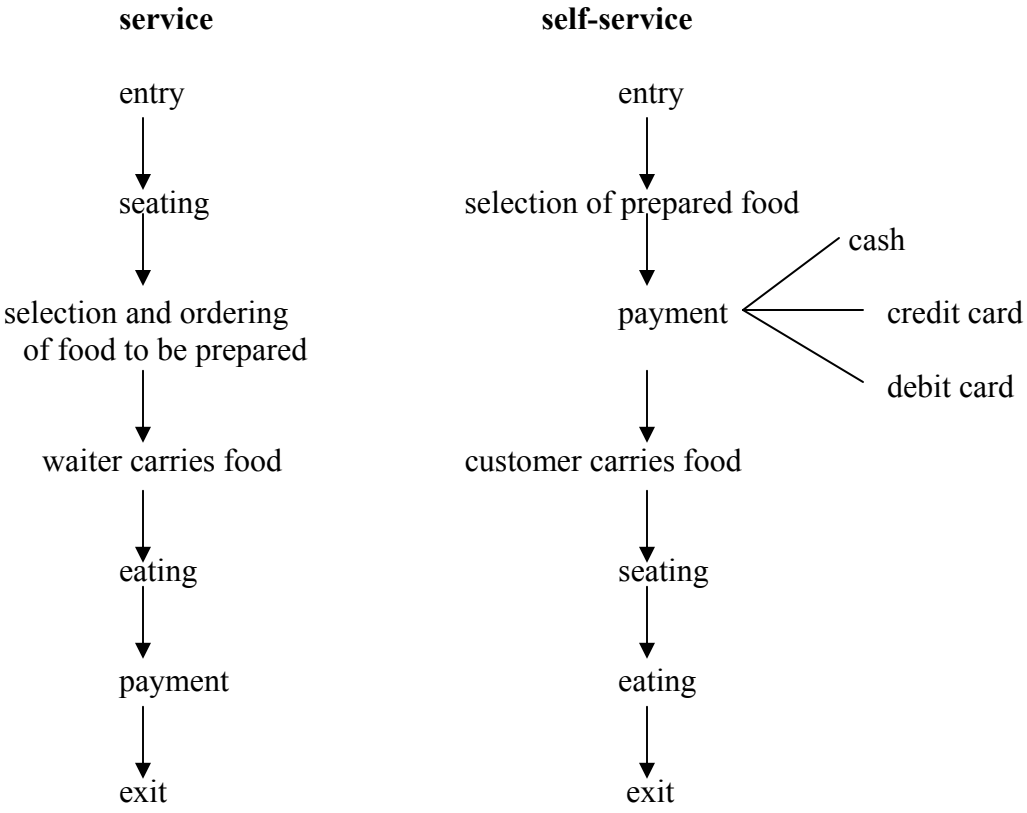

Figure Cycle of frame selection and attribution with scripts

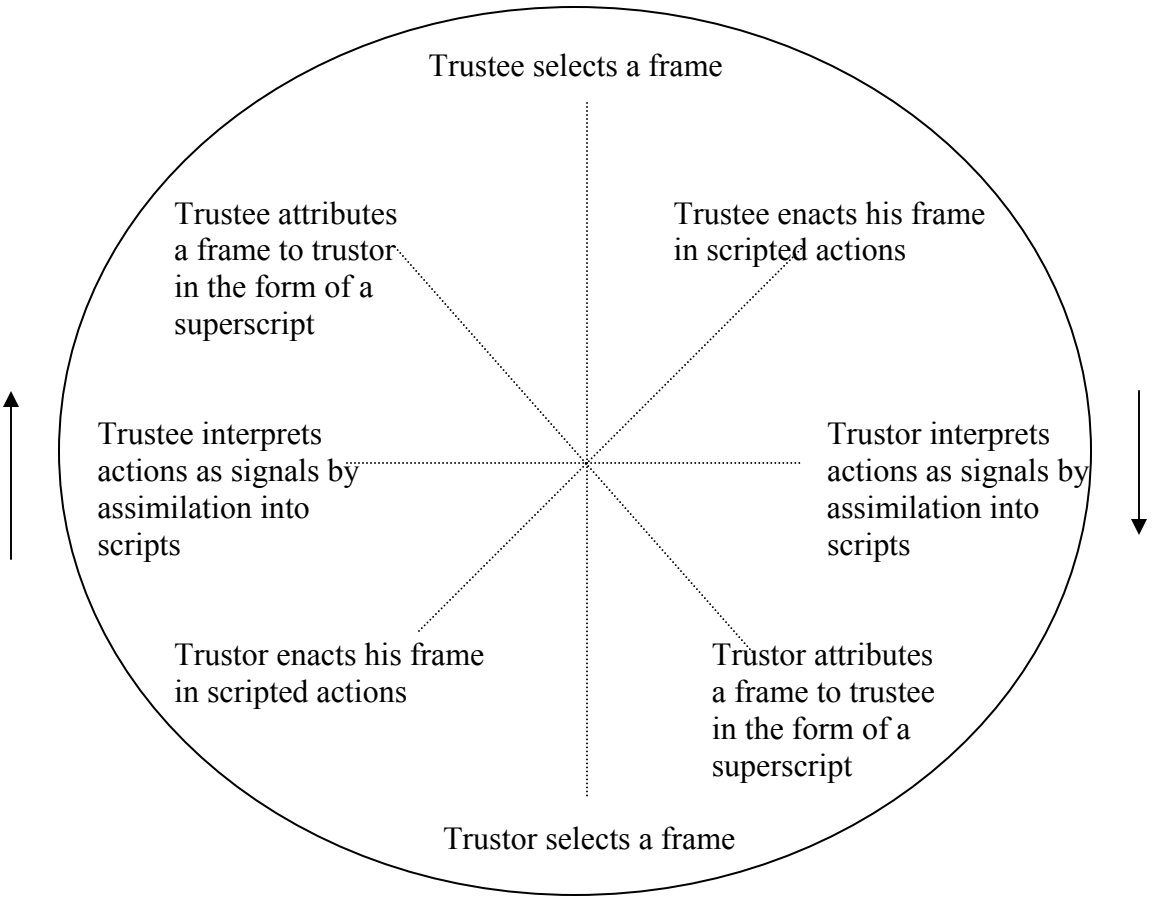

Table 1 Sources of (intentional) reliability

	macro; universalistic institutional	micro; particularistic, relation-specific organizational
self-interest opportunity control	contracts, legal enforcement	hierarchy, managerial 'fiat',
incentive control	reputation	dependence: unique partner value, switching costs, hostages
altruism benevolence	social/moral values/norms of proper conduct, sense of duty, bonds of kinship	empathy, routinisation, identification, affect, friendship

source: adapted from Nooteboom (2002).

Table 2 Attribution and selection

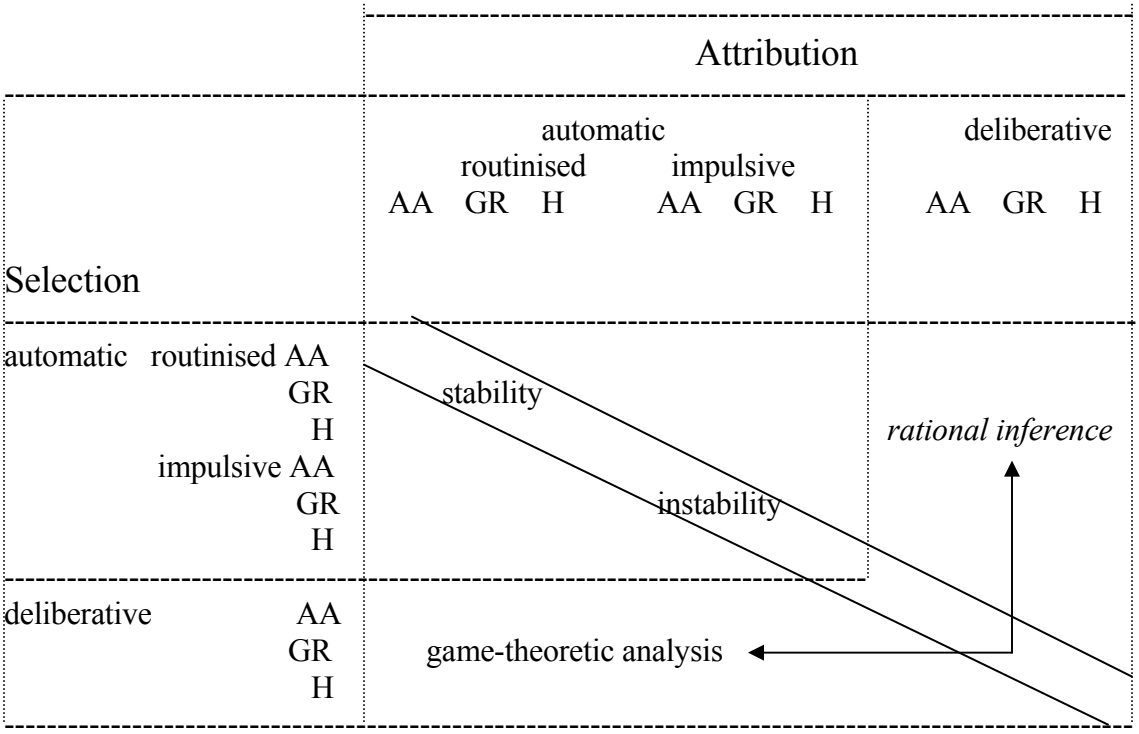

Table 3 Inference of trustee's benevolence (AA)

---

1. was the positive outcome of his action intended by the trustee, or was it an unintended result of his action?
2. did the action entail significant risk to him?
4. was he aware of the risk, and was it not neglected out of impulsiveness?
5. did he attach a positive value to this risk, out of masochism, sensation, (self) image?
6. did he have a choice, or was the action dictated by compulsion or conformity?
7. was it out of enlightened self-interest?
8. was it out of enjoyment of trust relations?
9. was it out of habitual, routinised conduct
10. or based on morality, ethics, friend- or kinship

---

source: adapted from Deutsch (1973)

References

Barkow, J.H., L. Cosmides & J. Tooby (Eds,1992), The adapted mind (pp. 19-136). Oxford: Oxford University Press.

Bazerman, M. (1998), Judgement in managerial decision making, New York: Wiley.

Bradach, J.L. and R.G. Eccles (1984), 'Markets versus hierarchies: From ideal types to plural forms', in W.R. Scott (ed.), Annual Review of Sociology, 15: 97-118.

Chiles, T.H. and J.F. McMackin (1996), 'Integrating variable risk preferences, trust and transacton cost economics', Academy of Management Review, 21/7: 73-99.

Cohen, M. D. and D. A. Levinthal (1990). 'Absorptive capacity: A new perspective on learning and innovation', Administrative Science Quarterly 35, 128 - 52.

Cosmides, L. and J. Tooby (1992), 'Cognitive adaptations for social exchange', in H. Barkow, L. Cosmides and J. Tooby, The adapted mind, Oxford: Oxford University Press: 163-228.

Dietz, T. and H. Nieswand (2005). Cognitive expectation structures and inter-firm relationships, discussion paper.

Deutsch, M. (1973), The resolution of conflict: constructive and destructive processes, New Haven: Yale University Press.

Elster, J. (1989), The cement of society., A study of social order, Cambridge UK: Cambridge University Press.

Esser, H. (2005), 'Rationalität und Bindung – Das Modell der Frame-Selektion und die Erklärung des normativen Handlens', in M. Held, G. Kubon-Gilke, R. Sturm (eds), Normative und institutionelle Grundfragen der Okonomik, Jahrbuch 4, Reputation und Vertrauen, Marburg: Metropolis: 85-112.

Henderson, R. M. and K. B. Clark (1990). 'Architectural innovation: The reconstruction of existing product technologies and the failure of established firms', Administrative Science Quarterly, 35, 9 - 30.

Heiner, R.A. (1983). The origin of predictable behaviour, American Economic Review, 73/4: 560-595.

Johnson-Laird, P. N. (1983), Mental models, Cambridge: Cambridge University Press

Khaneman, D., P. Slovic and A. Tversky (eds 1982). Judgment under uncertainty: Heuristics and biases, Cambridge UK: Cambridge University Press.

Klein Woolthuis, R., Hillebrand, B. and Nooteboom, B. (2005). 'Trust, contract and relationship development', Organization Studies, forthcoming.

Klos, T.B. and B. Nooteboom (2001), 'Agent-based computational transaction cost economics', Journal of Economic Dynamics and Control, 25: 503-526.

Lindenberg, S. (2000), 'It takes both trust and lack of mistrust: The workings of cooperation and relational signalling in contractual relationships', Journal of Management and Governance, 4: 11-33.

Lindenberg, S. (2003). 'Governance seen from a framing point of view: the employment relationship and relational signalling'. In: B. Nooteboom and F.E. Six, 2003. The trust process; Empirical studies of the determinants and the process of trust development. Cheltenham UK: Edward Elgar, pp. 37-57.

Lounamaa, P. H. and J. G. March (1987). 'Adaptive coordination of a learning team', Management Science, 33: 107 - 23.

Luhmann, N. (1980). Rechtssociologie, 2, extended edition, Reinbeck bei Hamburg.

McAllister, D.J. (1995), 'Affect- and cognition based trust as foundations for interpersonal cooperation in organizations', Academy of Management Journal, 38/1: 24-59.

Minsky, M. (1975), 'A framework for representing knowledge', in P.H. Winston (ed.), The psychology of computer vision, NY: McGraw-Hill.

Nooteboom, B. (1992). 'Towards a dynamic theory of transactions', Journal of Evolutionary Economics, 2, 281 - 99.

Nooteboom, B. (2000). Lerarning and innovation in organizations and economies, Oxford: Oxford University Press.

Nooteboom, B. (2002). Trust: Forms, functions, foundations, failures and figures, Cheltenham UK: Edward Elgar.

Nooteboom, B. (2004). Inter-firm collaboration, learning and networks: An integrated approach, London: Routledge.

Nussbaum, M.C. (1986). The fragility of goodness, Cambridge UK: Cambridge University Press.

Pagden, A. (1988), 'The destruction of trust and its economic consequences in the case of eighteenth-century Naples', in D. Gambetta (ed.), Trust, the making and breaking of cooperative relations, Oxford: Blackwell: 127-141.

Pettit, Ph (1995), 'The virtual reality of homo economicus', The Monist, 78/3: 308-329.

Rosch, E. (1977), 'Human categorization', in N. Warren (ed.): Advances in cross - cultural psychology, vol.1, New York: Academic Press.

Shank, R. and R. Abelson (1977). Scripts, plans, goals and understanding, Hillsdale: Lawrence Erlbaum.

Shank, R. and R. Abelson (1995). 'Knowledge and meaning: The real story', in R. S. Wyer jr. (ed.), Advances in Social Cognition, Hillsdale NJ: Erlbaum, 5 - 81.

Simon, H.A. (1983), Reason in human affairs Oxford: Basil Blackwell.

Six, F. (2004). Trust and trouble; Building interpersonal trust within organizations, PhD dissertation, Erasmus University Rotterdam.

Smith Ring, P. and A. van de Ven (1994), 'Developmental processes of cooperative interorganizational relationships', Academy of Management Review, 19/1: 90 - 118.

Tversky, A. and Kahneman, D. (1983). Probability, representativeness, and the conjunction fallacy, Psychological Review*, 90/4:* 293-315.

De Vos, H. and R. Wielers (2003). Calculativeness, trust and the reciprocity complex: is the market the dimain of cynicism?, in B. Nooteboom and F. Six (eds), The trust process in organizations, Cheltenham: Edward Elgar: 75-104.

Williamson, O.E. (1993), 'Calculativeness, trust, and economic organization', Journal of Law & Economics 36: 453-486.

Wittek, R.P.M. (1999). Interdependence and informal control in organizations, PhD dissertation. University of Groningen, the Netherlands.

Wittgenstein, L. (1976, first published in 1953), Philosophical investigations, Oxford: Basil Blackwell.

Zand, D.E. (1972), 'Trust and managerial problem solving', Administrative Science Quarterly, 17/2: 229 - 239.