Problems of trust

PROBLEMS OF TRUST
A QUESTION OF TRUSTWORTHINESS

An ethical inquiry of trust and trustworthiness in the context of the agricultural and food sector

Vertrouwensproblemen: een kwestie van betrouwbaarheid
Een ethische analyse van vertrouwen en betrouwbaarheid in de context van de landbouw- en voedselsector
(met een samenvatting in het Nederlands)

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Universiteit Utrecht
op gezag van de rector magnificus, prof. dr. J.C. Stoof,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen
op vrijdag 12 september 2008 des middags te 12.45 uur

door

FRANCK LOURENS BERTUS MEIJBOOM

geboren op 5 november 1975 te Rotterdam

Promotoren:

Prof. Dr. F.R. Heeger

Prof. Dr. F.W.A. Brom

Voorwoord

Dit proefschrift gaat over vertrouwen en betrouwbaarheid. In het dagelijkse leven vertrouwen we op vele anderen: familie, vrienden en collega's, maar ook op experts die we niet persoonlijk kennen, op bedrijven en overheden. In veel gevallen zijn we ons niet bewust van dat vertrouwen. Daarom is het opmerkelijk dat vertrouwen volop in de belangstelling staat. Dit geldt bijvoorbeeld voor het vertrouwen in de veiligheid van voeding. Enerzijds zijn de methoden om voedselveiligheid te beoordelen de laatste jaren alleen maar beter geworden en is de informatie die we als consumenten daarover krijgen alleen maar toegenomen. Anderzijds blijft het vertrouwen dat consumenten hebben in de veiligheid van hun voedsel kwetsbaar. Die spanning is niet alleen terug te vinden in de voedselsector, maar ook in andere sectoren van de samenleving. Mijn stelling is dat dit niet gezien moet worden als een fout van de consument of de burger, maar als een probleem van degene die vertrouwd wil worden. Het is een betrouwbaarheidsvraagstuk. Voor die stap van vertrouwen naar betrouwbaarheid heb ik conceptuele en morele argumenten. In het proefschrift werk ik die argumenten uit en analyseer ik wat betrouwbaarheid inhoudt.

Wie dit boek in handen heeft ziet het resultaat van het onderzoek naar vertrouwen en betrouwbaarheid. Wat u niet meer ziet, zijn de vele mensen die een belangrijke rol hebben gespeeld in de afgelopen jaren. Een aantal van hen wil ik graag expliciet bedanken. Allereerst wil ik Frans W.A. Brom bedanken. In 1998 wees hij mij op de landbouw- en voedselethiek als een interessant en nieuw onderzoeksgebied van de toegepaste ethiek. Het was de start van een inspirerende samenwerking. In diverse onderzoeksprojecten, adviestrajecten en bij het opzetten van de Europese vereniging voor landbouw- en voedselethiek (EurSafe) heeft hij mij enthousiast gemaakt voor de ethiek en scherp gehouden als onderzoeker. Als begeleider van dit proefschrift heeft hij niet alleen veel tijd geïnvesteerd. Hij was ook degene bij wie ik meer dan eens met halve ideeën en hele vragen kon binnenlopen om vervolgens met concretere ideeën, maar ook nieuwe vragen, weer naar buiten te gaan. Dat heeft mijn denk- en schrijfproces enorm geholpen.

De persoon die ik daarnaast veel dank verschuldigd ben, is Robert Heeger. Grondigheid en betrokkenheid kenmerkten zijn begeleiding. Zijn feedback was compleet en precies. Hij heeft niet alleen oog voor de grote lijnen, maar ook voor de 'kleintjes'. We vonden elkaar in de details, die het verschil kunnen maken. De afspraken met Robert waren zeer constructief. Hij verstaat de kunst om te sturen wanneer het moet, maar ook om de vrijheid te geven, zodat het resultaat je eigen proefschrift wordt. Dat heeft er voor gezorgd dat ik veel aan de begeleiding heb gehad. Ik vind het een eer om zijn laatste promovendus te zijn.

CONTENTS

CHAPTER 1

TRUST ON THE PUBLIC AGENDA: A QUESTION OF

TRUSTWORTHINESS

*Nach dem Aufstand des 17. Juni
ließ der Sekretär des Schriftstellerverbandes
in der Stalinallee Flugblätter verteilen,
auf denen zu lesen war, dass das Volk
das Vertrauen der Regierung verscherzt habe
und es nur durch doppelte Arbeit
zurückerobern könne. Wäre es da
nicht einfacher, die Regierung
löste das Volk auf und wählte ein anderes?*
Bertolt Brecht

1.1 Trust on the public agenda

Trust is on the public agenda. The Dutch Scientific Council for Government Policy, for instance, published a number of reports that have trust as the central theme or address it in the context of contemporary societal issues such as the introduction of biotechnology and the role norms and values play in the society (WRR, 2003a; 2003b; 2004; 2005). Furthermore, the Queen's speech from the throne of 2004 presented trust as an essential concept for both policy and a stable society. Another example is the focus on trust in the annual series of lectures on innovation of the Dutch Ministry of Economic Affairs in 2006.

This is an interesting development, especially since such attention to trust is not restricted to the Netherlands. All over Europe and in the United States trust in public and market institutions is in the spotlight. In the academic debate publications on the topic are proliferating. In 1968, Luhmann was one of the first to write on trust from a sociological perspective.[1] He suggested that sociology might profitably use words found

---

[1] Almost a hundred years earlier, Georg Simmel wrote some works that addressed the issue of trust. He never did write a monograph on trust, but his ideas on the subject have been influential in the debate up until now. Traces of his account can be recognised in the works of many authors including Luhmann and Giddens (Simmel, 1950; Möllering, 2001).

in everyday usage and concepts employed in traditional ethical contexts.[2] Numerous sociologists followed him in this, including Giddens, Coleman, and Sztompka. From the perspective of philosophy and ethics, attention to this concept is evidently intensifying. Baier's article 'Trust and antitrust' (1986), one of the first renewed attempts of the last century to address the concept that 'traditionally has an ethical connotation' (Luhmann, 1968), notes that trust remained a stranger in the field of ethics. Although moral philosophers have always been interested in cooperation among people, Baier observes that when we look at moral philosophy 'what we find can scarcely be said to be even a sketch of a moral theory of trust' (Baier, 1994a [1986], p. 97). It would be an exaggeration to claim that trust has not been discussed or analysed in moral philosophy before (e.g. Horsburgh, 1960; Løgstrup, 1997 [1956]), yet mostly it has been addressed as a kind of second–order theme, which is merely discussed in the context of another, major theme. Baier has been followed by many others such as Lagerspetz, Lahno, Potter and Hardin. They all addressed trust as a philosophically relevant concept.

This general attention to trust is remarkable, because trust is not a new phenomenon. It has a long tradition and has been considered as an essential element within a society for ages (cf. Frevert, 2003). On top of this, trust is essential for social life, but also implicit by nature. If we trust others, we mostly do not explicate our trust to the trustee. This also holds for the trustee. Once there is trust and one party starts to emphasise explicitly its trustworthiness, the truster[3] will probably wonder what reason may underlie this. For instance, if the meat industry, without any clear cause, were to state explicitly that they can be trusted never to use a specific carcinogen, one will probably will become suspicious rather than assured that they are trustworthy. Therefore, Baier compares trust to an atmosphere: we notice trust as we notice air 'only when it becomes scarce or polluted' (1994a [1986], p. 98).

All this attention to trust is a signal putting us on the alert. Trust apparently has become 'scarce or polluted'. This idea of scarcity and pollution is not just a rhetorical expression. It can be recognised in the everyday practice of trusting. Trust has become scarce in the sense that an individual is confronted with an increased need to rely on others. Many sociologists have shown the complexity of social life in our late-modern society; globalisation and technology result in changed and increased levels of risk and uncertainty (Giddens, 1990; 1991; Luhmann, 1968). This has changed the character and scope of the need to rely on others. On the other hand, scandals and affairs have polluted trust in a range of institutions. The financial scandal at Enron and the affair with respect

---

[2] Luhmann writes: 'Ob man der Soziologie raten sollte, Worte des täglichen Sprachgebrauchs und Begriffe der traditionellen ethischen Vorstellungswelt zu verwenden, ist ernsthafter Überlegung wert. Bei einer solchen Umrüstung moralischer in soziologische Begriffe scheinen zunächst Vorteile und Nachteile sich die Waage zu halten' (2000, p. v).
[3] Hereafter, I will use the terms 'truster' for the person who gives or aims to give trust to another agent and 'trustee' for the agent who is the trusted person or institution. I should mention that, unless noted otherwise, I use 'he' and 'she' as interchangeable.

to the human health consequences of the BSE-outbreak in the UK are only two of many examples of situations in which institutional agents turned out to be untrustworthy. In these cases a lack of trust or hesitance to trust can be interpreted as a 'realistic attitude towards the behaviour of institutions' (Marris et al., 2001, p. 69). These two aspects result in a problem. People have to rely on others, but often do not know whom to trust. I call this the *problem of trust*.

At this point it is important to stress that the 'problem of trust' does not imply that we are confronted with a crisis of trust. Surveys show that there is no strong indication of a real crisis of trust in governmental institutions in Europe (e.g. Gaskell et al., 2003; Kjaernes, 2004; 2006; Trustbarometer, 2008). Nevertheless, empirical research gives rise to concerns with respect to trust in political institutions. This holds for society as a whole (e.g. Dekker & Van der Meer, 2004) as well as for specific sectors such as the agricultural and food sector (cf. de Jonge et al., 2007; TNO, 2003).

In this study I analyse the 'problem of trust' as a question of trustworthiness. Rather than conceiving the problem of trust as a failure of the truster, I treat the public attention to trust as a sign of questions of trustworthiness. In the analysis I am especially interested in the moral dimensions that play a role in a trusting relationship, and I argue that trustworthiness on an institution level has to start in respect for the truster as an autonomous person.

In the course of this chapter I introduce my claims and present the outline of this study.

1.2 The problem of trust as a question of trustworthiness

The problem of trust, i.e. the problem of people who have to rely on others but often do not know whom to trust, can be addressed from two sides. First, it can be defined in terms of a problem of the individual truster. Trust enables an individual to perform actions such as buying and consuming food in spite of the uncertainty and the lack of personal control he is confronted with. For instance, although I lack any kind of knowledge regarding the possible carcinogenic effects of a certain food product, I buy and consume it. This is not because I am indifferent about my health, but because I trust certain agents in the food sector, such as the local supermarket, companies like Unilever, or the governmental regulations on food safety. Here trust is a way of 'managing uncertainty' (Becker, 1996, p. 45). In trusting, one acts 'as if' certain possible state of affairs will not occur (Lewis & Weigert, 1985). This does not imply that the problem of control and the uncertainty evaporate. However, in trusting it is possible to go beyond this uncertainty and lack of control because trust 'brackets ignorance or lack of information' (Giddens, 1991, p. 244). This acting 'as if' is not an escape in a make-believe world of certainty and control. The attitude of trust is predicated on actual belief

that the other agent is trustworthy, i.e. competent and adequately motivated to act in the expected way.

The problem of trust is often addressed as a dilemma of the individual. He is the one who is confronted with the need to rely and who is uncertain about whom he can trust. From this perspective, ways and methods to increase *trust* are introduced as the most effective way to address problems of trust. For instance, many elements in the current European food policy, such as a revised food safety policy, the establishment of food safety authorities and the increase of information services, have the aim to improve the level of trust (cf. Dreyer & Renn, 2007, pp. 551-552).

In this study however I start from the other side and argue that the problem of trust is best analysed as a problem of trustworthiness, that is, as a problem that focuses on the competence and motivation of the trustee. Confronted with the problem of trust, our question should not be 'how to increase trust?' but 'why would an individual agent trust the other agent?' and 'is this agent worth being trusted?' (Meijboom et al., 2006). For this change of focus I have three arguments.

First, there is a conceptual argument. In practice, the problem of trust often remains intangible if it is addressed as a problem of the individual truster. As long as the approaches start from the aim to change the attitude of the individual, building and regaining trust turn into a mission impossible. This has a conceptual background. As I will argue in Chapter 3, an individual cannot decide to trust. Trust results in beliefs and expectations, but is not itself a belief. One can want to trust, but one cannot trust at will. It is not a stance one simply decides to adopt all things considered. For that reason, you cannot make others trust you. To enforce trust is impossible. Nevertheless, it is possible to show oneself trustworthy.

The second argument starts in the trustee's assessment of the lack of trust as problematic. Unless a trustee hopes that he will be trusted blindly, he will acknowledge that trust usually is based on an (implicit) assessment of competence and motivation. Therefore, if a trustee assesses the lack of trust as problematic, it includes an implicit claim about his trustworthiness. If, for instance, a governmental agency worries about lack of consumer trust, this is more than just a description of the consumer's attitude. It incorporates an evaluative element about their own trustworthiness. They consider the lack of trust as problematic, because, according to them, the truster has good reasons to trust them, i.e. they consider themselves trustworthy. From this perspective, it would be too easy to define the problem of trust as a problem of the individual truster only. The trustee has a problem too. If the implicit claim is legitimate and the trustee is competent and adequately motivated to do what is entrusted, he obviously failed to make this point sufficiently clear to the truster. The problem of trust cannot be reduced to the problem of the truster only.

Finally, there is a moral argument for picking up the other end of the stick based on the autonomy of the truster. In a trusting relationship the truster is always depending on

the trustee. Trust relationships include by definition an asymmetry in knowledge and power. This results in truster vulnerability, precisely this is constitutive for trust. Without this vulnerable position, there would be no need to trust. Nonetheless, the vulnerable status does not change the moral status of the truster. As I will argue in Chapters 4 and 5, the truster should be treated as a person who is capable of autonomous agency. The truster is worthy of respect, in spite of her vulnerable and dependent position and her imperfect knowledge with regard to the object of trust.

When we assume such respect as a start in a trusting relationship, a lack of or hesitation to trust can no longer be defined as a failure of the truster only. Defining it as a failure would disregard the autonomy of the truster in two ways. First, it does not take the assessment of an autonomous agent seriously. Given moral respect for the autonomy of the truster, lack of trust or hesitation to trust should be acknowledged as a legitimate point of view, rather than as failure only. This does not imply that the truster cannot be wrong, but respect for autonomy involves a responsibility for the trustee to take this assessment seriously. Second, referring to the problem of trust only in terms of a problem of the truster ignores the trustee's responsibility with respect to the vulnerable status of the truster. Acknowledgement of the truster as autonomous implies responsibility for the trustee to take additional care for the truster in his vulnerable position. The fact that the truster is in a vulnerable position does not change his moral status. Accordingly, this vulnerability cannot be a reason for the trustee to treat him as non-autonomous, but should be an incentive to enable the truster to act as autonomously as possible. This results in a specific interpretation of trustworthiness such that exploitation and manipulation of whatever kind are excluded.

On the basis of these three arguments, I address the public attention to trust as a signal alerting us to issues of trustworthiness.

1.3 The agricultural and food sector as focus and case study

To focus on trust in institutions in general runs the risk of making the discussion rather indefinite. It raises the questions like 'whose trust?' and 'which institutions?' To forestall such vagueness I focus my analysis on a specific sector: the agricultural and food sector. The problem of trust can easily be recognised in this sector.

Since the last decades of the twentieth century, governmental authorities in Europe, but also commercial organisations that produce and process food products have been confronted with diminishing trust. Especially the many well-publicised food scandals in recent years have resulted in the problematic status of trust. Events such as the BSE affair, the introduction of genetically modified food products, and the dioxin and MPA scandals all include the two elements of the problem of trust. These events underscore the vulnerable position of consumers in the sector and their need to rely on many others

in order to perform everyday actions and demonstrate that some agents turned out to be not as trustworthy as they were supposed to be.

Consequently, there have been repeated calls for restoring and maintaining consumer trust. In 2000, the OECD reported after a meeting on GM-food that 'a strong sense emerged that there was a need to take steps to rebuild trust among the various actors, particularly governments, industry, scientists, regulatory agencies and the public' (OECD, 2000). Two years later the FAO stated, 'the food safety system … must be able to both manage risks and create trust' (FAO, 2003), and in 2004 the European Commission emphasised that there is a clear 'need to develop trust' (Byrne, 2004).

These statements are signals of the problem of trust, but are also relevant because they tend to reduce the problem to one of trust rather than conceiving the problem as one of trustworthiness. This has resulted in several measures that aim to increase levels of trust. However, these have not yet resulted in the desired effect. I take this as an indication that problems of trust in institutions have the same nature: essentially they are problems with respect to trustworthiness. Problems of trustworthiness require a different approach, as I will show in Chapters 2 and 3,

These characteristics make the agro-food sector an interesting case and a relevant starting point for my analysis of the problem of trust. Furthermore, the relevance of such analysis is not limited to this sector. The context, the dimension and the nature of the questions of trust in the food sector are similar, or at least closely linked to those in domains such as security and safety, social welfare, and health care.

1.4 The moral dimension in the problem of trust and the consequences for trustworthiness

In this study I explicitly elaborate on the fact that the problem of trust can be interpreted as a signal of questions on a moral level. I see two general contexts in which moral questions underlie the problem of trust. First, a problem of trust can be a signal of a lack of respect for the truster's position as an autonomous agent. In Chapter 5, I will argue that trusting or withdrawing trust illustrates how a person judges his own autonomy and that of the other agent. For instance, someone who rashly trusts everyone without any deliberation is easily accused of 'trusting blindly'. The problem of blind trust is not a flawed risk calculation. The truster can be faulted because he did not act in a way that fits his capacity of autonomous agency. Similarly, certain responses of a trustee can raise feelings of resentment in the truster and can result in reproaches towards the trustee. These feelings and reproaches are not merely signals of disappointment about a mistaken evaluation of the risks at stake. They start in the belief that one is wronged as a person. The trustee's behaviour is considered as incompatible with an attitude of respect for the

truster as a full participant in the trusting relationship. This lack of respect can be the reason for the truster's uncertainty about whom he can trust.

Second, the problem of trust can be a signal of a disagreement about what one can expect of each other. Entrusting and responding to trust reflect ideas on what one considers as legitimate moral expectations. This explicates that trusting is more than prediction and anticipation. Trust includes a judgement about what can be expected of another agent. This includes moral judgements. As a truster, one is not merely struggling with the practicalities of uncertainty. One also has ideas about how the trustee should react given the uncertainty and given the truster's vulnerable position. On the other hand, the trustee also has ideas and beliefs about what is morally right and wrong and about what legitimately can be expected of him. Trusting can be considered as a way of expressing moral claims and responding to trust as accepting the claims.

The emphasis on the moral dimension of the problem of trust does not mean that each problem with respect to trust can be reduced to moral questions. It only means that the reasons that underlie the problem of trust include moral questions and that a neglect of this dimension will result in ongoing problems of trust. In this thesis, I elaborate on the consequences of the recognition of this moral dimension on the level of trustworthiness.

1.5 The denial of the moral element: Trust as risk calculation

An extensive part of the literature on trust denies the moral dimension. The accounts referring to the problem of trust as a matter of rational risk calculation mostly conceive the issue as 'unmoralised' (Hardin, 1996, p. 28, also Coleman, 1990).

These accounts start out from the uncertainty of the individual and concentrate on the risk the truster runs. It is not hard to see why this focus is chosen. After all, a truster never can be sure that the trustee will behave in the favourable way. If you knew that the other agent can only act in the expected way, then you can simply anticipate on this and trust would not be necessary. However, in a case of trust there is no such certainty. In trusting you always run a risk: your trust can be harmed. Accordingly, trust is referred to as a risky matter (Gambetta, 1988, p. 235) and as a venture (Luhmann, 2000, p. 31). The close link between trust and risk leads some authors to the conclusion that trust is a bet about the future actions of others, or that trust is just a subclass of situations involving risk (Coleman, 1990, p. 91). It is argued that trust enables us to deal with uncertainty, but entails new risks at the same time. In other words, it 'copes with one type of risk by trading it for another type of risk' (Sztompka, 1999, p. 25, 32). If we frame trust in terms of risk, it is mainly a technical matter of calculation. It requires the assessment of the risks and benefits of trusting in the light of the aims and goals one pursues. Consequently, there is not much morality in this calculation. Morality only plays a role

on two levels. First, the value of the object of trust can be determined by moral reasons and thus influence the risk one runs. Second, moral norms may steer the behaviour of the trustee and thus make her more predictable, which simplifies the risk assessment. However, in both cases morality only has a procedural role to play.

This view on trust as a rather technical process of risk calculation and lacking a moral dimension is problematic. It (a) starts from a naïve interpretation of the relation between risk assessment and ethics and (b) it starts from a distorted view on the relation between risk and trust.

First, the rational-choice accounts of trust often start from a rather naïve view on the moral dimension of risk assessment. Trust as the result of a risk/benefit assessment suggests that what counts as risks and benefits is relatively clear and is knowable for all rational agents. However, identification of risks implies decisions and presuppositions that include cultural, social, political, and normative considerations. Thus, to evaluate something as a risk is not morally neutral (Jensen & Sandøe, 2002; Slovic, 1999). It presupposes a moral view on and evaluation of what is and what is not morally desirable. There is no such thing as a kind of non-moral risk assessment on which all rational agents would agree and that can serve as the undisputed input for the risk/benefit calculation that underlies trust. This equally holds for the element of benefit. This shows that a risk approach to trust cannot ignore the moral dimension of the calculation. A calculation is possible if and only if one has a moral framework as guide in the identification of risks and benefits and in the assessment of these elements. For instance, the consequentialist principle of maximising utility – often part of rational-choice accounts – can serve as such a normative guideline.

The second problem of this view lies in the claim that trusting and taking risks can be equated. Notwithstanding the relevance of trust in situations of uncertainty, trusting and taking risks are on different levels. 'Considerations about risk taking can only motivate risk taking, not trusting' (Lagerspetz, 1998, p. 56).[4] Risk calculation and trusting are two complementary, yet different mechanisms to deal with uncertainty. A risk approach aims to clarify the uncertain aspects of the situation in which one has to rely on another agent. In this context the aim is to translate the problem of known uncertainty into one of risk. Consequently, one can make a personal assessment and does not need to trust another. The ultimate aim is to prevent that trust is necessary and, if this appears to be beyond reach, to enable the individual confronted with uncertainty to calculate whether it is worthwhile taking the risk given one's own interests and preferences. Risk analysis does not provide us with reasons to trust, but can show that we are confronted with a risk although the situation appeared as uncertain at first hand. A trust approach to uncertainty, on the other hand, starts where a risk focus ends. It focuses on those situations that remain uncertain even after the uncertain aspects have turned into risk

---

[4] On the basis of empirical research Eckel & Wilson (2004) come to similar conclusions.

factors as much as possible. The aim is not to try to make the uncertain dimensions as certain as possible, but it is a way of bracketing 'ignorance or lack of information' (Giddens, 1991, p. 244) that enables us to act despite the uncertainty. A truster does not see trust as a problem of risk, but rather 'as a problem of judgement' (cf. Eckel & Wilson, 2004). Once I trust a company that offers novel foods, I do not assess risks, but I make an assessment of the company's competence and motivation with respect to the new food product. Suppose that even though the risk at stake is low or the risk information is of a high quality, I do not consider the agent on whom I have to rely to be competent or to have the proper motivation. Then I will not trust him. Maybe I consider it worthwhile taking the risk and act nonetheless. However, I take a risk, yet I do not trust. Trust is not a matter of a calculation of gains and losses, but an assessment of the trustee's competence and motivation. In this assessment risk-related arguments may serve as relevant input only as far as it serves as a signal or proof of the competence and motivation of the trustee. One can argue that the assessment of trustworthiness can be conceived as a risk calculation too. However, if the truster perceives the assessment of the other's competence and motivation in terms of what risk he runs, the assessment will not result in trust, but only in the evaluation of the risk as worthwhile taking. Therefore, in this study, I address trust as an attitude towards (collective) humans which is based on the assessment of the trustworthiness of the trusted agent and which is distinct from a risk or from risk taking. Consequently, trustworthiness is not about risk assessment, but about competence and motivation.

1.6 Trust: More than its function in a well-operating society

To make myself clear it is not only relevant that I stress the difference with those accounts that conceive trust as a matter of risk calculation. My approach also differs from accounts that explicitly ascribe a moral dimension to trusting. Authors who define trust as 'social capital' often emphasise the importance of moral values with respect to building and maintaining trust. Nonetheless, I have three general problems with the approaches that address trust as mainly a vital element of the good society. My problems concern (a) the strong emphasis on the functional element of trust, (b) the presumed restriction of the moral evaluation of trust to its contribution to a stable society, and (c) the focus on the need for a community to have shared moral beliefs.

1.6.1 The problem of reducing trust to its function

Let me start by underscoring that trust clearly has an important function in society. The practical problems of a lack of trust are evident both on the individual and the public level. A complete absence of trust would paralyse life as a whole (Luhmann, 2000, p. 1).

Organisations and communities are more effective if they are based on trust, since trust facilitates exchanges among individuals and enhances cooperation (Coleman, 1990; Fukuyama, 1995; Kramer & Tyler, 1996; Putnam, 1992). When we can confidently rely on others, we save a lot of time and effort we otherwise would have to spend on making contracts or risk calculations. Consequently, problems of trust can easily be framed in terms of problems of social cohesion and cooperation. The problem of this frame is not that it is incorrect, but that it tends to be presented as if this aspect is the core characteristic of trust. More than once, the essence of trust has been equated to its function. Kenneth (1974) for instance evaluates trust as 'a very important pragmatic value, if nothing else. Trust is an important lubricant of a social system. It is extremely efficient; it saves a lot of trouble to have a fair degree of reliance on other people's word' (pp. 22-23). I do not deny these positive effects of trust, but I do not agree with the idea that trust is just a pragmatic value, or merely an instrument which is essential to achieve certain aims. Such an instrumental approach of trust becomes even clearer in the title of the *2006 Innovation Lecture* organised by the Netherlands Ministry of Economic Affairs: 'Trust. Making innovation work.' It plainly illustrates the functional view on trust: trust as an instrument to pursue the aim of innovation.

With respect to this instrumental view on trust I see a conceptual problem. If we evaluate trust mainly in terms of its function we see simultaneously too much and too little trust. On the one hand we do not see enough, because trust includes a process and cannot be restricted to the product of that process. In focusing on the product only, we easily miss important aspects that underlie trust, such as the truster's interests and moral beliefs. Moreover, we miss those forms of trust that do not directly result in visible or measurable effects. Suppose I trust a certain food company but, because I do not need their products, the effects of my trust are go unrecognised. On the other hand, we run the risk of seeing too much if we focus on the effects of trust in society, because the effects that are ascribed to trust are not exclusive to trust. Stability, cooperation, and reducing complexity can also be achieved by other phenomena, such as power and coercion. The former Union of Soviet Socialist Republics for instance remained rather stable for over 70 years; one may seriously wonder however what role trust played in this result. Focusing on the functions of trust may lead to unwarranted attribution.

1.6.2 The problem of reducing the evaluation of trust to its contribution to a good society
The emphasis on the function of trust also has implications for the evaluation of the concept. The value of trust is regularly defined in terms of its presumably valuable contribution to a certain kind of society. Consequently, trust is conceived as 'social

capital' (cf. Coleman, 1990; Putnam 1992; 1995; Fukuyama 1995).[5] However, if trust is evaluated as desirable because of its function in a society, there is the danger of an implicit and unjustified fact–value leap. From an empirical perspective one can argue that trust is constitutive for the stability and the welfare of a society. From this perspective, Fukuyama is correct when he observes 'a nation's well-being, as well as its ability to compete, is conditioned by a single, pervasive cultural characteristic: the level of trust inherent in the society' (1995, p. 7). This importance of trust can quite easily serve as the basis for evaluating trust as desirable, even as morally desirable, because trust contributes to the level of well-being in a society. This last aspect, however, needs a justification that is not present in the empirical observations. The empirical claim that trust contributes to a well-operating society is not yet a sufficient argument to evaluate trust as morally desirable. To substantiate such a claim, a view on the good society is needed. At this point we run into another problem.

If trust is related to an idea of the good society or to the idea of maximising a certain moral or non-moral good, it is only a small step to argue that trusting is morally desirable because of its contribution to that ideal society or good. Then trust is no more than 'handmaiden' to the ideal society or a function to achieve a certain good. Such an evaluation of trust can have far-reaching and undesirable consequences. For instance, the African system of Apartheid was based on cooperation and trust among whites; it even resulted in an increase of overall welfare for those who trusted each other. Consequently, trust was highly valuable for that society, and it can be conceived in terms of social capital. However, this trust-based social order resulted in a systematic disrespect of the moral value of the black part of the population. In that society, less trust might have resulted in a more critical stance towards the society's attitude to the suppressed. Such cases are not just rare exceptions. Also in many other situations trust can disguise morally deviant situations (cf. Baier, 1991). Another example is that of child abuse. Even if one does not have any explicit indication for child abuse, it is better to address parents critically if they bring in an injured child to the hospital, rather than just to trust them and miss those situations in which child abuse is at stake. It seems appropriate, then, to be careful about general evaluations of trust in terms of its contribution to a society (cf. Lahno, 2002a, p. 398).

This critical position is the result of my emphasis on trustworthiness. We should not trust because of the contribution trust makes to the society or to the performance of a specific community. We should only trust if the trusted agent is trustworthy. To encourage adopting a trusting attitude can be justified if and only if the trusted agent is trustworthy. The question should not be how to increase trust, but how to increase the level of trustworthiness, which makes trust an intelligible attitude to adopt. This

---

[5] 'Social capital' refers to 'features of social organisation such as networks, norms, and social trust that facilitate coordination and cooperation for mutual benefit' (Putnam, 1995, p. 67).

automatically illustrates why the question of trustworthiness cannot be reduced to the question of 'How can I be trusted?' There are many ways to make others believe that you are trustworthy and to help you being trusted (cf. Pettit, 1995).6 For instance, institutional agents can present themselves as trustworthy by launching information campaigns or commercials. However, trusting can only be morally valuable if the trustworthiness of the trusted agent can be substantiated in terms of competence and motivation. The result of the trusting relationship is only of a second-order value. Questions of trust have a pragmatic element, but I do not take this element as the core of the problem.

1.6.3 The position of the individual truster and moral pluralism

The presence of or search for shared values is often presented as crucial with respect to trust. Fukuyama and Etzioni are two among many who emphasise this element. This accent on the importance of shared values is legitimate.

Nevertheless, I do not share the communitarian ideas that underlie their attention to shared values. Let me take the account of Etzioni as an example. Etzioni starts from the recognition of people as human beings who ought to be treated as ends in themselves. According to the author this is a characteristic of a good society. Etzioni argues, 'a good society is one in which people treat one another as ends in themselves and not merely as instruments, as whole persons rather than as fragments; as members of a community, bonded by ties of affection and commitment, rather than only as employees, traders, consumers or even as fellow citizens' (2001b, p. 5). Up to this point I share his view. However, he stresses the importance of the embeddedness of the individual in a community to such an extent that one may wonder what the position of the individual in a society is. According to Etzioni, 'a good society fosters a set of core virtues that defines that which it considers good. The good society is not a neutral one that leaves it up to its members to decide on their own whether or not they wish to pollute the environment, abandon their children, abuse their spouse, drink and drive and so on' (2001a, p. xv). This raises the question of who is the society and who are its members? It almost sounds as if there were 'a society' independent of its members, which could determine the core values that ought to be observed for its benefit. If this really would be the case one may question the value of the acknowledgment of the individual as someone who is to be treated as an end in himself. Etzioni's view is of course somewhat more nuanced. He would not argue that norms are imposed. Nonetheless, the question about the relation between the interests and beliefs of the individual and those of the

---

[6] From this perspective, it is remarkable that Potter's virtue account of trustworthiness is entitled *How can I be trusted?* The main question of the work is less instrumental and aims to address the question of 'How can I be trustworthy?' Potter, N.N., 2002, *How can I be trusted? A virtue theory of trustworthiness,* Lanham etc: Rowman & Little Publishers.

community remains. The justification of moral values in terms of their contribution to a stable society and the consequences of that view for the position of the individual lead me to conclude that this is not a desirable way of establishing shared moral values. Shared morality is highly relevant in trusting relationships. The acknowledgement of this importance, however, does not require the strong emphasis on the community at the cost of the status of the individual.

Moreover, the focus on shared values and the community does not take the problem of moral pluralism seriously enough. Western societies harbour a striking plurality of moral views. Although we have tools and ways to address this plurality, e.g. by improving communication, increasing transparency or enhancing the level of reflection, a real problem remains. Even if we are transparent about our moral viewpoints, reason well and are open to further reflection, there remains a pluralism of inherently incompatible 'good and important' values and principles (Benjamin, 2001, p. 27). This cannot easily be reduced by further deliberation. Trust and trustworthiness can be complicated by this moral pluralism. On the one hand it is complicated by the conflicts of moral views. On the other hand, it is complicated by the doubts about whether a good reason for one person is a good reason for all of us. Consequently, it is questionable whether a community perspective can provide an overall good or overarching principle which guides the truster and trustee in situations of moral conflict.

In Chapter 5 my argument is that, given this moral pluralism, being trustworthy is not a matter of searching for shared values only, but should also include a way of balancing the acknowledgement of the truster's moral values and principles and those of the trustee. Striving for such a balance does not need the strong emphasis on the role of a community. A more modest basis, such as Rawls's liberal concept of 'overlapping consensus' (1972; 1993) can be sufficient. Rawls accepts the evident moral differences, but he claims that it is possible to come to practical agreements. This does not imply that all share each other's values, but that there is an overlap in the discourse which makes it possible to come to an agreement (Rawls, 1999; Brom 1998).

1.7 Summary and outline

In this chapter I have made five claims that will be elaborated and substantiated in the following chapters.

First, I treat the public attention to trust as a signal of the problem that people have to rely on others, but often do not know whom to trust. I call this the 'problem of trust'.

Second, I claim that this problem of trust can best be approached as an indication of problems with respect to trustworthiness. I argued that we have conceptual and moral reasons to change the focus from trust to trustworthiness. Consequently, the main

question is not how the individual can be changed so that he will trust, but what conditions the trustee has to fulfil in order to be worthy of trust.

A third claim is on the conceptual level. Even though situations in which trust and risk are relevant often overlap, there is a fundamental difference between trusting and taking risks. I treat risk calculation and trusting as two complementary, yet different mechanisms to deal with uncertainty.

A fourth claim is that the problem of trust can be interpreted as a signal of questions on a moral level. On the one hand, a problem of trust can be a signal of a lack of respect for the truster's position as an autonomous agent. Trusting or withdrawing trust illustrates how a person judges his own autonomy and that of the other agent. On the other hand, the problem of trust can be a signal of a disagreement about what one can expect of each other. Entrusting and responding to trust reflect ideas on what one considers as legitimate moral expectations.

The fifth claim combines the second and the fourth claim. The moral dimension has direct consequences for what counts as trustworthy behaviour. I will argue that any attempt to be trustworthy should start in the respect for the autonomy of the truster. Trustworthiness has to start in the recognition of him as a moral agent and as a moral equal. Furthermore, this claim comes with a demand to take the moral dimension of what is entrusted seriously. The incorporation of this moral dimension by the trustee is a necessary condition for being trustworthy.

In the course of this thesis these five claims will be elaborated and substantiated. For reasons stated above, I focus the discussion on questions of trust in the agricultural and food sector.

Chapter 2 introduces the agro-food sector. This chapter shows that individuals cannot but rely on other agents in order to operate in today's society. I present the agro-food sector as an incontrovertible example of this. Developments that characterise the agro-food sector, such as the process of globalisation, the increase of the use of technology in food production and the recent food-related scandals, result in what I have called the problem of trust. The developments have stressed the individual's need to trust. At the same time, they have illustrated that these developments complicate trust, since it results in problems concerning trustworthiness. It has become more difficult to identify the agents in whom one is required to place trust and has raised all kinds of practical problems with respect to the assessment of the trustworthiness of identified agents. Analysis of the background of the current attention to public trust in the agro-food sector shows that the problems of trust can best be addressed as problems at the level of trustworthiness.

From this specific case, Chapter 3 turns to the conceptual analysis of trust. In spite of my claim that the problem of trust can best be addressed as a problem of trustworthiness, we need conceptual clarity on what we mean by trust. The concept is often used in vernacular and academic debates, but it is still not a clear one. In fact, this broadness is

an inherent characteristic of the concept. However, the chapter marks out trust by delineating it from concepts such as faith, hope and reliance. I will elaborate on the differences between trusting and taking risks and show that trust is an attitude, rather than an act or a belief. Finally, I define trust as an attitude which enables an agent to deal with uncertainty or lack of control, and show why a start at the side of trustworthiness is more promising in addressing the problem of trust.

Chapter 4 starts with a refinement of the above definition. I argue that, in contrast with reliance, trust is a way of dealing with a specific kind of uncertainty. It is an attitude that enables dealing with autonomous agents in situations of uncertainty, rather than with uncertainty as such. I propose freedom, agency and a participant attitude as constitutive for trust. This implies certain constraints to the tools and methods to build and maintain trust. If trust were about dealing with uncertainty as such, then power, coercion, or controlling behaviour would be relevant methods, which would help to establish trust. I use the Hobbesian contract to show why coercion, power and sanctions may be relevant to establishing public order, yet not to establishing public trust.

Chapter 5 builds on the conclusion from the former chapter that genuine trust starts from the recognition of the other as an autonomous agent. In this chapter, I explicitly focus on the consequences of that recognition for trustworthiness. I argue that trustworthiness involves a duty to show due respect for one's own autonomy and that of the truster. To ignore this duty has far-reaching consequences. It either implies disrespect of the truster as a moral agent and as a moral equal, which will not easily motivate that person to come to trust, or it is a signal that the trustee does not take his own autonomy seriously, which raises questions with respect to his competence and motivation. This duty results in some clear constraints on what counts as trustworthy behaviour. In general, it implies that a truster ought not to be considered just as a vulnerable person, but that she should be treated as person who has the capacity to choose her goals and values personally. At the same time, to be trustworthy the trustee cannot ignore her own autonomy. In cases of moral conflict this complicates the duty of respect for autonomy.

This problem is intensified the more by the fact that we are confronted with a plurality of moral views that is irreducible, and for which we lack an easy priority rule which tells the trustee whose moral position is overriding. To address the problem of trustworthiness that arises if the moral expectations of the trustee and the truster conflict, I propose a balance between accommodation to the truster's expectation and preserving the trustee's integrity. This balance includes two principles of accommodation and some constraint on these principles from the perspective of integrity.

Chapter 6 focuses on the consequences of the conclusions of the former chapter on an institutional level. Trust as it has been presented in the second chapter is mostly trust in institutions, such as governmental bodies, non-governmental organisations and market parties. In order to buy and consume food consumers increasingly depend on the competence and motivation of numerous, mostly anonymous others in the food chain.

Consequently, the trustee is no longer one person, but is a collective of agents, which operates towards a common aim and according to internal, shared rules. I can trust specific persons as representatives of the institution. At the same time, we also trust institutions because of the behaviour of their representatives. The difference with personal trust in this case is that I trust the institution because of the behaviour of the individual agent, yet my trust is not in the individual himself.

This change from personal to institutional trust raises different questions with regard to the analysis of trust and trustworthiness that has been presented in the former chapters. Chapter 6 argues that, in spite of the differences, we can speak about 'trust' on an institutional level and that institutions can be trustworthy rather than reliable only. On the one hand, institutions can meet the formal criteria for trust that I defined in Chapter 4. In addition, the discussion of collective responsibilities shows that a truster can have expectations of an institution that cannot be reduced to expectations towards individual agents. On the other hand, an institution can be trustworthy. It can promote trustworthy behaviour of those who operate within the organisation. Moreover, the institution itself can be competent and adequately motivated. The latter aspect is elaborated in some detail by the discussion of the importance of institutional integrity.

Finally, I deal with the question whether institutions can meet the moral condition of showing due respect for the truster as an autonomous agent. I argue that the condition as such is not problematic, but that the obligating character of this respect for autonomy and the related protection of the truster cannot be grounded in the autonomy of the institution. An additional argument is needed to safeguard respect for the truster. At this point, I present the notion of obligation, which follows from the implicit invitation to trust. Institutions regularly create, by communication or by their organisational structure, expectations about their trustworthiness. This invitation is close to a promise, which results in an obligation to show respect for the truster and to take his vulnerability seriously. The validity of this obligation, which directly results in action-guiding principles, is independent of the autonomy of the institution.

The final chapter of this study presents a case of trust in novel food products at the interface between food and health. As a result of different developments, the food sector and health sector become more and more intertwined. This creates many prospects, but it also raises questions. One of them relates to the implications for public trust in food and health. I show how the different steps taken in the thesis are helpful to address this practical situation.

CHAPTER 2

PROBLEMS OF TRUST AS PROBLEMS OF TRUSTWORTHINESS:

THE AGRO-FOOD SECTOR AS AN EXAMPLE

We may have no choice but to continue
to rely on the local shop for food,
even after some of the food on its shelves
has been found to have been poisoned with intent.
Baier, 1994a [1986], p. 98

2.1 The agro-food sector as focus and case

This chapter is a first step in the substantiation of my claim that the current attention for the problematic status of trust in institutions can be addressed fruitfully from the perspective of trustworthiness. I do so by focusing my analysis on the problems of trust in a specific sector: the agricultural and food sector. For several reasons this is an interesting case. This sector has been characterised by a number of highly publicised food safety problems that evidently have affected public trust in both governmental and market institutions. Moreover, several developments in the sector, such as globalisation and the expanding use of technology have resulted in an increased need for individuals to rely on institutional agents in the sector, but also make traditional ways of assessing someone's trustworthiness, like face-to-face contact, more difficult. The combination of these aspects already is an indication of why problems of trust have occurred recently: there is an increased need to trust, yet, at the same time there is uncertainty about the trustworthiness of the trusted agents.

The relevance of an analysis of trust in political and market institutions in the agro-food sector is not limited to this sector only. The context, the dimension and the nature of the problems of trust in the food sector are similar, or at least have close links to problems of trust in other domains and concerning many other themes, such as security, social welfare, and health care. For this reason the agro-food sector is an interesting case that illustrates that problems of trust are most effectively addressed as problems of trustworthiness. I first sketch the importance of trust in the agro-food sector, especially

for consuming food (Section 2.2). Next, I elaborate on the context of the problems of trust. I analyse why the implicit attitude of trust has become explicit. Three general characteristics of the agricultural and food sector provide us with the structure to give an answer to this question (Sections 2.3–2.5). Thirdly, I discuss the dimension of the problem. We are said to be confronted with a crisis of trust. I argue that this is not the case. Nevertheless, problems of trust have to be taken seriously (Section 2.6). Finally, the focus is on the nature of the problem. Along the lines of a discussion of the current European policy measures that address the problem of consumer trust, the last part of this chapter aims to substantiate my claim that problems of trust had best be addressed from the perspective of trustworthiness (Sections 2.7–2.9).

2.2 The specific context of food: The need to trust

Consuming food implies trust. To be aware of the evident importance of trust, we only haveto imagine what would happen if we really were to lack all trust in food. We emter a supermarket and are confronted with dozens of articles about which we just do not know enough to assess their safety, their nutritional value, or to what extent they fit within our lifestyle. All products, from fresh spinach to frozen pizza will raise questions on safety that we, as consumers, are unable to assess and answer. Steaks or chicken wings will confront us with our ignorance regarding the levels of animal welfare and health in the production system. Moreover, as consumers we are hardly able to determine whether the content of our shopping basket meets the standards of the recommended dose of essential vitamins, proteins, and fibres. In all these cases we have to rely on many others, since we lack the knowledge, the expertise, physical abilities and time to assess all the risks and benefits of all products. This need to rely is not limited to those from whom we directly buy products, but also includes reliance on many others, the entire range from family, friends, and acquaintances to trust in complete strangers and in abstract systems like product boards and the government. In this context a real absence of trust would not just make food consumption problematic; it even would greatly reduce the possibilities for consuming food. In the consumption of every product, we run a risk regarding our health or that of our kin, or it may infringe on our personal values. A complete absence of an attitude of trust would paralyse food consumption. In order to be able to act in spite of this dependent position an attitude of trust is indispensable, for trust is a way to act in spite of uncertainty and in spite of a lack of personal control. With regard to food, this means that despite all uncertainty and all ignorance food consumption is still possible.

The consequences of a lack of trust illustrate the importance of this attitude. However, reference to the importance of trust is not enough to explain the current attention for trust in this sector. The need to trust does not hold for food alone, nor is it a new phenomenon. Trust is equally important in other contexts, e.g. when we drive a car

or go on holiday. Nor is it new; trust has always been a precondition for food production and consumption. Those who lacked that trust had to find other ways to be able to consume their food, like Marcus Antonius, about whom it is said that he never dined with Cleopatra without his food taster, because of Cleopatra's dubious poisoner fame.

From this perspective, the need to trust is clear, but the attention for consumer trust in food in recent years remains remarkable. Different organisations on many levels have stated that there is an evident need for rebuilding and maintaining trust (e.g. FAO, 2003; CEC, 2000; FSA, 2001; Kettlitz, 2001; VWA, 2002; 2003). Why has it become such an issue in all kinds of public debates and policy documents? Why has the implicit notion of trust turned into an issue on the public agenda and is even related to the establishment of food authorities all over Europe? Three general characteristics of the agricultural and food sector provide the structure to answer this question.

2.3 The context of the increased attention for public trust

'The European Union needs to re-establish public confidence in its food supply, its food science, its food law and its food controls.' This is one of the main introductory statements in the European Commission's *White paper on food safety* (CEC, 2000, p. 7). It is evident that the importance of trust in food has been recognised on the European level, but does not yet explain why consumer trust has been put on the public agenda. To answer this question, it is helpful to distinguish three characteristics of the agricultural and food sector (e.g. Brom, 2002, Meijboom et al. 2006, RLG, 1998). These three have clear parallels with similar developments in society as a whole (e.g. Giddens 1990; 1991). The first is the development of the growing distance, in both time and space, between production and consumption. The second characteristic is the increasing importance of technology for the food sector. Finally, the food sector has strongly been associated with different food-related scandals in recent decades (Section 2.4).

2.3.1 The distance between production and consumption: A global sector

The agro-food sector is characterised by a growing distance, in both time and space, between production and consumption. The gap between the system of producing and processing food and the consumer is widening (e.g. Korthals, 2001, p. 208; FAO, 2003, p. 8). A dinner in the Netherlands may include beans from Egypt, potatoes from Malta, a steak from Argentina, and red wine from South Africa. This is what Giddens (1990, pp. 63-65) describes as the 'stretching' process in which different social contexts or regions become networked across the earth's surface as a whole. Yet, the dinner example does not yet explain the current preoccupation with trust in Europe. A lot of well-known fruits, vegetables and other food products were initially imported from countries from all

over the world. Potatoes and tomatoes are only two among many others. Looking for easier ways to bring spices to Europe was one of the main rationales for many of the discovery voyages in the 15th Century (Busch, 1997, p. 2). As Lang rightly notes, 'there have been many previous phases of global transfer of foods, habits, and techniques' (1999, p. 170).

Global transfer has radically changed in pace and scale however, and resulted in a widening gap between farm and fork. These radical changes mark the difference between international interaction and trade in food on the one hand and globalisation of food on the other. In the first case, there may be important and strong contacts between different social contexts, yet the focus is still local, regional or national. In contrast, globalisation is the 'intensification of worldwide social relations which links distant localities in such a way that local happenings are shaped by events occurring many miles away and vice versa' (Giddens 1990, p. 64).

The global character of the food sector is caused by (a) the increased scale of *production*, (b) the increased scale and pace of the *transfer* of products, and (c) the increased *number of actors* within the production system.

First, the increased scale of production strongly contributes to the global character of the food sector. In order to meet the need for food in the world, large multinational companies have been formed that operate in large parts of the world and transfer products around the world daily. Not only has the volume of different products increased, but also the number of different kinds of products. While fifty years ago one had the choice between one or two brands of soft drink, a consumer in the Western world now has the choice between many different kinds, brands and tastes. This requires a global production and automatically shows the increase of scale and pace of the transfer of products. Global transfer is no longer a special way of distribution in order to import exotic products, but part and parcel of the food sector. The more the production is global the more transfer is an essential precondition for production. Just as part of the regular food processing, transportation over long distances is common. Only two among many examples are the Dutch pigs that are transported to Italy to be prepared as 'Parma ham', and the Dutch prawns that traditionally were peeled by the fisherwomen at home, but in the 1990s were transferred to Eastern Europe and currently are peeled in Morocco.

Second, the increasing scale of production also influences the pace of the transfer. Whether you produce food in a local, regional or global setting, food is needed daily. Therefore, global production needs global distribution with a high rate of circulation, day after day.

Third, the increased number of actors within a production system is important. The line between producer and consumer is long, not only literally but also with respect to the number of actors who are involved in the food chain. One product, e.g. a bottle of beer, is the result of activities of many different agents in many different countries, such as farmers in Scotland and the Champagne region in France who grow wheat or barley,

the researchers in the laboratories that monitor the quality of the yeast strain, a malting company in the Netherlands, a water supplier in Germany, and the staff of the brewery, not to mention all who are involved in the bottling process and the distribution and retail systems. This example can be duplicated by many others. It is not an exception but illustrates that food production has become a matter of many agents, in many institutions and countries.

These circumstances of globalisation imply that in everyday actions such as buying and consuming food the individual's need to rely on others is vastly increased. As a result of the long and complex food chains no individual has the ability to fully control the production process leading to the final food product as consumed. Especially the consumer who is at the end of the production chain often have little or no idea of all the agents who are involved in the process preceding the moment one buys and consumes a product. Zwart describes it as an 'actual loss of intimacy with food production' (2000, p. 125). This illustrates the importance of trust. When buying and consuming food products, consumers cannot but rely on many others. The consumers' 'loss of intimacy' is not the result of indifference with respect to the production process, but of the inability to know and check all actors in the production chain. The need to rely is not only caused by the increased number of involved agents. Globalisation also confronts individuals with new and more complex situations that are difficult to assess by an individual consumer. This surfaces, for instance, in the discussion on the environmental and social consequences of the global trade in terms of the so-called 'food miles'.[7] Moreover, the growing gap between production and consumption introduces questions from miles away to our local dish and implies answers from miles away to local questions. For instance, the poultry consumption of a consumer in Germany is intimately linked with a whole system of global trade and transport that at the same time links that person with the local safety and welfare standards of humans and animals in Asian countries such as Thailand. The discussions in the context of the Word Trade Organisation show the difficulty of addressing themes like safety, working conditions, and animal welfare on a global level (Vedder, 2003; Heeger & Brom 2003; Meijboom & Brom, 2003; Staman & Brom, 2000).

Partly the increased need of trust has been addressed by strong state regulation. Different measures on the national, European, and global level have been taken to regulate issues such as food safety, e.g. the establishment of the Codex Alimentarius Commission that aims to protect the health of the consumers, to ensure fair trade practices in the food trade, and to promote coordination of all food standards work.[8] On

---

[7] The term entered the debate in 1994 when the SAFE Alliance (Currently: Sustain) launched the 'Food Miles Report'. Paxton A, 1994, *The Food Miles Report: The dangers of long distance transport of food*, SAFE Alliance.

[8] The Codex was created in 1963 by FAO and WHO, http://www.codexalimentarius.net/web/index_en.jsp (accessed April 2007)

the other hand, the food chain is still characterised by self-regulation in which commercial parties such as retailers play an important role. Therefore, trust in a whole range of agents is still necessary to operate in the global food market. This illustrates that one has to be almost omnipresent and omniscient to control one's own food. Since this is not given to humans, one cannot but rely on several agents.

This situation does not only underscore the need of trust however. It also complicates it, because it complicates the assessment of trustworthiness. Traditionally, trust is conceptualized in personal relationships. In these contexts, we often have rather direct access to information on the competence and motivation of the trustee, for example, your friend or the local greengrocer. However, in a global market short and personal relations are rare. Instead of assessing the trustworthiness of only a local farmer, one now has to trust a range of agents from numerous countries, e.g. scientists from Japan, or a big multinational with its headquarters in another part of the world. Consequently, it is more difficult to assess the trustee's competence and motivation. Thus, one knows that one has to rely on others, but often is uncertain whether the other can be trusted.

In short, the increased complexity that is the result of the widening gap between production and consumption illustrates the need to trust, but also makes trust more difficult to maintain and achieve, because it complicates the assessment of trustworthiness.

2.3.2 Food and technology: Risks, experts and information

A second characteristic colouring the food sector is the increasing importance of technology. Since the 19th century, several technologies have been introduced to enable people to transport food over longer distances and improve the storage life of products. These were technologies like canning, pasteurising and freezing. Yet, the current use of technology is no longer restricted to improving the 'use before' date. We are all used to food additives that enhance the colour or flavour of a product, that change the nutritional composition or even make for health claims that are completely new to that food product, e.g. enhancing bread with folic acid. Moreover, technology is used in the production process in order to shorten, intensify and simplify the production process. All this leads to a kind of 'artificialisation' of food as an end product (RLG, 1998) that contributes to the above-sketched situation in which personal ties with those involved in production and processing become rare.

The use of technology in the food sector has a double effect on trust. On the one hand, it provides personal control over situations in which one had to rely on or had to trust others. On the other hand, modern technology intensifies the dependency of individual agents because of the complicated nature of the topics involved in dealing with issues such as food safety, quality and health. These call for abilities that most of us do not have. Consequently, we cannot but rely on others. This double effect of

technology on trust can be recognised with respect to the influence (a) on risks and uncertainties, (b) on knowledge and information and (c) on predictable patterns on which one can anticipate.

First, the dual effect of technology can be recognised at the level of risk and uncertainties. On the one hand, it contributes to the reduction of uncertainties in the food sector. Technologies are often introduced in order to address uncertainties and risks. With the help of technologies, it is possible to get grip on a situation. For instance, in the past one could only hope that the water quality in a river was good enough to be used for the preparation of meals. Nowadays, technology makes it possible to monitor the quality and to intervene if the public health is endangered. Consequently, we are in a less dependent and less vulnerable position. Nevertheless, the use of technology also raises new uncertainties and risks. When we define risks in terms of chance and hazard,9 technology can affect both levels. Due to the effects of technology on the structure of the food chain, risks surface because the chance of defects and mistakes can increase exponentially. The introduction of technology often contributes to longer and more complex production chains that are interrelated with many other systems, such as transportation and storage. Due to this complexity, minor effects of a technology can have consequences with a major impact for the society. For instance, one ingredient that contains toxic residues may affect the safety and quality of hundreds of products all over the world. Furthermore, uncertainties arise along with the introduction of new technologies. This is the case when the risk of hazard is unclear or when the precise hazard is not yet well defined. The discussions on the threat of antibiotic resistance because of the use of antibiotics in animal husbandry or the adverse effects of the use of growth hormones in meat production are clear examples. There both the probability that something goes wrong and the precise hazard are at issue. Such discussions show that the use of technology entails risks and uncertainties which are often difficult to assess, especially because of complicating factors like unknown carry-over effects and the possible long-term effects. Therefore, assessing issues like safety, quality and health is a task that requires powers that most of us do not have. Only few have the expertise and can assess and evaluate these problems. All others cannot but rely on these experts. This shows a shift in focus. The problem is not merely the dimension and acceptability of the risk at stake, but also one of the reliability of the experts. The experts make a calculation, not the consumer. Consequently, problems do not only occur at the level of the risk itself, but also with respect to the trustworthiness of the experts on whom one has to rely. For instance, even when the assessed risk might be low one still can be extremely uncertain about the other's trustworthiness.

Second, technology has a double effect on trust because it affects two other relevant aspects: the validity and distribution of knowledge and information. Since new

---

[9] The relation between risk and hazard is discussed more extensively in Section 2.8 below.

technologies and their results are presented with a high rate of circulation both the validity and the distribution of information and knowledge have been fundamentally changed. First, the validity of both tends to change, and at a very high pace at that. The soundness of knowledge claims always has been subject of discussion. However, because of the use of technologies, such as information technologies, genomics or nanotechnology, the validity of knowledge and information is persistently under pressure. What is considered safe today can be questioned tomorrow and what is argued to be environmentally friendly by several authoritative experts can be refuted by a same number of other distinguished experts. The debate on the safety of genetically modified food shows the striking tensions between the opinions of experts on the risks of modern biotechnology for humans, animals and the environment. 'Just' relying on the expert opinion is often not possible. Hence, even tough more knowledge and information is available it is difficult for an individual to decide whom to trust. Furthermore, the distribution of information has dramatically changed. As a result of new information technologies and tools, such as the internet, and on-line communities such as 'Hyves', information spreads over the world at such a high speed that information can easily be shared and exchanged. As it is easier to obtain information, individuals are better informed about many themes such as health claims, production methods and dietary advice. This knowledge can contribute to trust, because one is in a better position to asses a situation without the need to rely on others. However, in combination with the question of the validity of the information one still is confronted with the question of whether the (provider of the) information is trustworthy.

Third, the double role of technology with respect to trust can be highlighted by way of its effect on predictable patterns. Regularly, technology results in procedures that make a situation more predictable. For instance, the use of technology can standardise a production method so that one can anticipate that the quality of the product is similar at any time and any place, provided that a certain technology is applied. However, the introduction of new technologies also can thwart existing predictability and familiarity in the food sector. When a technology is introduced, there is often no predictability that can serve as a basis for trust. The example of the introduction of food products with a health claim is a case in point. Although the relation between food and health is not new, lowering one's blood cholesterol with the help of a dairy product is new. Normally one drinks milk for several reasons, but not for lowering one's elevated level of LDL cholesterol, just as one usually does not take liquid drugs when one is thirsty. Both for drugs and for dairy products there are rather clear patterns and traditions that provide a certain predictability explicating what one can expect regarding issues of safety and justice. These patterns and routines help to trust another with regard to both food and pharmaceutical products. However, since a food product with a real health claim can be categorised in both groups, there is not one unambiguous pattern available upon which one can base trust. Trust is either based on patterns from the food sector, although it has

a health claim, or it is based on the patterns of the pharmaceutical domain, although it is a food product. Thus, the introduction of such a diary product complicates the possibilities to trust.[10] It seems no longer possible to rely on existing roles and patterns that serve as a basis for trust. Hence, as a consumer, one has to find other reasons for trust, yet it is often not clear what can function as such a basis.

To summarise: the increasing use of technology does not necessarily lead to problems of trust. Nevertheless, it intensifies the need to rely on others, especially well-informed experts, and complicates trust, because of its effects on the validity and distribution of knowledge and information and on predictable patterns.

2.4. Concerns and scandals: Problems of trustworthiness

A third influential characteristic of the food sector is the number of food-related scandals and food safety problems. 'Food scares are never far from the news', as Tenant (1997) claims, but the number of cases of foodborne diseases and food-related scandals in Europe during the last decade(s) is striking. Although cases of foodborne illness and food safety related problems occur daily, some have attracted a lot of media attention, raised serious consumer concern and affected trust in several agents in the food chain. Roughly, the food-related concerns and affairs can be classified in three categories: (a) concerns related to foodborne diseases that are the result of microbiological hazards, (b) concerns and scandals related to chemical hazards of food consumption, and (c) concerns and scares with respect to zoonotic diseases.

2.4.1 Concerns on microbiological hazards: New dangers, new cases of trust

There have been serious concerns related to the microbiological hazards of food consumption. Especially adverse effects of microorganisms such as *Escherichia coli* O157:H7, *Campylobacter jejuni* and *Salmonella typhimurium* DT104 have caused serious consumer concern. This is not without reason as this category of foodborne diseases takes many deaths each year. According to the WHO, 'approximately 1.8 million children in developing countries (excluding China) died from diarrhoeal disease in 1998, caused by microbiological agents, mostly originating from food and water' (2002, p. 10). Nevertheless, in Europe the most serious concerns were not raised because of these deaths, but were caused by the fact that some microbiological hazards were unknown and others turn out to be very dangerous. The *E.coli* O157:H7 bacteria are an example of the first case. It was first identified as pathogen associated with food in

---

[10] This holds not only for this specific case. It may also affect our expectations with regard to other food and pharmaceutical products, since we are no longer sure whether our 'normal' expectations still apply to those products.

1979.[11] Prior to that many deaths and cases of disease could not be diagnosed as related to food consumption. Other bacteria were known as pathogen, yet appeared to be more hazardous. For instance, *Salmonella typhimurium* DT104 had developed resistance to five commonly prescribed antibiotics and had a rapid spread during the 1990s (WHO, 2002, p. 10). The issue is complicated further by the question of the best way to address these food hazards. It has been argued that the tendency to focus on analysing and controlling individual diseases caused by micro-organisms ignores the fact that these organisms are part of an ecosystem. Removing one pathogen can have the effect that 'new pathogens will simply substitute for those which are deleted. Hence, dealing with microbiological hazards is not just a problem of particular microbes, but 'one of niches created by a particular way of organizing the agri-food system' (Waltner-Toews & Lang, 2000, pp. 122-123).

   The detection and recognition of these new dangers raises new questions and concerns with respect to the safety of food, and stress the public's need to rely on expert views since most if not all consumers lack the competence to assess these microbiological hazards. It also indicates the complexity and uncertainties these scientific experts are faced with. This entails questions of trustworthiness, especially concerning the competence of the experts, but also in relation to their motivation to take consumer concern seriously.

2.4.2 Chemical hazards: New responsibilities and the case of the dioxin scandal

'Public awareness about chemicals in food is relatively high' (WHO, 2002, p. 11). At this level three different types of concerns have been raised. First, there are concerns related to the uncertainty and unclarity about the precise hazard of chemical contaminants in foods, such as natural mycotoxins, and the impact of environmental contaminants, such as lead and dioxins. Like in the case of the microbiological hazards, some dangers were relatively unknown and appeared to cause foodborne illnesses only after new research, or turned out to be especially dangerous for specific subpopulations such as children, pregnant women and the elderly. Moreover, new assessments revealed that even a low-level exposure to multiple chemicals has adverse effects that one would not expect if one focuses on the single chemicals only. This also holds for the cumulative effect of the intake of one chemical contaminant over a long period (WHO, 2002, p. 11). This better understanding of the hazards of chemical contaminants in foods has been crucial for the improvement of food safety, yet it also implied the explication of food dangers and risks most consumers were completely unaware of. Consequently, this again

---

[11] It turned out to be present in ground beef, unpasteurized apple cider, milk, lettuce, alfalfa and other sprouts, and drinking water in several countries (WHO, 2002, p. 10).

stresses the consumers' need to rely on expert views and raises questions regarding the trustworthiness of agents in the food chain.

With respect to chemical hazards the issue of trustworthiness is even more profound than in the case of microbiological hazards, because some chemicals like food additives, but also pesticides and veterinary drug residues are deliberately used in the food chain. The addition of chemicals makes it possible to help preserve food, improve nutritional value, enhance quality and increase the food supply (cf. Scheuplein, 1997, p. 422). This introduces new responsibilities for those who use these chemicals in the food chain, but also for the consumer. Consequently, two other types of concern can be recognised. On the one hand, 'consumers continue to express concern about the risks to health due to the deliberate addition of chemicals to food' (WHO, 2002, p. 11) since the introduction of chemicals is considered as unnatural and undesirable. On the other hand, concerns and scandals surfaced because of malpractice and other untrustworthy behaviour.

With respect to the former concerns, the problem is not the uncertainty about the precise probability and consequence of the presence of a chemical, but the belief that such chemicals do not belong in the food chain. It faces parents for example with the question whether such products are appropriate for their children and whether they are entitled to take such risk on their children's behalf (Tennant, 1997, p. 383). This can be the reason to switch diets and to choose vegetarian or organic products only.[12] If such choices are not possible serious problems of trust surface, since consumers then have to rely on the safety of products that they do not consider as safe or acceptable.

The latter concerns are related to malpractice or negligence and receive the most public attention. A clear case is the dioxin scandal in Belgium in 1999.[13] A processing company sold fat for use in animal feed from a dioxin-polluted storage tank of fats and oils. Between 15 and 31 January 1999, oil from discarded transformers was admixed to fat that was delivered to ten animal-feed producers. As a result of the global character of the agro-food sector and the related increase of the pace and scale of transportation, the 500 tonnes of animal feed contaminated with approximately 50 kg of polychlorinated biphenyls (PCBs) and almost 1 g of dioxins[14] found their way to an estimated 1,400 Belgian poultry and pig farms[15] over just a two-week period. Food produced from animals given this contaminated feed found its way to every continent within only weeks

---

[12] Such specialized diets do not guarantee that there is no intake of additional chemicals from the diet. A vegetarian diet, for instance, will have 'lower levels of intake of substances associated with animal products, such as veterinary drug residues, but could have higher intake of substances associated with plants, such as pesticide residues' (Tennant, 1997, p. 383).

[13] This example is not unique. Several cases of contamination with PCBs have occurred in the past in Japan (1968), Taiwan (1979) and Michigan (1983) (Covaci & Grob, 2002, p. 51). Another example of contamination with a chemical that raised a serious food scandal was the addition of toxic ethylene glycol to wine in Austria in 1985, the so-called *Glycolwein-Skandal*.

[14] Polychlorinated dibenzodioxins (PCDDs) and dibenzofurans (PCDFs)

[15] To a lesser extent also to farms in the Netherlands, France, and Germany.

(Van Larebeke et al. 2001; Covaci & Grob, 2002; WHO 2002). As a result, consumer trust was harmed. Some agents within the food chain appeared to be untrustworthy, not because they were incompetent, but because they deliberately acted against the law and did not respond to the trust of many partners in the chain.

2.4.3 Animal diseases and zoonotic hazards: More than concerns on health

Thirdly, the emergence of food-related zoonotic diseases has profoundly contributed to an atmosphere of concern and anxiety. The intensified way of animal husbandry resulted in an increase of food supply, but also led to 'the emergence of new zoonotic diseases, which affect humans' (WHO, 2002, p. 10). BSE[16] or 'mad cow disease' is probably one of the most striking examples of zoonotic diseases that have raised serious food-related scandals. Since 1986, BSE is known as a fatal neurodegenerative disease that affects cattle, but it was only in 1996 that the UK government announced that BSE was linked to a novel human disease that is fatal for humans: the variant Creutzfeldt-Jakob disease (vCJD). Moreover, this disease appeared to be almost certainly caused by consuming BSE-contaminated food. This had serious consequences for the consumer confidence in the safety of meat (cf. Millstone & Van Zwanenberg, 2002). This, however, was not the only effect at the level of trust. It also harmed public trust in the UK government. The governmental authorities first assured the public that no safety issues were involved in the consumption of beef and that all BSE-related health issues were fully under control. Most members of the public confidently relied on this message. However, when the government had to announce that beef consumption could have serious adverse health effects and that they had already known this for some time but had not conveyed the uncertainty they faced, trust was seriously harmed.[17]

In this context other outbreaks of animal diseases, such as classical swine fever in 1997-98, of foot-and-mouth disease in 2001 and avian influenza (bird flu) in 2003 and 2006, resulted in new concerns related to human health and animal welfare. Even though the threat of these animal diseases is not, or only marginally, linked with food consumption, such occurrences certainly do not help reduce the food safety concerns and the worries with respect to foodborne illnesses. Moreover, its marginality in relation to food did not migtigate the problems of trust related to the competence and motivation of the government.

---

[16] Bovine spongiform encephalopathy (BSE).
[17] Below I will elaborate on this example and focus on the role of the UK government in the BSE-vCJD case.

2.4.4 Scares and scandals as a problem of trustworthiness: The MPA-scandal

The effect of these problems and scandals on public trust in the government and several other agents in the food chain has been mentioned widely. The FAO, for instance, states that highly publicised food safety problems 'have given rise to a general state of distrust among consumers' (2003, p. 3). The UK Food Ethics Council observes that 'in the aftermath of BSE and Foot-and-Mouth Disease, and as the GM controversy rolls on, nowhere is trust lower than in food and agriculture.'[18] The food-related affairs, as an explanation of the problems of trust, are a signal of and complementary to the above-mentioned characteristics of a global and technology-driven food sector. All food-safety problems and scandals are a direct confrontation with the need of the individual to trust many agents in the food sector. The above analysis illustrates that some concerns have been raised because research has identified new hazards that face consumers with additional situations in which one has to rely on others in order to buy and consume food. Faced with this situation questions of trust automatically materialise, either because it is unclear whom one can trust with these 'new' uncertainties or because it is unclear to what extent a particular agent is competent and sufficiently motivated to deal with the issue. Moreover, the food scares are a direct confrontation with the fact that trust is mainly trust in a small group of experts. One often has no other option than to trust experts; consequently one is vulnerable. This especially holds for trust in food products, since food is something very valuable and intimate (cf. Caplan, 1997) and we do not have the option not to consume food. Consequently, when trust in expert authorities has eroded, it makes trust not just a bit more complicated, but can make the difference between an attitude of trust and a state of uncertainty or even distrust.

Most important however, the scandals stress the importance of trustworthiness. The majority of the affairs came about because trustees had proven not to be as trustworthy as they were thought to be, either because they were not competent enough or lacked the adequate motivation to respond to what was entrusted to them. The above-mentioned dioxin scandal and the BSE affair illustrate this.

The MPA affair in 2002 in Europe further confirms this and shows the impact of a global sector on the complexity of the issues at stake. In July 2002, Dutch authorities noted the presence of MPA hormone[19] when pigs were tested. It turned out that this was not due to the use of an illegal livestock growth promoter. A serious food contamination scandal was discovered. The hormone was traced back to waste water from an Irish pharmaceutical firm, which was sold on to a waste management firm and then shipped to Belgium where it was mixed into pig feed. By the time Dutch farmers bought it, it was contained in treacle or glucose syrup used for pig feed. Ultimately MPA was found in

---

[18] http://www.foodethicscouncil.org/projects/agrifood/introduction.htm accessed at 1 December 2005.
[19] MPA: synthetic progesterone medroxyl progesterone acetate.

products ranging from animal feed to soft drinks and may have been distributed to many other countries, including Germany, Italy, France, Spain, and possibly Luxembourg and the UK (BBC, 2002; Birchard, 2002; PDV, 2002).

This contamination resulted from lack of competence and a clear abuse of the safety regulations and laws, rather than as a result of the lack of a clear legal framework on food safety. Some agents within the sector were less competent than they were considered to be or even deliberately abused the confidence of others, i.e. they acted untrustworthily. For instance, the company that takes care of the export of the waste from the pharmaceutical company incorrectly classified it as 'green' rather than 'orange' waste,[20] which implies that certain safety checks are not compulsory and the waste could enter the food chain. Furthermore, there is a clear element of offence and crime. The Belgian company that processed the waste was not licensed to treat pharmaceutical waste, but nonetheless deliberately processed it. In addition, the carelessness of certain feed producers and farmers was mentioned because they did not buy their products from a certified company even though they participated in a quality assurance system (Rogers, 2002). The Dutch 'Product Board Animal Feed' correctly placed the problem at the level of the animal feed companies and farmers in the Netherlands: 'at the companies in question there was a lack, on the one hand, of sufficient insight into the risks of their actions to the subsequent links in the chain. On the other hand, in a number of cases, there was a lack of respect for the rules and the way in which they should be implemented.' Furthermore, they recommend that 'every business in the animal feed chain puts the required quality assurance into daily practice' and that this 'requires not only professionalism but also integrity from businessmen' (PDV, 2002, p. 2).

This clearly highlights the two elements of acting as a trustworthy partner and illustrates why problems of trust occur after a scandal like the case of MPA contamination. Some agents have proven to be unworthy of the trust of the public and other partners in the food chain. The well-known scares and scandals such as the BSE case and the MPA contamination are not isolated problems, but are the most explicit moments at which the consumer's need to trust in a global and technology-driven food chain is not responded to or even abused.

---

[20] This was possible since the MPA content in the sugar waste stream was below the so-called 'amber' threshold of 0.5%, meaning it did not have to be classified as hazardous. Later in the investigations it was argued that the fact that a waste is non-hazardous does not automatically imply that it can be labelled as a 'green list' waste (Birchard, 2002, p. 235).

2.5 Trust: Crucial, but complicated

The three characteristics of the agro-food sector analysed above provide the background for the public concern with trust in the agro-food sector. The distance between food production and consumption and the increasing use of technology in the food sector result in a situation in which an individual consumer has to rely on many known and unknown parties with regard to an increasing number of issues. This increase of complexity and perceived uncertainty and the decrease of personal control demand trust, since trust is the way to deal with uncertainty and lack of personal control. However, the same characteristics complicate trust in four ways that are all problems of trustworthiness. The first two are related to inabilities of the individual truster to assess trustworthiness. The latter two problems concern the two elements that constitute the trustworthiness of an agent: competence and motivation.

First, the process of globalisation and the use of technology results in situations in which it is unclear whom one can trust. It might be obvious that one is in a dependent position, but it is not always easy to identify a trustee. For instance, with respect to food safety one may be faced with the problem whether one has to rely on the local supermarket only or also on the multinational that produces the product, or whether trust in the national government is sufficient or whether one has to trust the European government too.

Second, when it is clear whom one has to rely on it is rather difficult to assess the trustworthiness of agents in the current agro-food sector. Since most agents are institutions that operate at a great distance from the individual, and since many agents are involved, assessing trustworthiness has become quite complex. One needs a lot of information and knowledge that is not always available, or hard to evaluate even if accessible.

Third, the process of globalisation and the use of technology raise questions about the competence of those whom one trusts. For instance, one may have trusted the national government with matters of food safety for many years, since they were considered to be competent and motivated by the cares and concerns of the individual. However, since food safety has turned out to be a supranational issue because of the global market, the question is whether the notional government is still competent enough to be trusted. The national government can still be motivated by truster's concerns. However, if they are not competent they are not trustworthy (cf. Baier, 1994a [1986], p. 104; MacLagan, 1998).

Finally, the food scandals have illustrated problems about the competence, motivation and intentions of some trustees in the sector. These scandals have proved both that new situations may arise in which agents are no longer sufficiently competent to be trusted and that competent agents can sometimes faulted for deliberate misconduct.

In short, trust in the agro-food sector is crucial and complicated at the same time. It is crucial because individual actions without trust are hardly possible. It is complicated since the individual truster is confronted with difficulties to identify the trustee, to assess the competence and motivation of that agent, and is faced with a sector in which trusted agents have proven to be less trustworthy than they were expected to be.

2.6 The dimension of the problem: A crisis of trust?

The above section illustrates that trust in the food sector is under pressure: there is a clear need to rely on others, but it is often not so clear who can be trusted. This raises the question of the dimension of the problem. This has been regularly addressed in different works. [21]

According to some, we have to do with a crisis of trust in the food sector or, more generally, we live in an age of distrust. Especially public trust in government and science is in a critical phase after the outbreaks of food-related animal diseases such as BSE (House of Lords, 2000, Ch. 1). But others argue that we still consume food and since consumption without trust is not possible, there obviously is still trust. However, this conclusion is too easy. Food consumption is complicated, but not impossible when one lacks trust in one of the key agents of the sector. First, one can shift one's trust to another party, for instance to the local butcher whom one still considers as competent and properly motivated to act in the expected way, i.e. one considers him trustworthy. Second, even when one does not trust an institution, one still may have reasons to rely on it in practice. One such reason is that everyday life would become too difficult if one did not rely on some institution or a certain production method. Another reason is that it would be simply impossible not to rely, because the alternative is to live in a state of permanent uncertainty and not to consume at all. Therefore, it is possible to answer in a survey that one does not trust the current food system, but does buy food in the supermarket. Thus, the fact that consumers still buy products does not yet settle whether we are faced with a crisis of trust.

Empirical evidence can help at this stage. The results of empirical studies show that the notion of a crisis of trust is not applicable to the European food sector in general. For instance, empirical data from the Netherlands show that a vast majority of consumers still have trust in their food products and in the sector (De Jonge et al. 2005; TNO, 2003). The term 'crisis of trust' does not seem appropriate in this context. On the other hand, when we consider the UK data for public trust after the BSE affair, the situation seems to be very close to such a crisis. Yet even here we should be hesitant to draw

---

[21] See: House of Lords, 2000; Millstone & Van Zwanenberg, 2000; O'Neill, 2002b; Byrne, 2002; Harbers, 2003

general conclusions. A recent European study has put forward evidence that public trust in the UK has substantially improved (Kjaernes, 2004; 2006). In general, empirical studies highlight that there is considerable evidence in contrast to 'the widespread belief about a "crisis of confidence" in scientific and technological institutions in European society' (Gaskell et al. 2003, p. 32; also O'Neill, 2002b, p. 15-17; Dekker & van der Meer, 2004).

This illustrates that we are not faced with a real crisis of trust. We do not trust less today, nor is there a state of distrust because of which one cannot trust food anymore. However, there is an increase in the number of situations where individuals are uncertain about what they can expect of another and whether the other is trustworthy. The problem is that one is dissatisfied with one or more stakeholders, because they have not met the expectations inherent in a trusting relationship. If we define problems of trust in terms of problems with respect to trustworthiness, then we can see why in the Netherlands trust is considered by many as a central issue for the agro-food sector that needs (more) attention, although we lack empirical evidence for a real crisis of trust (Beekman & Brom 2000; Commissie Wijffels, 2001; Wubben et al. 2003).

2.7 Assurance based on technocratic policy

In the preceding sections the problem of trust has turned out to be a problem at the level of trustworthiness. To elaborate this observation in further detail a discussion of the different food policies, especially those on food safety, is helpful.

Traditionally, governmental policy with respect to agricultural and food issues has been a rather closed system based on science and preserved by public trust. The governments presented problems of issues such as food safety, quality, and animal disease as technical matters that they had completely under control. On the basis of 'sound science' policy makers decided what products were safe or what animal disease prevention strategy was most efficient. Risks did not play an explicit role, not to mention uncertainties. Based upon the scientific input, policy makers could assure the consumer or citizen that there was nothing to worry about. The public could trust the government since the authorities pretended to have full control over the situation and were not confronted with uncertainties at all. This behavioural pattern, however, turned out to be completely intractible and to raise serious problems of trust.

The policy on BSE in the United Kingdom is a clear example of this idea of policy: a rather closed system based on science only. Before 1996, both UK and European officials insisted that the policies on this animal disease were purely based on sound science and concerned only with risks to public health (Millstone, 2006, p. 40; Millstone & Van Zwanenberg, 2002, p. 598). Citizens were reassured that no problems were to be expected and no risks concerning pubic health or food safety were communicated. The

government could be trusted because all was under control. This 'excessively reassuring language about the risk' had far-reaching consequences as the *BSE Inquiry* has observed: it 'sedated those who needed to act' (Phillips, Bridgeman & Ferguson-Smith, 2000, p. 205). Within this closed system there was of course an element of uncertainty, but this was not communicated, partly because one had the idea that one had control over the situation with the help of scientific risk assessment, and partly because one was frightened about what would happen if one would inform the public about the known and unknown risks and dangers. There was a 'widespread belief that the general public were unable to conceptualize uncertainties associated with risk management processes' (Frewer et al. 2003b, p. 75; Millstone & Van Zwanenberg, 2000).

This talk of science-based policy led the public to trust the government. The view that science is value-free, objective, and capable of providing an indisputable input into policy, was combined with a rather technocratic idea of policy making. This resulted in a genuine belief that the government could sincerely assure the public that they could be trusted with matters of food policy. They presented themselves as a genuine trustee: fully competent and adequately motivated for the protection of public health. They assured the individual citizen that they were in full control of issues that an individual could not deal with and had to entrust to the government.

The problems of this policy became evident on 20 March 1996, when the UK government had to announce that cases of the human Creutzfeldt-Jakob disease are linked to exposure to BSE. Suddenly, governmental policy turned out to rest on the advice of a closed and relatively small group of scientific experts that was not under democratic scrutiny. Moreover, the government appeared to have made decisions that were not all the result of scientific risk assessment only. Furthermore, it turned out that in contrast to what had been communicated, the protection of public health was not the main, or at least not the only consideration in drafting policy measures. At the end of the day, the government proved not to be in control of the situation as they claimed to be (cf. Millstone & Van Zwanenberg, 2002). Consequently, the reassurances that had been given to the public turned out to be empty. The result was widely observed and analysed: the public felt that the government had not been telling the truth and that they had been deceived (Phillips et al. 2000, p. 235). The decline of trust in political institutions was evident. Even years later the BSE affair, together with some subsequent food scandals, continued to have noticeable effect on trust. For instance, in 2003 the then German Minister of Consumer Protection, Food and Agriculture stated that 'consumer trust in the safety of foodstuffs as well as in official risk assessment was badly shaken in the past'[22] (Künast, 2003).

---

[22] 'Das Vertrauen der Verbraucherinnen und Verbraucher in die Sicherheit ihrer Lebensmittel und auch in die offizielle Risikobewertung wurde in der Vergangenheit schwer erschüttert.'

From 1996 on, it has become clear to many in Europe that the UK approach to the BSE problems should give cause in reforming past policies, so that risks and uncertainties can be dealt with better, a clearer view on the science–policy–society relation is obtained and rebuilt and maintained.[23] The newly formulated policies that aim to address these issues contain roughly four elements: (a) increased attention to the assessment and management of risk, (b) the establishment of new organisations, (c) increased emphasis on transparency and information and (d) encouragement of public participation (cf. Korthals, 2004; LNV/ FAO, 2004; Rowe, 2007). I discuss the first two and the latter two as pairs for the reason that they are strongly interrelated and illustrate that these measures will be effective only if the problem of trust is defined as one of trustworthiness. The next section analyses the increased attention to risk assessment and management and the consequences for the reorganisation of governmental organisations (Section 2.8). The elements of transparency and public participation are discussed in Section 2.9.

2.8 Independence and a division of labour

The BSE affair, like many other food-related scandals, illustrated that it is no longer tenable to withhold risks and uncertainties from the public and to present the competence of the government as almost omnipotent. It has been stressed that governmental bodies should not only speak about safety, but also about risks, i.e. about the considerations that underlie the evaluation of a product or compound as safe. Moreover, it has been argued that the role of science should be explicated. Hence, a clear distinction has been introduced between independent scientific advice on risks on the one hand, and the policy decisions that are made by the administration or the parliament on the other. In line with this distinction, new governmental organisations have been established that focus on the scientific assessment of risks rather than on the policy decisions. These measures are intended to lead to science-based policies that at the same time take into account that policy decisions on matters of safety will have to include the social, political and cultural context. This distinction between science and policy is based on a division of risk analysis in three parts: risk assessment, risk management and risk communication – the 'Red Book Model'.[24]

Before the discussion of the three steps of risk analysis, some remarks are needed on the relation between (a) risk and danger and (b) risk and hazard. The idea of a risk

---

[23] For a clear analysis of the interplay between facts and evaluative judgments in the policy on BSE see Karsten Klint Jensen (2004), 'BSE in the UK: Why the risk communication strategy failed', *Journal of Agricultural and Environmental Ethics,* 17*,* pp. 405-423.
[24] Officially the book is called *Risk Assessment in the Federal Government: Managing the Process* that was issued by the US National Research Council in 1983. The volume had a red cover.

analysis marks the conversion or transition of danger into risk. While a danger is something that can happen to us, a risk can be analysed, evaluated as acceptable, and one can choose to accept a risk. What is considered as a risk and what as a danger depends on the role human action plays in the occurrence and probability of a hazard. An outbreak of an animal disease has long been considered as a danger. For the individual farmer it is something that happens to him. However, nowadays it can also be seen as a risk. The way the animal sector is organised and animal transportation is structured has a profound influence on the number of outbreaks. Human interaction can both increase and lower the probability that an outbreak occurs. Consequently, it is still a hazard, but no longer a danger. It has turned into a risk that can be analysed and evaluated. Although none of the affected farmers probably chose to accept the outbreak of a disease, the occurrence is no longer seen as the result of nature only. It can be attributed to human choices, which is a precondition to speak about risk (Luhmann, 1988).

Furthermore, the relation between risk and hazard needs some clarification. Okrent (1980) presented an illuminating example to show the relation between them. In the example, three people are crossing the Atlantic in a rowboat. They face the hazard of drowning. At the same time, three hundred people are crossing the Atlantic in an ocean liner. They face the same hazard of drowning. Clearly, the hazard is the same for each individual, either in the rowboat or on the ocean liner. However, since the risk is given by the size of the hazard multiplied by the probability of the hazard, the risk is greater for the individuals in the rowboat than in the ocean liner (p. 372). The transition from a hazard to risk is the inclusion of the probability that a hazard occurs. In short, these examples clarify that risk (R) can be defined as the probability (p) of the occurrence of a hazard (h): $R = p \; x \; h$.

2.8.1 The three parts of risk analysis

'Risk assessment' as the first part of the process is defined as 'the scientific evaluation of known or potential adverse health effects resulting from foodborne hazards. The process consists of the following steps: (i) hazard identification, (ii) hazard characterisation (iii) exposure assessment and (iv) risk characterisation' (FAO/WHO 1995, p. 6). The definition clearly emphasises the scientific and descriptive character of this stage. The aim is to provide quantitative expressions of risk. This entails that the assessments aim at quantifying the probability of occurrence of a hazard.

In contrast to the descriptive character of the assessment, risk management is defined as 'the process of weighing policy alternatives in the light of the results of risk assessment and, if required, selecting and implementing appropriate control options, including regulatory measures' (FAO/WHO, 1997, p. 4). The context here is one of regulation and policy. It is important to note that this management part is not simply

included to minimise risk, but to manage the identified and assessed risks that are associated with food in such a way that the management protects public health (cf. Brom, 2002). At this level it is decided what can be considered as an acceptable risk, i.e. what is considered as safe food and what is not. The expression 'in the light of the results of risk assessment' is sometimes understood as if the assessment were to be the only input that plays a role in managing risks. However, many other aspects such as cultural identity, values, and political considerations play a role in this part of the risks analysis too.

Finally, the third stage of the risk analysis process is risk communication. This communication is defined as 'the exchange of information and opinions concerning risk and risk-related factors among risk assessors, risk managers, consumers and other interested parties' (FAO/WHO, 1998, p. 3). Although it is mentioned at the end of the process, it is considered as an integral part of risk analysis and is also essential as a tool to 'define issues and to develop, understand and arrive at the best risk-management decisions' (1998, p. 4).

2.8.2 Newly established food authorities

The above 'division of labour' between risk assessment and risk management, combined with a strong emphasis on the independence of the scientific advice, can be clearly recognised in the several food authorities[25] that have been established all over Europe (Millstone, 2006, pp, 42-43). For instance, in the Corporate Brochure of the European Food Safety Authority (EFSA) the aim is defined as to provide 'independent scientific advice and clear communication on existing and emerging risks.' The independence is stressed. The authority 'serves but operates independently from the EU institutions.' They claim a certain autonomy, a position that is not directly steered by policy issues. Implicitly this is a claim to their trustworthiness as a scientific body. They are not only competent as they claim to deliver 'the best science at the right time and in the most appropriate manner', but they are also adequately motivated to do so since their motivation is not interfered with by any policy interests. Hence, it is not surprising that managing risks and taking decisions are clearly distinguished from EFSA's aim and scope. EFSA aims to undertake risk assessment and risk communications 'to enable

---

[25] When we focus on Western Europe we find the French food authority, L'Agence française de sécurité sanitaire des aliments (AFSSA), started in 1999. In 2000 Denmark, Belgium, and the UK established the Fødevaredirektoratets, the Voedselagentschap (FAVV), and the Food Standard Agency (FSA) respectively. The year 2002 brought a kind of establishment boom of agencies. On the national level the German Bundesamt für Verbraucherschutz und Lebensmittelsicherheit (BVL), the Österreichische Agentur für Gesundheit und Ernährungssicherheit (AGES) in Austria and the Voedsel en Warenauthoriteit (VWA) in The Netherlands were all established in 2002. In the same year the European Food Safety Authority (EFSA) was the legal offspring of the European Parliament and Council regulation (EC) No 178/2002.

effective and timely risk-management decisions to be taken by the European Commission, Member States and the European Parliament' (EFSA, 2006). In this way, they aspire to improve the quality of the risk analysis, but also to 'ensure a high level of consumer protection through which consumer confidence can be restored and maintained' (CEC, 2000, p. 14).

The fact that the problem of trust is taken as an incentive to improve the quality of risk analysis shows that it is recognised as a problem of trustworthiness. The British Food Standard Agency (FSA) is even more explicit when it states that one of its key aims is to 'earn people's trust by what we do and how we do it and to address the issue of consumer confidence' (FSA, 2001). In addition, the French AFSSA hopes to be considered as independent and trustworthy.[26] In all three cases, the aim is to present the authorities as competent and adequately motivated. Here 'adequately' implies that the vulnerable position of the consumer is taken as the origin of the motivation to act in a trust-responsive way. This suggests acknowledging the autonomy of the truster, which requires that the vulnerability of the individual agent who is confronted with uncertainty is taken seriously. Here a serious step has been taken in comparison to the 'pre-BSE policies'.

Nevertheless, the current approach runs the same risk as the pre-BSE policies, i.e. the risk that it starts as a way to assure the public, but ends as the cause of problems of public trust. If the food authorities start from too optimistic a view on science and an overly strict division between assessing and managing risks, they have to claim a competence that they can hardly live up to. Consequently, problems of trust will arise. The next sections show why an optimistic view on science and too strict a division of labour are positions that have to be avoided.

2.8.3 Optimism on science

The first position that has to be avoided is the view that risk assessment is a matter of independent and value-free science. The problems of this position are that it starts from too optimistic and positivistic a view on science and incorrectly defines the problem of trust as a problem of the individual.

It is obvious that the public often perceives problems of risk and safety ways other than scientific experts do. This can easily lead to problems of trust and is, according to Dreyer and Renn, even the main reason for the policy attention to trust. 'Trust-building' they say, 'appears as the "general cure" for public pressures resulting from perceptions of food threats which diverge from the science-based expert perspective' (2007, p. 552). This difference in perception has to be acknowledged. The question however is how.

---

[26] AFSSA, 2001, *Etude notoriété et image, novembre 2001*,
http://www.afssa.fr/ftp/afssa/divers/maxicom.pdf (accessed 1 July 2006).

There is still a strong tendency to define the expert view on risks as the correct and objective one and the perception of the public as hypothetical, emotional and irrational (Slovic, 1999, p. 690). Forsythe claims the main problem is not at the level of science, but at the level of the public's risk perception. According to him, 'even good science does little to reduce these perceptions of risk. The public does not understand dose response and looks at risk from a different point of view than scientists' (1993, p. 1152).

From this perspective the problem is that people do not understand what the real risks are, or act irrationally with regard to risks. In order to address this situation adequately it has been suggested that the public be informed about the 'real risks' that are identified by science and on scientific grounds alone. Consequently, the problem of trust is defined as a problem of the individual. He does not understand the real risks at stake, or has incorrect beliefs. Hence, he cannot trust. In order to address this problem the individual's beliefs and perceptions have to be adjusted to scientific evidence, since only sound scientific evidence can show the risks involved and indicate how they should be assessed. Therefore, improved risk assessment and an increase of information are proposed to address the public hesitance and distrust.

This approach starts from a view that underexposes the use of fundamental assumptions in science. Results of scientific research are often taken at face value. Given the 'overwhelming success of science and its importance in our lives', it might be understandable that 'we are so impressed with science that we give quite amazing credibility to any claim that is successfully presented as being scientific' (Newton-Smith, 2000a, pp. 1, 2). It seems that science can provide answers that are beyond discussion because they are based on scientific facts only. Traditionally, science was considered as free from social factors or value-laden considerations. Sometimes it is granted that social factors may play a role in the generation of theories, but that such 'corruptions' are filtered out in the process of testing. The idea is that science has the ability to filter social, moral and political factors. This view, however, is untenable, because research from many disciplines has proven that 'science seems to be shot through with social factors' (Brown, 2000, p. 448). It is inevitable that science is influenced by social, ethical and political factors. There is no option to do science in a manner that is value-free, since the choice to filter out values or social factors from the scientific debate is a choice in itself that does not result from science only but is a decision on values. For this reason the inclusion of these broader factors in science cannot be avoided and reflection on these factors is constantly needed (Brown, 2000; Okruhlik, 1998; Newton-Smith, 2000a; 2000b).

The persistent influence of social factors and decisions on values also holds for scientific risk research. Assessing risks or uncertainties[27] is more than a matter of a

---

[27] In the case of risk both damage and probabilities are known. In the case of uncertainty the damage possibilities are known, yet we lack information about the probabilities (Wynne, 1992, p. 114).

calculation based upon mere facts and probabilities. An emphasis on the independent and scientific nature of risk assessments 'tends to signal, misleadingly, that decisions are made exclusively on scientific grounds' (Rasmussen & Jensen, 2005, p. 9). This signal is misleading because not only scientific dimensions play a role in risk assessment. In defining hazards as risks and in the evaluation of probabilities, decisions have been taken (implicitly). Such decisions are based upon arguments other than mere facts. Jensen and Sandøe accurately state that risk assessments are determined by 'the exact choice of putatively hazardous ... to be assessed for possible unwanted consequences, and by the exact demarcation in time and space of the possible consequences to be addressed' (2002, p. 247). In addition, the ways in which the results are presented do not directly follow from scientific arguments exclusively. For instance, the choice to express the risks in terms of the number of illnesses or the number of deaths substantially influences the ranking process (Fischer et al. 2005, p. 504), but the choice is not in itself part of the risk assessment. Such a choice does affect the outcome, but is not the result of a scientific process. Each of these choices and every theoretical method requires the inclusion of theoretical and factual premises. Given this influence of value assumptions and other social and political factors it is not only the awareness of the role these factors play that is important, but there is also a need for critical reflection in order to 'lessen the chances of one social factor being systematically present in all the rival theories' (Brown, 2000, p. 448).

The explication of these broader issues does not only show that risk assessment is more than an issue of mere science or even an analysis of facts only, but it also illustrates that it would be too easy to claim that the gap between the views of scientific experts and the public is the result of irrationality or just ignorance on the part of non-scientists. Scientists have scientific qualifications that make them experts in their field of research. But this does not entail that they are experts 'on moral and ethical issues, or on problems outside their area of expertise' (Hayward, 1997, p. 344). The suggested distinction between the scientist who assesses risks and the public who perceive risks is untenable. The public has its own models and assumptions to assess risks. This is what Slovic calls 'intuitive risk assessment' (1996, p. 3). The situation, then, is not one of correct and incorrect assumptions, or rational and irrational views, but between different types of risk perception. In defining hazards as risks and risks as reasons for safety measures, decisions are (implicitly) taken that are based upon views on science, society, and politics. Slovic even goes a step further and argues that we cannot speak about 'real risks' (1996; 1999). The suggested clearness and objectivity of the input and results of risk assessment is often not justified. This does not imply that the concept of reality does not play any role with regard to risk. The danger is real, yet risks are socially constructed. For instance, a hurricane can cause real adverse effects. Risks can serve as a way to handle these dangers, to anticipate on hazards and to go beyond mere uncertainty. However, this requires the use of basic assumptions and theoretical models that go

beyond the level of 'mere' facts. There is no such thing as value-free science that only deals with the 'real' risks. This is no problem for a risk analysis, but has to be acknowledged. It would be a problem if the results of risk assessment were treated as the outcome of a fully objective and value-free process while they do not and cannot meet these standards. This can lead to a serious problem of trustworthiness even though the intention is to improve trust.

2.8.4 Problems of the division of labour

It has been widely argued that when food safety is entangled with commercial or political interests in way that ignores the position of the consumer/citizen, questions of the government's sincerity and integrity will immediately arise (cf. Millstone & Van Zwanenberg, 2002; Waltner-Toews & Lang, 2000; O'Neill, 2002b). Hence, the introduced division of labour is useful in the context of trust, because it explicates the relation between the input from science and from politics. It enables to acknowledge that policy decisions on matters of safety include the social, political and cultural context without denying that the measures are based on scientific research.

However, the above discussion on the inevitable impact of ethical and social factors in science also has implications for this division of labour. It shows that the distinction should be handled with care. If the division between risk assessment and risk management is referred to as a distinction between value-free and objective science on the one hand, and social, political and ethical considerations on the other hand, it will only generate false expectations, because this is inadequate. To take assessing and managing risks as two fully independent parts in the process of risk analysis burdens risk assessment with a weight that it cannot bear. Risk assessment cannot operate in a trustworthy manner within a strict scheme of a science–policy distinction. Current practice shows that assessing risks always involves ethical and epistemological conflicts that demand discussion of a broad range of issues rather than about risks only. Even more important is that risk assessment emphatically should not bear the weight of aiming at a science that is free of value judgements because this is not something worthwhile striving for. Risk assessment without elementary assumptions of science, society and morality is in principle impossible, as I have argued above. It is always based on factual, cultural and moral premises. There is no clear watershed between assessing and managing risks with respect to the role of social, political and ethical considerations. Former EU-commissioner Byrne (2002) emphasizes that risk managers, 'have to take into consideration not only science but also other matters – for example economic, societal, traditional, ethical or environmental factors, as well as the feasibility of controls.' He is right, and what he says is no less true for the assessment of risks. Attention to these factors cannot be postponed until 'purely scientific assessment' is done and risk management and policy-making begin. Neither can the policy-makers hide

behind the scientific and independent level of the risk assessment. They also have to take into account the fact that in risk assessment decisions that directly influence the policy options have already been made.

Moreover, a strict compartmentalization suggests that scientific risk assessment yields only one single answer that serves as the input for the stage of policy-making. This ignores the fundamental uncertainty that is inherent to scientific evidence. Rasmussen and Jensen observe a direct link between the opacity of uncertainty and problems of trust. They argue that 'a large part of the loss of confidence in risk assessments arises from the fact that such assessments seldom present their own limitations and uncertainties' (2005, p. 9). The risk assessors, more generally the food agencies, can be asked to calculate as accurately as possible, yet not to act as if no uncertainty cleaves to their findings and as if the evidence is beyond discussion. Waltner-Toews and Lang correctly stress that 'no agency, any more than any Ministry, can be expected on its own to resolve this philosophical problem' (2000, p. 125). If they try to do so, either voluntarily or by force of a policy system, they may be adequately motivated, but because they are not competent to live up to this standard, they will in time turn out to be untrustworthy.

In short, the division between assessing and managing risks and the establishment of new safety authorities can contribute to the maintenance and establishment of trust if and only if (a) the problem is recognised as a problem of trustworthiness rather than as a problem that results from the ignorance or irrationality of the individual, (b) the fundamental premises of science are acknowledged and (c) there is an awareness that the influence of social, cultural, political, and ethical considerations is not restricted to the level of risk management, but emphatically also plays an inherent role in scientific risk assessment. If these three aspects are taken seriously, problems of trust with respect to food safety are no longer addressed as problems of the ignorance or irrationality of the public. This change in the way of addressing problems of trust is a precondition to take autonomy of the individual seriously and to fully acknowledge his vulnerable position in the agro-food sector. Furthermore, it frees science from a requirement that it cannot live up to. Science cannot get rid of all uncertainty and cannot avoid making fundamental assumptions with respect to a whole range of issues, including those of ethics and politics. Therefore, acting as if it could would only result in more problems of trust as science lacks this competence. Finally, the inclusion of social and ethical factors and the acknowledgment and explication of uncertainty in the assessment process also enable risk managers to act more trustworthily. Their job becomes more complicated, because they can no longer hide behind the presumed objectivity of the scientific evidence. However, they are now in a position to present differences in the risk management as the result of discussion with respect to political, social and ethical considerations and beliefs.

2.9 Transparency and participation

The restructuring of the food safety policies in Europe does not only lead to an emphasis on risk analysis, but also to an increased emphasis on transparency and information, and the encouragement of public participation in the policy process. In the *Corporate Brochure* of EFSA these elements can clearly be recognised: 'Transparency and independence are key principles of good science and fundamental to building consumer confidence.' The authority aims to 'listen to views through public consultations that have included guidance on the risk assessment of genetically modified organisms and harmonised approaches for risk assessment of additives used in animal feed' (EFSA, 2006, also EFSA, 2005). Openness and transparency are essential with respect to trust, since 'confidence was undermined by a climate of secrecy in the past' (FSA, 2001, p. 23). Nevertheless, transparency and public participation do not automatically improve the level of trust.

2.9.1 Openness and transparency: More than addressing a deficit

Information facilitates trust. In order to assess the competence and motivation of a government that enforces food safety regulations we need information. If we lack all information, we have nothing to go on to form expectations, either negative or positive. However, openness and transparency will have little impact on the problems of trust if they are restricted to factual information and framed in terms of the education of the public.

When openness and transparency are referred to as parts of public education and information they are often a signal of what is called the 'deficit model' (Frewer, 2000; Scholderer & Frewer et al. 2003a). This model starts from the assumption that the public has a gap in its knowledge, a deficit which among other things leads to problems of trust. Since the individuals are no longer in a position to assess complex issues such as food safety and food quality, they have a serious knowledge gap. This disability yields 'unfounded (negative) judgments that need to and can be corrected by providing the "right" information' (Paula & Birrer, 2006, p. 261). The hope is that additional and correct information provided by the scientific elite will fill this gap and enable a layperson to make a better assessment, or convince him that his initial reactions were unfounded or even irrational. From this perspective, communication is a matter of education and correction by an increase of the flow of information from the experts to the lay audience. The result is a top-down way of providing factual data in order to change the opinion of consumers. However, Paula and Birrer claim that 'even comprehensive, genuine efforts to provide "hard" data to the public… will, in itself, not suffice to build trust' (p. 261). They are right because the interpretation of openness and transparency in the 'deficit model' is problematic: it frames problems of trust as

problems of the inability of the individual, and these are best addressed by factual information. This is a distorted image since (a) trust is about more than facts, (b) the image mainly ignores the element of trustworthiness and (c) it starts from a view of consumers that does not recognise them as full participants in the sector.

If problems of trust are considered to be due to lack of information and knowledge, the best way to address these problems is to educate and inform the truster. It is supposed that the truster lacks knowledge about what is going on and what is at stake. As a result, much effort is spent on giving factual information to people. However, these 'facts' are not as factual as they are presented. Moreover, trust is not only based on facts.[28] Former EU-commissioner Byrne stated that 'the science-based message simply fails to get across, or if it does, it is ignored [by the public, FM]' (2004, p. 3). This predicament is not the result of unclarity about the facts at stake or the unwillingness of the public. Rather, it arose because communicating in order to build or maintain trust is more than filling knowledge gaps with factual information. Analyses of public debates on controversial issues in the agro-food sector have illustrated that communication is more than just an exchange of factual data that are the result of scientific research only. What is at stake are not only facts, but also value judgements, theoretical backgrounds, and political deliberation (cf. Sandøe & Jensen, 2002; Korthals, 2004). The debates during and after the incidence of foot-and-mouth disease can serve as an example.[29] The outbreak of FMD in 2001 has shown that the policy not to vaccinate animals, which was common up until 2001, was not based on sheer facts alone. Decisions whether or not to vaccinate the animals, or which animals had to be killed, were presented as measures which had been taken on the basis of facts only. However, in retrospect they turned out to include many evaluative judgements too, such as the weighing of economical consequences versus the impact of an animal disease on animal welfare, or versus the social impact of a crisis. Based on the facts that were available at the time that the decisions were made, one also could have concluded that vaccination was the preferable solution to address FMD. Hence, communication as a tool to enhance trust cannot be limited to showing *what* has been done or what is normal, but should also include *why* a particular choice has been made and how a specific decision can be justified. There has to be explicit attention to arguments and evaluative judgments that underlie the position or decision. This implies that both normative assumptions and factual premises have to be explicated and have to be subject of discussion. This view can be recognised in a

---

[28] Another aspect is that trust is not just knowledge based. Trust is an attitude that has a clear affective element that colours the perception of information. Consequently two agents can interpret the same type and amount of information completely differently. I will elaborate on this aspect in further detail in the next chapter.

[29] For another example of a clear analysis of this interplay between facts and evaluative judgments with regard to food-related issues see the elaboration of Karsten Klint Jensen (2004) on the BSE crisis: 'BSE in the UK: Why the risk communication strategy failed', *Journal of Agricultural and Environmental Ethics,* 17, pp. 405-423.

report by the FAO that stresses the importance of transparency for trust in the food sector. The report emphasises that 'the values embedded in the decisions that underpin that system therefore need explication, to make decision making more transparent and provide a better understanding of the choices we exercise in the governance of food safety' (FAO, 2003, p. 5).

Another difficulty of the interpretation of trust as an issue of knowledge and information is the underexposure of the element of trustworthiness. The deficit model locates the problem of trust in the individual truster. The information that is necessary to trust is available, but people lack this information or have a deficient perception of the situation. It is widely acknowledged that this entails a responsibility on the part of the trusted agents to be as open as possible and to act transparently. Nevertheless, the individual is conceived as the source of the problem. This view makes the problem of trust a rather intangible issue. Public communication is highly difficult, not to mention the aim to change the beliefs and perceptions of individuals. Nevertheless, openness and transparency are relevant, but their relevance lies in the fact that they enable the truster to check and assess the competence and motivation of the trusted agent rather than to fill knowledge gaps. The Dutch Ministry of Home Affairs and Kingdom Relations mentions the connection between trust and trustworthiness in the right order: 'Greater transparency in governmental performance forces the government not only to be careful and trustworthy, but also contributes to trust in the government' (BZK, 1999, p. 35).[30] Transparency is relevant for trust building if it is understood as a tool to assess trustworthiness. The aim is not so much the full disclosure of all available information since more information does not necessarily lead to more trust. Transparency as a way to asses trustworthiness should involve 'uncovering, describing, documenting and communicating all the argumentative steps in the line of reasoning and the weighing of evidence leading to and justifying the final decision.' and this should also include 'limitations, weaknesses and uncertainties' (Rasmussen & Jensen, 2005, p. 11). This broader interpretation of transparency enables an agent to assess the competence and motivation of the trusted agent. For both competence and motivation, the inclusion of normative aspects in the concept of transparency is crucial. With regard to competence, being transparent as to the relevant moral aspects is important because only then it will be clear that the trustee is able to handle the normative deliberation he must engage in to deal with issues such as the safety of genetically modified food or the policy on the prevention of foot-and-mouth disease. Moreover, communication of the evaluative judgements will give the truster an indication of the other party's motivation to act in a certain way. When we know that an agent is motivated by social or normative rules that we agree on, or at least that we consider as relevant in the specific case, then this makes

---

[30] Original quote in Dutch: 'Een grote transparantie in het functioneren van de overheid dwingt de overheid niet alleen zorgvuldig en betrouwbaar te zijn in haar optreden, maar draagt ook bij aan het vertrouwen in de overheid.'

her trustworthy. Part of this is that the trustee is more predictable since one may expect that she will act in line with her moral framework. Another part of this is that the truster gets an indication that the trustee will recognise him as an individual who is worthwhile protecting for his own sake and who is acknowledged as an agent.

Finally, the deficit model starts from a view on consumers that does not recognise them as full participants in the sector, that is to say, the idea is that they need the additional information and knowledge to be enabled to function as a full participant. This, however, ignores that trust starts from a participant attitude towards the other agent. Only if one considers the other person as a participant is it possible to enter a trustful relationship. Let us explore this a little further.

2.9.2 Participation and the recognition of the moral agent

In the policies that aim to address questions of trust, there is a clear tendency to encourage public participation, for instance by organising hearings and public debates and by appointing members of the public to official committees that evaluate parts of the food sector. The efficacy of these measures depends once again on the way the problem of trust is defined. If it starts from the idea that the individual consumer has to be involved since he does not know what is going on, we are back in the 'deficit model'. Then participation is no more than an extended tool of education and information. The public may be involved, but they are no real participants. The problems of this perspective were discussed above.

If public participation aims to be relevant as a way to address issues of trust, the individual truster should not be merely involved, but has to be treated as an agent, i.e. as a full participant in the agro-food sector in spite of his vulnerable position. The reason to invite the truster to be involved is not his need to be educated, but because of his entitlement to participate in the policy process as citizen and consumer. This has immediate consequences for how problems of trust are to be addressed. If agents, who as citizens are entitled to be involved in policy developments in the agro-food sector, cannot trust the governmental authorities, then it is not the problem of the citizen, but the problem of the government to take care that these individuals can act as full participants. This does not mean that the preferred situation would be one in which trust is not needed. In most cases this would be simply impossible, given the complexity of the issue and the broad range of situations in which an individual cannot but rely on others. It does mean that the individual citizen should be enabled to trust institutions in order to enable them to act in spite of the uncertainty and the lack of personal control. From this perspective, governmental and market institutions can build trust by showing themselves trustworthy rather than by starting up information campaigns to educate the 'lay audience'. If genuine (public) participation presupposes the recognition of the other as an agent and participant it should be no surprise that Poortinga and Pidgeon concluded

from empirical research that involving the public, in for instance a public debate, will not automatically improve the level of trust (2004, p. 1485). When meetings of a Food Safety Agency are open to attend by the public, or even when members of the public are part of the meeting, this does not automatically build trust. This kind of participation is only relevant with respect to trust as far as it offers the opportunity to check the agency's competence and motivation.

When public participation starts from the recognition of the truster as an agent there is another aspect that needs to be emphasized. With respect to trustworthiness, it is not only important that one is recognised as an agent as such, but also as a *moral agent*. The agency of individuals in the agro-food sector is not restricted to the realm of risk and safety. Consumers regularly express their moral beliefs and ideals with respect to food production and consumption in the market and in the public debate. Such 'consumer concerns' reflect public uneasiness with regard to a whole range of issues, such as animal welfare, sustainability, and the introduction of technological innovations. When the public is involved in the process of policy or in a public debate restricted to economical terms only, the moral dimension is disregarded and the truster is not taken seriously as a moral agent. The consequence is that problems of trust remain unresolved because the moral dimension of what is entrusted is not recognised or it its importance not acknowledged.

2.10 Summary: Trustworthiness as the key

Emphasised above was that consumers cannot but rely on other agents in order to operate in today's society. In this chapter I presented the agro-food sector as an incontrovertible example of this. Trust is essential in this context, because it enables an individual to act in spite of the uncertainty occasioned by the lack of personal control.

The analysis of the background of the current attention to public trust in the agro-food sector showed that the problems of trust are mainly the result of problems at the level of trustworthiness. Developments that characterise the agro-food sector, such as the process of globalisation, the increase of the use of technology in food production and the recent food-related scandals stress the individual's need to trust. At the same time these developments complicate trust, since they give rise to or intensify problems concerning trustworthiness. It has become more difficult to identify the agents in whom one is required to place trust and has raised all kinds of practical problems with respect to the assessment of the trustworthiness of identified agents. Furthermore, the global and technology-based food sector with its recent food scandals confronts trusters with questions concerning the competence and motivation of the trusted agent.

The discussion of the policy measures that have been proposed and implemented in order to deal with problems of trust underlines the relevance of addressing the issue from

the perspective of trustworthiness. The increased attention to risk analysis, the establishment of new organisations, the emphasis on transparency and information, and the encouragement of public participation will all be effective if and only if the problem of trust is defined as one of trustworthiness. Otherwise, one has to (i) burden some domains, such as science, with a responsibility that they cannot meet, (ii) presume that individual consumers are mainly ignorant and irrational which is incorrect, and (iii) exclude certain themes, such as morality, from the debate without sufficient justification. The results of former food policies in Europe have shown that failure to address trustworthiness inevitably leads to a serious increase of trust problems.

In conclusion, the core of the 'problem of trust' is not the individual's failure to trust, but concerns questions of trustworthiness. To address this problem the aim should be to prove oneself trustworthy, rather than to focus on the question of how to increase consumer trust.

CHAPTER 3

PROBING THE CORE OF TRUST

Trust has received considerable attention in recent years,
resulting in a confusing potpourri of definitions
Susan Shapiro, 1987, p. 625

3.1 If one starts by saying 'trust is….'

'One finds that if one starts by saying "Morality is…" nothing one says afterward seems to be quite right.' According to Bernard Gert (1988), this is the feeling that arises when discussing morality. A similar feeling emerges in the case of trust. In daily conversation, we use 'trust' many times, in many contexts and regarding quite different subjects. However, what we mean by trust exactly is rather difficult to formulate. Trusting a car seems to be quite different from trusting a friend. Entrusting the care for your child to someone else often is perceived rather differently from entrusting a part of a research project to a colleague. And trust in the personal relationship among close relatives seems quite unlike trust in the public relation between a citizen and the government.

It is likely that Hardin has this diversity in mind when he writes, 'the notion of trust in the vernacular is often vaguely warm and fuzzy' (1999, p. 429). This vagueness is not restricted to the vernacular. Also at the level of academic analysis there is a 'conceptual jungle' (Lindenberg, 2000) and a lack of conceptual clarity (Gambetta, 1988; Hosmer, 1995; Sevenhuijsen, 1999; Hartmann, 2001; Nooteboom, 2002; Harbers 2003). Even though many works on trust were published in sociology, philosophy, studies in political theory and economics during the last decades this situation has not substantially changed. More than once the discussion of trust results in an analysis of other concepts, such as power, rationality, promises or friendship. Yet, even when one sticks to the theme of trust, the concept seems to be intangible, as is reflected in the many attempts to define trust. Back in 1987 Shapiro said that the considerable attention for trust resulted in a 'confusing potpourri of definitions' (p. 625). A brief look at the literature on trust substantiates this claim: for some trust is a bet about the future actions of other agents (Sztompka, 1999, p. 25), while according to others trust is an individual psychological state that entails a generalised and optimistic expectation towards others (e.g. Rotter, 1967). Also on the level of the characteristics of trust the diversity is striking: for some

trust is mainly cognitive (e.g. Hardin, 2002a; 2006), while according to others it is clear that trust mainly has an affective dimension (Jones, 1996; Lahno, 2001; 2002). In order to clarify this field many authors distinguish different types of trust. They distinguish for instance between anticipatory trust and responsive trust (Sztompka, 1999), or trust as habitus, passion, and policy (Misztal, 1996), predictive trust and normative trust (Hollis, 1998), calculative trust and personal trust (Williamson, 1993), and particularized trust and general trust (Uslaner, 2002). This clearly shows the difficulties of analysing trust because of its elusive and multifaceted character (cf. Gambetta, 1988, p. vii). This diversity and complexity would almost suggest that we had better look for more promising concepts, yet Newton seems to be right that 'it is unlikely that there is a trouble-free term amidst the constellation which serves the purpose very much better than the concept of trust' (2001, p. 203).

Given this context, the aim of this chapter is not to provide a definition of trust that all will agree on and that will end all conceptual discussion. My aim is (a) to discuss different characteristics of trust in order to differentiate this concept from related phenomena, (b) to provide a map of trust based on this first discussion, (c) to position this presentation of trust within the academic debate about the concept of trust, and (d) to turn my analysis into a working definition.

As a preliminary to the first analysis of trust, I discuss briefly the importance of basic trust and argue why I do not elaborate on this kind of trust (Section 3.2). In Section 3.3, some basic characteristics of trust are formulated by positioning the concept in the spectrum of control and hope and by discussing its relation to faith, reliance and confidence. This results in a first map of trust. This discussion of trust is still compatible with a wide variety of accounts. In the Sections 3.4 to 3.7, I further analyse trust and position my presentation of the concept within the academic debate. Trust has been referred to as an act, as an expectation or belief, and as an attitude. I discuss these positions and link them to three related themes: (a) trust as an act and the relation between trust and risk (Sections 3.4–3.5), (b) trust as a belief and the relation between trust and empirical evidence (Section 3.6), and (c) trust as an attitude and the emotional or affective element in trusting (Section 3.7). Finally, I present a working definition of trust that will be used in the next chapters (Section 3.8).

3.2 Brief comment on 'basic trust'

As a preliminary to a first map of trust the need for and the importance of so-called 'basic trust' or 'ontological trust' has to be mentioned. Trust is not innate and never emerges automatically. We learn to trust during our life. That process starts in early childhood. Babies are thought to have a primitive or elementary form of trust that is essential for them to thrive and develop. This basic or ontological trust has its origin in

the parent–child relation. The parents are conceived therefore as highly important in one's first experiences with trust. Especially the relation between mother and child is considered crucial for the development of an individual's possibilities to trust. According to Erikson, 'the firm establishment of enduring patterns for the balance of the basic trust over basic mistrust is the first task of the budding personality and therefore first of all a task for maternal care … the amount of trust derived from earliest infantile experience … [depends] on … the quality of the maternal relationship' (1963, p. 149). The patterns and regularities in the parent–child relation show the parents' trustworthiness. In line with Erikson, many authors evaluate this kind of trust as essential for the psychological and social development of each individual and as the most important basis for intimacy, which ensures emotional and moral development (cf. Misztal, 1996, p. 160). Consequently, if this development of basic trust is disrupted, then this may have serious consequences for one's ability to trust another and to participate in society.

In spite of the fundamental importance of the dimension of basic trust, I will not elaborate on it in further detail. For this I have two reasons. First, with respect to the specific account of Erikson, I share Govier's observation that Erikson 'does not really try to define trust but seems to have thought of it as a kind of confidence in regularities' (1997, p. 10). My aim in this chapter is to take a further step by providing an outline and proposing a definition of trust. Second, I acknowledge the importance of basic trust for every situation in which one has to trust, but I take it as a *ceteris paribus* for my thesis. There may be many individual differences concerning basic trust, but I will not go into these differences. The reason is that I approach the problems of trust in the agro-food sector as problems of trustworthiness. The individual differences in basic trust are not crucial from the perspective of trustworthiness. It is obvious that pathological problems of basic trust will frustrate any attempt of acting trustworthily. But the current problems of public trust cannot be limited to such pathologies. Therefore, my aim is to analyse the general conditions of trustworthiness rather than to focus on these specific cases of problematic basic trust.

3.3 Trust and related concepts

As a first step to analyse trust, I will provide a map by positioning trust within the spectrum of five related phenomena: *control, hope, faith, reliance*, and *confidence*.

3.3.1 Between control and hope

To sketch the contours of the concept of trust, we first have to look within the range between feelings of hope on the one hand, and complete control and knowledge of the ins and outs of a situation on the other hand. Trust is located somewhere in between, as

the following quotations from different authors suggest. According to Gambetta, 'Trust is particularly relevant in conditions of ignorance or uncertainty with respect to unknown or unknowable actions of others' (1988, p. 218). Sztompka stresses the element of uncertainty and adds the aspect of control: 'Trusting becomes the crucial strategy for dealing with an uncertain and uncontrollable future' (1999, p. 25). The element of control is also present when Govier says that in trusting another 'we are willing to go ahead without a guarantee. We feel that we can rely or depend on the other, even though there is always some possibility that he or she will act in unexpected ways, or even betray us' (1997, p. 4). These remarks on the relevance of trust illustrate that it is different from hope. This position of trust within the spectrum from control to hope is affirmed if we discuss it in further detail.

If one has full knowledge of and control over a certain subject in a specific situation, then trust is not relevant. When, for instance, two journalists have been taken hostage, it is no longer useful to say that these persons can be trusted not to go home. All measures have been taken in order to prevent them from escaping. Thus, complete control makes trust superfluous. Sztompka correctly states that the expression 'I trust my prisoner not to escape, sounds absurd' (1999, p. 20), but also on a more general level full monitoring and control of somebody's performance makes trust unnecessary (Giddens, 1991, p. 19). Still, the relation of trust and control is complex (Nooteboom, 2002, p. 11). Sometimes they function as substitutes as in the example of the prisoner. However, in that case they also function as complements. Complete control does not rule out trust in general. Even when one has full control over one person, it may imply trust in many others, e.g. in the guards or in those who designed and built the security system. It is possible to take a further step: complete lack of control does not automatically imply the need of trust. For instance, we cannot but rely on the sun to rise; we have no control over its direction or speed. However, given this situation, we have enough knowledge to predict with reasonable certainty when the sun will rise tomorrow morning. We calculate that the sun will rise at 5.30 am rather than trust the sun to do so. Even though we lack control regarding sunrise there is not much uncertainty about the time of its occurring. For this reason we can base our acting on this knowledge without trust. This shows two things. First, lack of control does not always imply that one needs to trust. Secondly, the reason why we do not need to trust that the sun will rise is by reason of the fact that the situation is not coloured by uncertainty.

This illustrates that situations of trust have an element of uncertainty. Trust always presupposes a situation of what one may call ontological uncertainty, that is, an uncertainty that is present and can be identified from a third-person perspective. This uncertainty is to be distinguished from an individually experienced uncertainty. The latter kind of uncertainty is not a necessary condition for trust. Even if one has good reasons to be sure that someone will act in a specific way, so that one does not feel uncertain, it remains possible to define this situation as one of trust, since the trusted

party still has the ability to act contrary to the expectations (cf. Pettit, 1995, pp. 204-5). Consequently, there is uncertainty even though it is not experienced as such by the truster.

In summary, trust plays a role in cases in which agents both lack control of and are uncertain about future actions and events. Trust enables agents to act in spite of this predicament.

On the other end of the spectrum, we find hope. Hope is an attitude that seems to be appropriate to the same situations as trust, that is, situations where one lacks the opportunity to control or predict future events and acts, where one is confronted with contingency. Nevertheless, there is a difference. If we only hope that things will go as expected, we are less certain about whether and why things will turn out for the good. Already Thomas Aquinas has noticed that trust (*fiducia*) is more than just hope.[31] It is possible to say that you hope your house will not be destroyed by a hurricane, yet it is questionable whether it is sensible to say that you trust the hurricane not to do so. While hope is 'a passive, vague, not rationally justified feeling that things will turn out to the good (or to the bad)' (Sztompka, 1999, p. 24), trust is more than such a feeling. Trust is a higher degree of hope, as the German philosopher Christian Wolff argued.[32] According to Luhmann, trust differs from hope in that it has a higher level of reflexivity. One can hope without reason and even in spite of the available knowledge, while trust has a stronger intentional and rational element (2000, pp. 28-29). It is important to note that the difference is phrased in terms of a comparative degree. We should neither reduce hope to mere irrationality and unfounded beliefs, nor trust to pure rationality. Trust can also be against evidence and have an affective element. Nevertheless, trust is always based on an individual's assessment of the trustworthiness of an agent, while hope does not necessarily imply such assessment and can be an attitude that is fully detached from the past and current performance of the subject of hope.

3.3.2 Faith

Having located trust within the spectrum of full control and sheer hope, we can turn to two other related concepts. First, 'faith' is a term that is often mentioned in relation to trust. Many scholars have mentioned the link between them (cf. Simmel, 1950; Holton, 1994). Sometimes trust and faith are used as interchangeable notions. The *Shorter Oxford Dictionary*, for instance, defines trust as 'faith or confidence in the loyalty … of a person or thing.' Uslaner also speaks in terms of faith when he defines his concept of 'generalized trust' as faith in others (2002, pp. 3, 76-77), and Firth argued that the

---

[31] Aquinas states: 'Ad fidem aute pertinet aliquid & aliqui credere; Fiducia autem pertinet ad spem secundum illud' (Summa Theol. 2,2 Q 129).

[32] Original quotation is in German: "Das Vertrauen ist ein hoher Grad der Hoffnung", quoted in Frevert (2003, p. 14).

problems concerning BSE in the UK 'caused the complete collapse of public faith in food-regulating authorities' (1999, p.41). These examples clearly show the interrelation between both: when we trust others, we can have faith in them. If we take a historical perspective, the resemblance between the two notions is even more striking: the Latin *fides* is used for both trust and faith, and the concepts at least have the same root when trust is rendered as *fiducia* (cf. Frevert, 2003). Hence even early modern writers, John Locke for example, use *fides* in a sense that we would currently translate with trust (cf. Dunn, 1988). Faith and trust have more in common than trust and hope. Faith and trust are more intentional and reason-based than hope. Faith and trust are more sensitive to knowledge and information. However, in both cases there still is a remarkable leap between the available information and the given state of affairs on the one hand, and the trustful or faithful attitude on the other hand (cf. Simmel, 1950; Giddens, 1991). Just like acquiring faith, adopting an attitude of trust requires a certain amount of knowledge. However, in both situations one lacks an objective standard, which indicates what amount of information is sufficient to come to trust or to faith. Both attitudes can be evidence based, but 'What I have will always be a belief' (Pagden, 1988, p. 129). When we trust we take a situation for certain or manageable, but we always act in a context of uncertainty and lack of control, i.e. within the spectrum of hope and certainty. With respect to religious faith there is a clear tradition that explains the leap between evidence and an attitude of faith by stressing that faith is not merely a matter of knowledge, but is a receiving (e.g. Ames, 1628, I,1). In a secular setting an explanation of trust as a (divine) receiving is not very illuminating however.[33]

Notwithstanding the resemblance between trust and faith, the precise relation between them is not immediately clear. They are neither synonyms nor interchangeable concepts. Faith always includes an element of trust,[34] but the reverse claim does not hold. Trust need not include an element of faith. A quote from Baier can clarify this point. The question she raises is: 'When, if ever, should we ask for an accounting from those we have trusted – when and for how long should we have faith that all will be well (1991, p. 173)?' Trust includes the possibility to check the trustworthiness of the trustee, and does not necessarily imply the end of the trusting relationship. But to ask for an accounting from those we have faith in would have a more profound impact on the relationship. Both trust and faith imply that one surrenders oneself to another agent. Yet the extent of this surrender differs importantly. Faith is a more categorical attitude. One has faith in a person or in God, period. It is a general attitude that holds for a wide range (or even all) situations and objects, while trust can be restricted to certain objects and indexed to specific contexts. If trust in a person or in God is unrestricted and unconditional, it is similar to faith, yet this is only one part of the spectrum of what we can call trust.

---

[33] I will return to the leap characteristic of trust later on in this chapter.
[34] Faith, however, does not coincide with trust. One's faith in God, for instance, is more than the belief that the Eternal is reliable.

3.3.3 Reliance

Within the range between mere hope and full control it is possible to distinguish between reliance and trust. However, not all agree on the necessity of making the distinction. Coleman (1990) for instance does not differentiate between reliance and trust as he considers both to be ways of taking decisions under risk. According to Hardin (2002) there is no fundamental difference between them either. On the other hand, Baier (1994a [1986]) emphasises the importance and need of differentiating trust from reliance. I agree with her on this point.

There is a difference between trust and reliance that it is useful to explicate in order to understand the notion of trust. Let me explain this with a brief example. Suppose you are travelling in a foreign country and during a walk you arrive at a very deep valley. The only way to cross is to use a wooden bridge. In this case, it sounds as an exaggeration to say that you *trust* the wooden bridge when walking on it. Nevertheless, you *rely* on it in the sense that you expect it be a safe way to cross, e.g. because countless local people have crossed this bridge many times before and met with no mishap. However, this is not yet the same as trust.

This illustrates that reliance is a broader category than trust. Reliance is a necessary, but not a sufficient condition for trust. One can rely without trust, but not vice versa. The difference between both is (a) in the ability of the trusted to refrain from what is entrusted and (b) in the need of an interaction between the truster and trustee.

First, trust requires that the trusted party has the ability to deliberately refrain from doing what is entrusted. This condition indicates why your 'trust' in the bridge is actually reliance. If you really were to trust the bridge not to collapse while crossing it, it should have the ability to refrain from performing according to its construction. The bridge lacks such ability. Therefore, trust is not applicable in this case. Of course, you run the risk that the crossing has a lethal fault and may unexpectedly collapse. Yet if this happens, there is no intentional element or decisional moment on the part of the bridge that results in the unfortunate situation. This condition also holds when humans, who have the ability to refrain, are forced to act in a specific way. For instance, it would be out of place to say that I trust a person not to move when his legs are tied. Even though he has the ability to refrain he does not have the opportunity to go against my expectation. Thus, I do not really trust him. I would trust him if he would be freed and I would confidently rely on him not to move. With respect to public trust in institutions, this condition does not entail that all actions that are regulated or part of specific structures and processes cannot be trusted, but only relied on. Many actions are structured and have specific patterns on which one can anticipate. Nevertheless, the agents within the system still have the ability to refrain from doing what is expected of them. If I trust a national food authority to take care of my concern of food safety, they still have the ability to act contrary to my expectations, e.g. by giving priority to the economic interests of some big multinationals. Structures, laws and sanctions can

substantially reduce the likeliness of such a decision, yet because the agents within the institution have the ability to refrain from doing what is entrusted we need trust. This illustrates that we have to do with a sliding scale. The more an agent is factually forced by structures and sanctions, the less space is there for trust.

Second, trust requires a relationship of interaction between those involved (cf. Lahno, 2002, pp. 138ff). We 'rely on' rather than 'trust' the bridge in the above example. Only if the bridge were to have the ability to respond to our attitude, we could define our attitude as trust and the bridge as a trustee. We depend on the bridge but the relationship is not mutual. Our trust does not change the performance of the bridge in any way, since the bridge does not have the ability to respond. It cannot change its performance in answer to your trust. The crucial element is the ability rather than the actual response. To define an attitude as trust the actual response of the trustee is not a necessary condition. If the trusted agent does not respond to someone's trust this does not turn trust into reliance. It would if the trusted agent did not have the ability to respond to what is entrusted. Furthermore, the response does not need to be directed towards one specific case of trust. For instance, it is unlikely that a national food authority responds directly to my trust. They may even not be aware that I trust them. Nevertheless, my trust is more than just a matter of reliance. If the actions and decisions of such an institution are influenced or even directly induced by the awareness that consumers entrust them the care of food safety, then it is possible to speak about trust, because the behaviour of the institution shows that they change their acts and policy as an answer to the fact that other agents depend on them. My trust may not change the outcome, but the fact that consumers such as I trust them influences their policy. This feature makes them a party that can be trusted.[35]

The two conditions that distinguish trust from reliance explain why predictability is crucial to reliance. Predictability helps to anticipate the behaviour or performance of the other person or object. In contrast to reliance, trusting includes the ability to 'read' others and to make a decision about whether someone is trustworthy (Eckel & Wilson, 2004, p. 464).

3.3.4 Confidence

Confidence is often used analogously with trust. Both entail a positive expectation towards the other party. Nevertheless, some argue that they are not synonymous. Authors such as Luhmann (1988) and Seligman (1997) have underscored the importance of a distinction between them.

---

[35] This is a formal condition. The fact that they can be trusted does not yet have implications for their trustworthiness.

Luhmann was the first to distinguish confidence and trust systematically. He emphasises that both involve the formulation of expectations in cases of uncertainty and lack of control. The difference however is related to the transition of danger to risk.[36] If one were confronted with a situation of danger, acceptance would be a case of confidence. To deal with this danger one has no other option than to form a positive expectation with regard to someone or something; otherwise one will be paralysed by the apparent danger. This is the case, as Luhmann says, when we rely 'that politicians will try to avoid war, that cars will not break down or suddenly leave the street and hit you on your Sunday afternoon walk.' In adopting this stance of confidence, you 'neglect, more or less, the possibility of disappointment. You neglect this because it is a very rare possibility, but also because you do not know what else to do' (1988, p. 97). If we only have the formal ability but not the opportunity to decide whether to accept the danger among alternative options of action we are have to do with confidence. Trust, on the other hand, presumes a more active stance. According to Luhmann, trust deals with risks rather than with danger. If an individual perceives a situation as one of risk rather than of danger then this enables him to address the uncontrollable as controllable (p. 100). One has the opportunity to decide how to deal with the situation rather than the restricted choice to take or leave the danger. The situation of danger turns into an object of action, because one has the option of making an assessment of the danger and of addressing it as a risk that one can analyse and manage. Consequently, one has the opportunity to decide on the best way to deal with the uncertain situation. This in turn entails that the possibility of disappointment, which is implicit in confidence, gets a more prominent position. In a case of confidence 'you will react to disappointment by external attribution,' while in a case of trust 'you will have to consider an internal attribution and eventually regret your trusting choice' since it has been your own choice 'or so it seems if a situation is defined as a situation of trust' (pp. 97-98, 100).

According to Luhmann, trust is not necessarily preferable to confidence. Both can exist simultaneously in a certain domain. Confidence does not rule out trust and vice versa as the example of money illustrates. He shows that one needs confidence in money to participate in the economy. However, it still implies trust with respect to the way you spend or invest your money (p. 98). It is not desirable to turn our confidence in money into trust as we cannot and probably do not want to decide on the risks related to using money. We simply confidently rely on a monetary system. By building a system or a contract we can have confidence rather than that we have to choose ourselves.

In Luhmann's account, confidence turns into trust when individuals get the opportunity to make a choice, to influence a situation, or to decide personally whether or not to accept a situation. In other words, when individuals are able to address a danger as a risk. Elections are put forward as an example of a tool that may convert political

---

[36] See Section 2.8 for an elaboration on the transition of danger into risk.

confidence into political trust: because of such elections it becomes possible (or is seen to be possible) to avoid a relation of confidence (p. 98). One has both the ability and the opportunity to decide whether one can rely on certain politicians. From this view, the enhancement of public participation in the food sector focuses on the shift from confidence to trust. The consumer is provided the opportunity to avoid certain situations and to address the danger as a risk. Consequently, trust may return to confidence if one has the impression that the invitation to public participation is merely a kind of window-dressing and that one does not have a real opportunity to decide personally how to deal with the uncertain situation.

Luhmann's view on the relation between confidence and trust as linked to the relation between danger and risk raises questions. The suggestion is that confidence is accepting a danger because one does not know what else to do than to rely on another, while trust is a way of dealing with risks. This claim, however, is questionable as Luhmann's own example of elections in a democracy can illustrate. An election provides the public with the opportunity to assess a situation, but it hardly changes the danger one is confronted with into risk. The complexity of the issues at stake is so evident that tools such as elections or an official public debate do not really provide the individual the ability to address these issues in terms of risks that one can decide to accept or not. For instance, problems of employment, state finance, global warming, or global trade are mostly so complex that even professionals are still confronted with dangers rather than risks and uncertainties. The recent detection of new microorganisms that cause food-borne illnesses is a clear example that even (scientific) experts are confronted with ignorance and 'known unknowns'[37] next to risks (cf. Wynne, 1992; 2001). My right as citizen to elect someone to deal with these issues on my behalf does not change the fact that I am still confronted with problems that I am not able to address as risks.

Nevertheless, instruments of public participation may turn confidence into trust, though not because the dangers at stake turn into risks. It is adequate to speak about trust, because the instruments of public participation enable the truster to assess the trustworthiness of scientific or political experts. The danger entailed by relying on these experts turns into a personal assessment of the competence and motivation of the trustee. Personally, one is still confronted with a danger, yet considers the experts capable of addressing the issue as risks, or at least as having more competence to assess the danger. Thus, trust is possible because the other agent turns out to be trustworthy, rather than that the danger of, for instance, zoonotic diseases has disappeared.

This indicates that confidence requires other reasons than trust. A social contract, laws, and the structure of policy can provide the reason to be confident. When I have *confidence* in a company, I expect them to live up to their own code of conduct or to the European laws. In contrast, when I *trust*, I do not need information about the system or a

---

[37] "Known unknowns" is a matter of partial or incomplete knowledge.

contract, but about the other agent. The question is no longer whether the agent is part of a social arrangement, but whether the trusted actor is competent and adequately motivated. I expect him to be capable and motivated to respond to this trust. Therefore, I expect that he will not only act in line with what is common or issues from a contract, but also that he will take care of what is entrusted to him, even if this were to require that he had to do something different from the usual. This illustrates the element of motivation. In trusting another agent, I do not only expect him to be competent to act in the expected way, but I also assume that he has the proper motivation to do so.

This element of motivation illustrates why we are hesitant to say that we 'trust' criminals. Suppose that I am regularly confronted with some criminals. In the course of time their doings and internal codes of ´honour´ become predictable for me. Nonetheless, I would not say that I trust them to burgle the university building, although I expect them to do so, given my knowledge about their routines. It is possible to rely on them and even to be confident that they will act in this way, yet I do not trust them since I do not expect them to have a genuine motivation. Deutsch (1958) makes this point implicitly when he argues that it is quite easy to predict, and thus be confident about, how many persons will die on a given day, but that this does not imply that one 'trusts' these people to die. Trust entails an element of motivation.

Baier (1986) defines this element of motivation as goodwill. In her view, trust presupposes that the other is competent enough to act in the expected way and have goodwill towards you as a truster. In the case of the burglars, this seems a proper criterion. You may expect them to burgle the university building, but if they do, they certainly will not do so out of goodwill towards you. Nonetheless, goodwill is a too strong condition for trust (cf. Holton, 1994; Lahno, 2001). First, trust does not necessarily presuppose goodwill towards me as a truster. When I entrust the care of my child to the day nursery, I do not necessarily presuppose that the employees of the crèche have goodwill towards me. Suppose that I even know that they have not. Then it is still possible for me to trust them to take care of my son. This trust is sincere, because I know that they have goodwill towards my child. Therefore, goodwill as an important condition for trust cannot be restricted to the goodwill towards the truster. It also includes goodwill towards the object of trust. However, this interpretation of the motivational element of trust is still too stringent to be a necessary condition. Without explicit knowledge about whether the staff has any goodwill towards me personally or towards my child, it is still possible to trust the crèche rather than to have confidence only.[38] When I know that they are committed to the care of children in general, or by the awareness that parents in general trust them, I expect them to be motivated by this commitment or this awareness when I entrust them with the care of my child. For this kind of trust it is not necessary to

---

[38] It would be a case of confidence if I do not rely on their motivation, but only on the fact that the staff is officially qualified and professionally trained and operates in a context that has been legally regulated.

know whether their motivation is directly related to my trust. The trustee's motivation, apart from underlying social contracts or sanctions, is crucial to defining a situation as one of trust, but need not be directly linked to one specific case of trust.

The element of motivation makes trust a higher degree of reliance than confidence. This affects the degree of certainty that we attach to our expectations when we trust (Misztal, 1996, p. 16). In the case of trust, one has a higher degree of expectation with regard to the other agent.

3.3.5 A map of trust

At the end of this section, it is possible to summarize some characteristics of trust. We have seen that trust can be localised in the range between control and hope. When one has full control over a situation or a person, trust is irrelevant. On the other end of the spectrum, we have indicated hope. Like trust, hope is a way of dealing with situations of uncertainty and lack of control, but it is a vaguer attitude, less based upon the available evidence, a feeling that appears to be difficult to justify on rational grounds.

Within this spectrum other phenomena occur, such as faith and reliance. Both concepts are broader than trust. Faith is broader with respect to the object of trust: it contains most if not all characteristics of an agent, while trust can be indexed to certain objects and specific contexts. Reliance is broader than trust with respect to the subject. One can rely on inanimate entities, parts of nature that cannot but behave in one specific way, or on human beings who are strongly restricted in their freedom. Trust requires that the other party has the ability both to respond to the truster and to refrain from doing what is entrusted. Finally, confidence is closest to trust, but is not exchangeable with it. Trust has a stronger element of choice at the level of the truster. If the truster does not consider having any other option than to rely on one specific agent, we had better describe the situation as one of confidence. Furthermore, the discussion of the relation between trust and confidence has shown the importance of the motivational element in responding to trust and in acting in the expected way. In a situation of trust, the trusted agent should not merely act in a predictable way, but should also be motivated to respond to what is entrusted.

In conclusion, I see four features of trust for the first outline:
- *Uncertainty and lack of control*
  Trust is relevant in cases in which agents lack control and are uncertain about future actions and events because it enables them to act in spite of this predicament. This shows the function of trust. It is a way of 'managing uncertainty' (Becker, 1996, p. 45). In trusting, one acts 'as if' certain possible states of affairs will not occur (Lewis & Weigert, 1985, p. 969).
- *Sincere belief based on assessment*
  This acting 'as if' is not an escape in a make-believe world of certainty and control.

One has the sincere belief that the other party can be trusted. This belief is based on one's assessment of the trusted party (Baker, 1987; Lagerspetz, 1998). In this way, it is different from hope.

- *Trustworthiness and human action*
  The discussion of the relation between trust and reliance shows that trust is an orientation that we have towards (collective) human agents rather than to machines or natural phenomena. Trust is a way to deal with the uncertainty that is the result of human agency rather than with the uncontrollability of natural mechanisms, such as sunrise or a tsunami. Trust is an assessment of the trustworthiness of agents, while reliance can also be based on other reasons, such as predictability or predetermination.
  Reliance can be the result of a calculation; trust, however, includes the ability to 'read' others and to make a decision about whether someone is trustworthy (Eckel & Wilson, 2004, p. 464).

- *Choice and motivation to enter a relationship*
  The discussion of the relation between trust and confidence has shown the importance of the truster's choice to enter a trusting relationship, and of the motivation of the trustee to respond to one's trust. Trust presumes a relationship that includes social interaction between the truster and the trustee.

Based upon these four features it is still possible to refer to trust in very different ways. For instance, it is possible to refer to trust as a matter of taking the risks involved in cooperating with other agents. But within the above mentioned map of trust, it is also possible to define trust in terms of an emotional attitude that makes us perceive the world in a specific way and therefore enables us to cope with the freedom entailed by human agency. In the next sections, these different views on trust are discussed and analysed. I start with the idea of trust as a choice to act (Section 3.4). This discussion is linked to the debate on the relation between trust and risks (Section 3.5). Consequently, I analyse the view on trust as a belief. This raises epistemological questions (Section 3.6). Finally, I discuss the emotional element of this concept (Section 3.7). The aim is to analyse the different views on trust in terms of the question to what extent they fit with the above-mentioned features of trust.

3.4 Trust as a choice to act

The idea of trust as a choice to act is widespread. When we take an action as 'a unit of intentional behaviour that produces expected outcomes' (Creel, 2001, p. 160), it is clear that trust is often referred to as the choice to act in a certain way. Sztompka (1999, pp. 66, 69), for instance, defines trust in terms of a decision, more specifically, the decision of making a bet (pp. 25, 69). He obviously links trust to making a choice with respect to

risk. However, the idea of trust as the choice to act in a particular way is not restricted to authors who refer to this concept in terms of taking risks, as the article by Eckel & Wilson (2004) shows. They infer from their study that 'subjects do not see trust as a problem of risk, but rather as a problem of judgment' (p. 464). Nonetheless, they define trust as the choice to act in a situation of uncertainty rather than as an attitude or a belief. Also the focus in the literature on 'entrusting' entails a moment of action (Baier, 1994; Sevenhuijen, 1999).

The background of the idea of trust as a choice to act is obvious. Trust can be traced by looking at the way people act and behave. When someone is confronted with uncertainty and a lack of control, yet acts nonetheless, it is quite easy to infer that this person trusts the agents on whom he or she has to rely. For instance, if a consumer buys a food product, this can be considered as sign that he must have some confidence in one or more agents in the food sector. Given the uncertainty and lack of control he is confronted with, the individual needs to rely. Thus when he buys food this appears to be a matter of trust. Defining trust in terms of the decision to act, however, raises the problem that only a marginal number of situations of trust will be visible. The focus on trust as the choice to act is problematic in two ways. First, the observed action can mistakenly be ascribed to trust. Trust enables agents to act in cases of uncertainty. This, however, does not entail that all actions in uncertain conditions are signs of trust. Equally, they can be the result of reckless behaviour or of a naïve decision, or there may be certain power relations or implicit sanctions that make the agent decide to act in spite of the uncertainty. For instance, the consumer who buys food might not have any confidence in the competence and motivation of the involved agents with respect to food safety, but as he knows that severe sanctions will prevent the agents from acting contrary to his expectations, he acts in spite of his lack of control. Thus, referring to trust as an act runs the risk of regarding too many situations as results of trust.

Second, trust does not always result in specific actions. One can trust an agent without the need or the possibility to express this trust in a specific action. For example, I can have perfect confidence in my doctor, yet as long as I am not ill my trust will not easily be noticed. Hardin distinguishes between the assessment of someone's trustworthiness and an action that is based on trust (2006, p. 33). Correctly so. For instance, my trust in my doctor is sincere since I consider her to be both competent and properly motivated. However, because I am in good health and have no need to consult her I do not act upon my trust. Therefore, from a third-person perspective it is difficult to discern my trust in my physician. Trust does not always produce (visible) effects. Consequently, trust easily remains unnoticed if the focus of the analysis is only on the choice to act. Baier takes a further step and argues that trust need not be purposive (Baier 1994a [1986], p. 100). Trust need not result in actions, because we do not necessarily trust another agent in order to achieve a specific a goal. I do not trust my physician because I need to consult her. I consult her because I trust her.

Thus, behavioural accounts of trust, i.e. accounts that define trust in terms of acts and behaviour are both too broad and too restricted. The choice to act in a specific way can be seen as evidence of trust, but it cannot be equated with trust. Certain behaviour, such as cooperation, can be the effect rather than the essence of trust. Hence, trust cannot be equated with its visible results. If we equate them, we will end up in a vicious circle: we can only explain the essence of trust by focusing on certain trustful behaviour, but we can only explain this behaviour by focusing on trust. It 'would make the thing we want to explain the explanation of it' (Hardin, 2002a, p. 7).

This does not rule out however that trust structures and guides our actions. Trust certainly steers action and behaviour. The point is that the content of the concept does not coincide with its visible results, as behaviourists would have it.

3.5 Trust as an act: Taking decisions and risks

The content of trust does not coincide with its effects that are visible in our actions, yet one still can argue that the concept can be defined in terms of an action. From this perspective trust is not the act of, for instance, consulting a doctor or buying a product, but trust is itself an act, e.g. the act of entrusting an object to a trustee, or the act of taking a risk.

This raises the question whether trusting is something we can decide to do. At face value, it seems to be possible to decide to trust. For instance, when I am in a foreign country it appears that I can decide to trust a man on the street to sell safe food. Another example is the moment that I decide to trust my child to walk to school alone. In a case of stage diving the artist apparently decides to trust others to catch him. In all situations, there is a moment of deliberation that results in a decision to buy the food, to let the child walk alone, or to jump from the stage. Especially, the characteristic that trust is relevant in cases of uncertainty and lack of control suggests that it entails moments of deliberation that result in a decision. Trust seems to include an assessment of the risks and benefits of acting in spite of the uncertainty. Trusting another involves the risk that the other agent may not act as expected, i.e. one is vulnerable. However, at the same time one may benefit from acting despite the uncertainty. Consequently, it is possible to refer to trust as a decision based on the assessment of the risks and benefits at stake.

The most prominent example of this view is the account of James Coleman, who defines trust as a sub-class of risk taking. To understand his position, we have to take his theoretical background of rational choice into account.

Coleman starts, in line with the rational choice theory, with the individual that he conceives as an atomistic agent for whom interests both indicate the levels of satisfaction and are the driving forces of action. The individual has an almost innate inclination to promote his interests and acts rationally, where acting rationally is understood in terms of utility maximising (1990, pp. 509-511). This individual is neither endowed with altruism or unselfishness nor does he act upon a shared normative system (pp. 5, 31-32). Coleman does not deny that individuals may sometimes have these characteristics or can build on a shared normative system, but his aim is to construct a theory that can explain the realization of norms and altruistic behaviour when it runs counter to private interests. The assumption of a normative system or the assumption that man is endowed with altruism would frustrate this endeavour.

His aim is to analyse how these individual actors pursue their own interests in the complex relationship they have with many others. In this analysis, norms play a role, yet remain implicit. It is at this point of interests and relationships that trust comes into the story. If actors were to have full control over situations, they would not have problems of trust. In those situations 'they merely exercise their control in a way that satisfies their interests' (p. 29). However, it is obvious that this element of control is a problem. If we want to work, to love, to eat, we have to trust others, since there is a 'competition for the resources held by each actor' (p. 131). In these situations, trust is useful, since it enables us to find alternatives and provides opportunities to promote our interests. Accordingly, trust can be best analysed as a strategy of a rational actor who is after optimally achieving and promoting his interests. In other words, it is 'a unilateral transfer by the trustor to the trustee of control over some resources or actions or events, with the expectation that this placement of trust will bring a gain in utility' (1983, p. 92).

The combination of the perception of the individual as both ideally rational and essentially self-regarding makes, according to Coleman, that trust is a matter of rational risk calculation. It turns out that trust itself is a matter of decision making under risk. Therefore, he claims that trust is 'nothing more or less than the considerations a rational actor applies in deciding to place a bet' (p. 99). Trust is referred to as a type of action rather than as an attitude that leads to certain actions. In this approach the idea is that when one consumes food, one makes a calculation of all risks and benefits and evaluates whether it is worthwhile taking the risk of taking the vulnerable position of a truster. The decision to take this risk is one's trust rather than the reasons that underlie this behaviour. Consequently, Coleman draws the conclusion that 'situations of trust constitute a subclass of those involving risk' (1990, p. 91).

From this perspective, trusting is acting and problems of trust are problems of risk. Hence, to ask when it is rational to trust is similar to the question of when it is rational to take a risk. In each situation of trust, the truster has to make a risk analysis by calculating

the possible gains and losses of the two possible options: to trust or not to trust.[39] In an ideal situation, this calculation is not that complex: you know your beliefs and interests, and you know whether an action will contribute to these interests, and thus you have to decide for the option in which the total sum of gain outweighs the loss. In line with this argument, Coleman argues that one has a reason to trust if the ratio of the chance of receiving gain[40] (p) to the chance of loss is greater than the ratio of the amount of the potential loss (L) to the amount of potential gain (G). Thus, one has a reason to trust if:

$$\frac{p}{1-p} > \frac{L}{G}$$

(1990, p. 99).

It is important to note that for the result the dimension of the gain and loss does not really matter: the combination of a small loss and a small gain can be defined as rational trust just like a high loss in combination with a high gain. In both situations, trust can be framed in terms of taking a risk. This assessment, however, is complicated by the problem of uncertainty. One entrusts a certain object of trust to the other at t=0 but the reaction that proves whether the trust has been well placed will only occur at t=1. Therefore, Sztompka's description of trust as placing a bet seems also accurate for Coleman's account (1999, p. 25). The only thing a truster can do is to assess the probabilities with respect to the risks and benefits of either trusting or non-trusting, and deciding whether it is worthwhile placing the bet, i.e. taking the risk of relying on another. This element of time and the definition of trust in terms of placing a bet also explain why Coleman stresses that the real problem in deciding to trust another agent is the insufficiency of information. The element of information is essential on different levels. First, it can provide better insight in the possible losses and gains of placing trust. Second, the more you know about the other person or institution, the better you are able to assess the likeliness that the agent will act in the favourable way. This will improve your ability to formulate expectations towards another person or organisation. Therefore, we need to search for the most accurate information. However, this need is equally dependent on a calculation of gains and losses: 'The search should continue so long as the cost of the additional increment of information is less than the benefit it is expected to bring' (1990, p. 104).

Next to the element of information, the aspect of rationality may help the individual to decide whether it is worthwhile to take the risk of trusting. Coleman defines trust as

---

[39] Coleman (1983) underlines that in this calculation the expected gain and expected loss are often measured relative to the present state, but that it is also possible that the expected values of different alternative are measured relative to one another, p. 92.

[40] Coleman also speaks about the probability that the trustee is trustworthy (1990, p. 99).

rational when it is purposive in the sense that it contributes to maximising one's interests.[41] In the calculation that precedes a decision to trust only the rational pursuit of preferences counts. Other aspects, like norms or social conventions, are only relevant as far as they contribute to this rational pursuit. For instance, a reputation that the trustee is loyal and dedicated does not have an inherent value in this model. It certainly is considered as valuable, but only as far as it contributes to one's preferences. A good reputation is relevant since it enables you as a trustee to cooperate easier with others, which help you to maximise your preferences. When you are in the position of a truster, a good reputation of the trustee is important too, since it can help in assessing the involved interests, and it can serve as a sanction when the trustee betrays your trust. One knows that breaking trust would have negative consequences for the other party as well, which leads to a 'negative sanction' (p. 115) that can directly influence the calculation. However, from both perspectives it is only relevant because it can serve as direct input in the risk–benefit calculation that underlies the decision to trust. The idea that a good reputation is inherently valuable because it is a sign of virtuous behaviour or the result of following an ethical code of practice does not fit within this calculation model in another way than as a possible gain or loss within the calculation. This shows that every element in a problem of trust can be modelled in terms of a factor in the risk–benefit calculation. The only way to steer the processes of coming to trust and behaving in a trustworthy manner is to introduce new sanctions or rewards that fit the agent's preferences.

In short, within this rational choice model the question of trust can be equated with the question of whether I have reasons to act, i.e. to take a risk given my rational pursuit of preferences.

3.5.2. The conceptual problems of trust as the act of risk taking

Prima facie, trust defined in terms of taking a risk has certain attractiveness. As mentioned before, trust is relevant in case of uncertainty and lack of control. Hence, trust is always related to situations in which one runs a risk to be harmed by the trustee, although the attitude of trust explicates that the truster does not deliberately take that risk, but is confident that the other person will not let him down. Consequently, it is rather easy to reframe situations of trust into a process of weighing risks and benefits. If we do so, it is possible to strip trust from its elusive character. Although the above analysis shows that both risks and benefits are complex issues, it is often easier to say

---

[41] At this level, Coleman's main endeavour is descriptive; he aims to explain people's behaviour. A more prescriptive element surfaces when the element of trust as placing a bet is combined with the view on rationality as choosing that action that contributes optimally to the utility function. From these claims it is possible to draw the conclusion that if trusting in the sense of placing a bet contributes to the overall utility and the agent is rational, it is not only likely that he will trust the other party, but he also that will place the bet (p. 99).

something about risks and benefits than about trust. This makes accounts such as Coleman's applicable in a wide range of situations and in a rather consistent way. Both personal trust between neighbours and public trust in the government can be analysed as a risk–benefit calculation that results in a decision to trust. Furthermore, the account is in line with the second characteristic of trust: the view that it is a sincere belief based on assessment rather than an escape in a make-believe world of certainty and control. Trust is not just a step in the dark, but the result of a calculation of risks and benefits.

This emphasis on risks and benefits, however, also implies some serious problems. First, in Coleman's account it is hardly possible to distinguish between trust and reliance. Above, however, I have shown that trust is a more restricted category than reliance. We can have reliance in a much broader category of entities. Furthermore, reliance can have its origin in different mechanisms, such as predictability or the availability of standard operating procedures. Trust, however, is based on an assessment of the trustworthiness of one or more agents. By making the risk–benefit calculation central to the process of coming to trust it is hardly possible to differentiate between 'relying on' an agent and 'trusting' him, since both are reframed in terms of ways of taking decisions under risk. The problem is that the assessment of trustworthiness is no longer the central issue. Central is the assessment of risks and benefits. The ability to 'read' others and to make a decision about whether someone is trustworthy is substituted by the ability to make a risk calculus. In this process, the motivation of the trustee plays only a marginal role. Consequently, there is no opportunity to differentiate between trust and reliance. Even if one does not have any idea about the motivation of the trustee, but has a clear indication that he will act in the expected way, the calculation may show that the risk is worthwhile taking and thus one 'trusts' him. This highlights that the emphasis is on risk taking rather than on assessing trustworthiness.

Another effect is that it does not provide any room for levels of trust. We trust some people or institutions more than others. However, in the calculation model there is no room for these differences. The calculation of gains and losses only tells you whether it is rational to take the risk, i.e. to trust or not, but does not provide a clear indication of the level of trust.

The risk focus does not only devaluate the importance of the motivational aspect in responding to someone's trust; other issues appear to be less important too. Trust as acting under risk based on a calculation entails that every element of trust has to be modelled in terms of a factor in the risk–benefit calculation. Consequently, the characteristics that do not fit in with the risk focus have to be left out of consideration or framed in terms of a risk or a benefit. Hence, aspects such as the affective element and psychological elements like the influence of basic trust or generalized trust (Uslaner, 2002) or ethical considerations hardly fit in with the risk frame as independent aspects.

The inability to distinguish between trust and reliance is not an unintended flaw in Coleman's account. It is the direct result of his analysis of rationality and individual

goal-oriented behaviour. The truster is first and foremost an individual who aims to pursue certain goals and has to rely on others in order to do so. Thus, he has to take certain risks that he has to evaluate in the light of the possible benefits. The individual and his goals and preferences are central rather than the social interaction with the trustee. The social interaction between truster and trustee, which I have presented as one of the features of trust, can be entirely absent in this account. Consequently, the incentives of the trustee are less relevant. Primarily the truster is less interested in why a trustee will act in the expected way. In bold terms: the trustee is just an obstacle that entails risks which have to be taken in order to achieve the individual's aim.

This view on rationality and agency is a necessary element to the risk focus on trust. The calculation of risks and benefits itself does not yet explain why someone takes the risk of trusting (cf. Barrera, 2005). To understand why one takes the risk one needs assumptions that are beyond the strict field of risk. As I have shown, Coleman's analysis contains crucial assumptions about human nature and rationality which keep other persons and institutions accountable, such as the premise that the other agent aims to maximise his preferences and that in striving for maximisation he will by guided by the function of utility. Such assumptions do provide the truster with a tool to assess whether it is worthwhile taking a risk. Furthermore, they provide information about what we may expect of a rational agent. For instance, the link between acting rationally and the striving for utility makes the rational individual more predictable. If we have to rely on him, we can predict what option he will choose, provided that we have the same information as he has, since we know his goal and how his preferences are ordered. The truster knows something about the motivation of the trustee to act in the favourable way. However, the incentives to act in a trustworthy manner are not the result of an expected response to one's trust. The incentives are rather that one can expect that a rational actor will be motivated to act in this way, given the specific context of preferences, information and goals. The addition of these substantive assumptions makes the risk focus widely applicable.

For the moment I leave the discussion of the fundamental assumptions on individuality, rationality and agency[42] and focus on the relation between trust and the act of risk taking.

3.5.3 Trust and risk taking: Complementary but fundamentally different

Since trust is related to situations of uncertainty and lack of control, it is a small step to argue that trust is a matter of risk. Even though not all authors follow Coleman in defining trust in terms of taking a risk, trust is often discussed in terms of risk and

---

[42] These issues will be discussed in further detail in Chapter 5.

uncertainty. Many authors have stressed the connection between them. When I trust someone I always run the risk that my trust will be abused, or in Hume's words, in a case of trust 'Tis impossible to separate the change of good from the risk of ill.'[43] Thus, it is not surprising that many scholars consider trust as a risky matter (cf. Gambetta, 1988, p. 235). Baier does not speak about risks, but clearly emphasises the vulnerability in trusting relations (1994a [1986]). Luhmann refers to trust as a venture, since it is 'a purely *internal* calculation of *external* conditions which create risk' (Luhmann, 2000, p. 31; 1988, p. 100). This leads Sztompka to the conclusion that trust 'copes with one type of risk by trading it for another type of risk' (1999, pp. 25, 32). Trust and risk are evidently related, yet their precise relation is far from clear.

A brief historical view on both concepts illustrates that trust and risk are at least no synonyms. The word risk 'derives from the early Italian word *risicare*, which means 'to dare'. In this sense risk is a choice rather than a fate (Bernstein, 1998, p. 8). Therefore, in the view of trust as the choice to take a risk, we are close to the original meaning of risk: one dares to do something. The term most commonly used for trust in Italian and Latin is *fede*. *Fede* is also the term for faith. Even though we have already noticed in Section 3.3 that not all trust can be equated with faith, this etymological link shows that trust is different from daring. As having faith in God is not a matter of fearlessness, so trust is not a matter of having the courage to take a risk based upon a calculation. Trust is an attitude that is based upon a belief about the trustee. This belief rather than his courage or our willingness to take risks enables the truster to deal with uncertainty.

Moreover, not only the historical context suggests differences between the two concepts, also in empirical research evidence has been found that trusting decisions are not always thought of as risky gambles (cf. Eckel & Wilson, 2004; Schechter, 2006). Eckel and Wilson infer from their study that 'subjects do not see trust as a problem of risk, but rather as a problem of judgment.... Although people may vary in their capacity to make accurate judgments, the choice to trust appears to be one of conditional judgment, not a calculated financial gamble' (p. 464). From another disciplinary background, Lagerspetz (1998) has drawn a similar conclusion. In his philosophical analysis he underscores that trust is something fundamentally different from risk taking. He does not deny that trusting is important in cases of uncertainty and that entrusting certain objects to someone else may be related to risk, yet he emphasizes that trusting is not the same as deliberately taking a risk. He states that 'considerations about risk taking can only motivate risk taking, not trusting' (p. 56).

This confronts us with a dilemma. On the one hand, when a person trusts he is in a vulnerable position. Thus, trust appears to be a risky matter and a venture, in the words of Luhmann, *'ein Wagnis'* (Gambetta, 1988, p. 235; Luhmann, 2000, p. 31). On the other hand, Lagerspetz's view can easily be recognised as well. When a person trusts, it will

---

[43] Quoted in Hardin (2002, p. 120).

appear that he does not perceive the situation as risky or as a gamble, although he certainly will run a risk. This, however, is not a real contradiction. The problem is that the authors have a different focus, which shows the importance of a distinction between first-person and third-person perspectives. From a third-person perspective, trust is certainly a risky matter: A truster takes a risk. In acting as if only one state of affairs were to be expected, one runs a risk and makes trust close to a gamble. Nevertheless, from a first-person perspective the picture is quite different. As a truster, one is not aware of taking this risk. If he was he would be a risk taker, not a truster. If you trust a colleague to write a report, you know that she has the ability, and maybe even good reasons not to write it. Nevertheless, when you trust her, you do not make a risk analysis and consequently take the risk, but judge her to be competent and properly motivated with regard to this assignment. From the perspective of the truster, risks are not the main element of trust. Only as an observer one may notice that the truster runs a risk. An example can clarify this point. In the case of faith, believers can sometimes take considerable risks – from a third-person perspective – without any sense of uncertainty. This is not the result of different views on risk management, but because they consider God as ultimately trustworthy. The individual believer may not be able to assess the risks involved in relying on God, but that is not the point. Because he trusts he is certain about his acts. Only if he were to lack such trust would he consider his way of acting as (too) risky and uncertain.

The distinction of perspectives can also shed light on a paradoxical element in Luhmann's analysis. He argues on the one hand that trust is risky, but on the other hand defines trust in terms of a phenomenon that reduces complexity (2000, p. 30). If one were to agree that trust is risky, one would take a third-person perspective. However, from this perspective the last element is questionable: the fact that one trusts does not really change the possible number of actions or reduce complexity. The world remains as complex as it was before. The reduction of complexity is, as Luhmann agrees, a reduction by a subjective process (p. 32). This clarifies a change of perspective that Luhmann applies. Only from a first-person perspective is trust a way to act in spite of complexity and uncertainty. Therefore, it 'reduces' complexity in the sense that it changes the truster's perception of complexity. Yet, from this perspective, trust is not a matter of risk, but of trustworthiness. Only if the potential truster considers somebody trustworthy, trust can serve as a mechanism that allows people to cope with complexity and uncertainty (*cf.* Giddens, 1991; Lagerspetz, 1998; Lahno, 2001). Hence, Luhmann's combination of trust as a risky matter and as a phenomenon that reduces complexity is the result of an implicit change in perspective. Trust conceived as a risky matter starts from an observer perspective, while trust as a complexity-reducing phenomenon suggests a first-person perspective.

This has direct consequences for the idea of trust as the act of risk taking. It has been argued that trust modelled in this way only allows that trust is either nothing but correct

risk analysis, when enough information is available, or just a step in the dark when one really is confronted with uncertainty and serious lack of information. Yet, in the first situation one does not need to trust, since 'trust begins where prediction ends' (Lewis & Weigert, 1985, p. 976). In the second situation, the question is whether one really trusts or only naively or under pressure takes a risk that cannot be valued properly. Becker accurately claims that in both situations we have eliminated trust. He argues that with regard to such a cognitive account trust is no more than a problem of knowledge and of power (1996, p. 49-50).[44] In contrast to a risk taker, a truster is not calculating but coping with complexity. He is not calculating risks, but dealing with the uncertainty he is faced with. Trust is not something that you decide on with a pc.

Trust is part of a relationship built up through time, rather than an act in which pros and cons are weighed. It is based on the personal assessment of the trustworthiness of the trustee. This, however, suggests opening the door once again for calculations. Trust may not be a matter of risk taking, but can still be modelled in terms of probability, for in every subjective assessment of trustworthiness there remains the probability that the expectations are not met. Therefore, trust can be considered as a 'particular level of the subjective probability with which an agent assesses that another agent or group of agents will perform a particular action' (Gambetta, 1988, p. 217; also Dasgupta, 1988). I do not deny this aspect, but this personal assessment is different from a calculation of risks and benefits. First, because considerations about probability of risks and alternatives are insufficient to motivate trusting, as Lagerspetz has made clear. Secondly, the focus on prediction and assessment underestimates the fundamental uncertainty that one is confronted with. The reason that we have to rely on other individuals or institutions is that we lack the knowledge about all possible outcomes of a certain choice and all alternatives from which we can choose (cf. Six, 2004; Nooteboom, 2002). Hence, we do not – and mostly even cannot – predict in trusting others, but we do formulate positive expectations.

I conclude that the acts of taking risks and trusting are two complementary, yet different mechanisms to deal with cooperating with other agents in cases of uncertainty. A risk approach aims to clarify the uncertain aspects of the situation in which one has to rely on another agent. In this context a risk–benefit analysis provides me with tools to assess this probability and to evaluate the hazard. In this way it helps to translate the problem of known uncertainty into one of risk. Consequently, I can make my personal assessment and decide whether it is worth to take the risk entailed by relying on another

---

[44] Becker mainly focuses his critique on the 1993 article by Hardin on trust and street-epistemology. Although I think that his general arguments against a cognitive approach still hold for Hardin's approach, this specific argument is more relevant in the case of Coleman's theory than with regard to the more recent work of Hardin (2002; 2004). At least the suggestion that trust can be equated with power in the case of Hardin's theory seems a little out of place since Hardin also argues that trust under pressure is no real trust (*cf.* Hardin, 2002a, p. 12).

agent. The ultimate aim is to enable the individual who is confronted with uncertainty to calculate whether it is worthwhile taking the risk given one's own interests and preferences. A risk–benefit analysis does not provide us with reasons to trust, but can show that we are confronted with a risk although the situation at first hand appeared as uncertain. Or it will show that the risk at stake is worthwhile taking, given our preferences. Then I can choose to take 'the risk of cooperating with you on some matter even if I do not trust you' (Hardin, 2002a, p. 11).

A trust approach to uncertainty, on the other hand, has a different focus and starts where a risk approach ends. It focuses on those situations that remain uncertain even after the uncertain aspects have turned into risk factors as much as possible. The aim is not to try to make a risk–benefit analysis, but to personally assess the competence and motivation of the trusted agent. When I trust a company that offers novel foods, I do not assess risks, but make an assessment of the company's competence and motivation with respect to this new food product. In this process of assessing the competence and motivation, risk information may contribute to trust only as far as it serves as a signal or proof of the competence and motivation of the trustee. Suppose that the risk at stake is very low or the risk information is of a high quality, yet I do not consider the other agent, on whom I have to rely, to be competent or to have any good-will to me. Then I will not trust him. Maybe I consider it worthwhile taking the risk, and act nonetheless. Then I take a risk, yet do not trust. This explains why risk information or improved risk analysis does not directly influence the level of trust. It only has direct influence on a mechanism to cope with uncertainty that is different from trust. In building trust, risk information only plays a secondary role.

3.5.4 Deciding to trust?

Trust is different from the act of risk taking. However, the more general question whether we can decide to trust, which is a precondition of defining it as an act, has remained unanswered. At the start of this section, I have provided three examples that cannot all be reduced to terms of risk taking. Especially the example of trusting your child to walk to school alone is not just a way of taking risks. Something else is at stake. In this case, it seems that I really can *decide* not to accompany him, but to trust him. This is what Sztompka calls 'evocative trust' (1999, p. 27-29). The aim is to evoke trust in order to show the child the importance of trust in human agency. Primarily it has an educational purpose: we trust to evoke trust, i.e. to teach the child the essence of trusting and of taking care. Horsburgh (1960) argues that this is not restricted to trust in children, but can be applicable in many relationships in which the aim is to increase the trustworthiness of the trustee. He defines this as 'therapeutic trust'.

This situation of evocative trust seems to affirm Holton's claim that, although not all trust is the result of a decision, 'in *some* circumstances we can decide to trust.… We are

constrained in the trust that we can extend; but within the constraints there is some room for choice' (1994, p. 68). The question is whether this idea of deciding to trust fits the four mentioned characteristics of trust. I do not think so. In the case of evocative trust, I make a decision, yet not one of trust. My decision is not based upon an assessment of trustworthiness, but upon my evaluation of the importance that my son knows what it implies to be trusted. I may not really consider him fully competent, but by entrusting him, I hope he will learn that trust is not self-evident. I decide to trust him, or maybe more accurately, to act *as if* I trust him. I treat him as someone who can be trusted, but his actual trustworthiness is of less importance. Consequently, I will not feel betrayed when he disappoints me. I may blame myself for not assessing his competence properly, but will not have feelings of resentment towards him. I make a decision, but what is at stake is a decision to use a pedagogical tool rather than to trust.

Nevertheless, it seems that there are other situations in which one can decide to trust. Holton starts his article with the example of participating in a game during a drama course. As part of the game you are blindfolded and stand in the middle of a circle formed by the others.

> They turn you round till you lose you bearings. And then, with your arms by your sides and your legs straight, you let yourself fall. You let yourself fall because the others will catch you. Or at least that is what they told you they would do. You do not know that they will. You let yourself fall because you trust them to catch you. If you are like me, there is a moment at which you weigh up whether or not to let yourself fall. How does it feel at that moment? It feels as though you are deciding whether or not to trust. (1994, p. 63)

It appears that you make a decision based upon deliberation, but is this really a case of trust? On the one hand, it fits in with some basic characteristics of trust, since it is a situation of uncertainty that is related to human agency, and it is more than just hope. On the other hand, if you really make a decision, I doubt whether you consider the others as trustworthy. Rather, it seems to be just a case of reliance or risk taking. For instance, you may decide to let yourself fall since you consider it the only way to complete this course and evaluate this benefit as important enough to take the risk. Or you rely on your classmates since you expect them to catch you because it is part of the game. Whatever the reason may be for letting yourself fall, it is not because you assess them as trustworthy. In Holton's example, it is either the decision to rely on the other participants or to take the risk at stake and thus you act *as if* you trust, yet then the decision is not with respect to the other person's competence and motivation.[45] Or you can decide to act *out of* trust. Then, however, you do still not decide to trust, but decide to act in a trustful

---

[45] Holton also seems to broaden the view on trust to reliance. After the claim that we can decide to trust, he continues that 'the idea of deciding to *rely* presents few problems' (1994, p. 68, my emphasis).

way since you already trust these others or have confidence in classmates in general. The presence of trust leads you to make certain decisions and to act in a certain way.

If we consider trust as intimately related to the assessment of the trustee's competence and motivation, it cannot be a *decision* to start trusting (cf. Lahno, 2002a, pp. 401-2). To be sure, trust can be based on decisions, for instance the decision to consider certain information as relevant evidence in the assessment of trustworthiness. Yet trust cannot be decided to, since it cannot be willed (Baier, 1994a [1986]; Jones, 1996; Lahno, 2002a). We can have the will to trust, but we cannot trust at will. We cannot simply decide to trust since we want to or are invited to. Baier elaborates on this latter situation. She writes, '"Trust me!" is for most of us an invitation which we cannot accept at will – either we do already trust the one who says it, in which case it serves at best as reassurance, or it is properly responded to with, "Why should and how can I, until I have cause to?"' (1994a [1986], p. 110).

Frankfurt's analysis of love at this point may equally hold for trust. With respect to love he argues, that it is 'involuntary' in the sense that it is not under the immediate control of the will. 'We cannot love – or stop loving – merely by deciding to do so' (Frankfurt, 2005, p. 195). The claim still holds if we replace trust for love in this proposition. We do not have a switch by which we can turn trust on or off and to decide to trust another agent. One only comes to trust if one has reason to think that the other person is actually trustworthy, since 'we can't just decide to believe regardless of the evidence' (Jones, 1996, p. 16).

To conclude: one cannot decide to trust and therefore trust is different from an action. In the process of coming to trust decisions certainly can play a role, and trust obviously results in actions, but trust itself is not an action that one can decide to perform. With this conclusion, two ways of conceptualising trust are still possible: trust as a belief and trust as an emotional attitude.

3.6 Trust as a belief and epistemological issues

The conclusion that we cannot decide to trust can be interpreted as a first signal that trust had best be analysed as a belief, i.e. as mental conviction regarding the truth or validity of something. As Cohen (1989) argues, 'belief is not normally achieved at will because it is caused in each kind of case by something independent of the believer's immediate choice: factual beliefs are the believer's willy-nilly feelings about physical or mental reality, moral beliefs are the welcome or unwelcome dictates of his conscience, intuitive beliefs are the immediate, unreflective, and untutored deliverances of his intellect, and so on' (p. 370). Consequently, trust defined as a belief is never something we can decide to have. Even if I have the mental capacity and the willingness to believe something, e.g. that an institution is honest, I am not able to decide that I believe this. However, Cohen

remarkably claims that it is nevertheless possible to decide to believe. In fact, his examples to substantiate this statement are related to trust. He argues, 'You may decide to believe a friend, that is, to trust his word.… You may even decide to believe in him, that is, to have confidence in his abilities' (p. 369). He distinguishes between beliefs, such as the belief that it will be sunny tomorrow, and the beliefs entailed by having confidence in someone's abilities. According to Cohen, we cannot decide to believe that tomorrow will be a sunny day, while we appear to be able to decide to believe a friend or believe in his abilities.

Cohen explains this difference by distinguishing between believing that *p* and accepting that *p*. Even though you cannot decide to believe that it will sunny tomorrow, or to believe that it will not, 'you can decide … to accept that it will, or to accept that it will not: the belief may then ensue, but it may not' (p. 370). This distinction is illuminating; it is questionable however whether it is applicable to trust as Cohen's examples suggest. When you really trust your friend's word, you believe him to be trustworthy, i.e. you have the mental *conviction* that he is competent and properly motivated. This, however, does not necessarily include an explicit decisional moment of acceptance. It includes the acceptance of certain evidence as relevant for the assessment of trustworthiness, but this acceptance is not yet trust.

Hence, in this section 'trust as a belief' is referred to as the mental conviction with respect to the trustworthiness of the other agent rather than the acceptance of the belief. It is this mental conviction that distinguishes trust from hope, because it entails that the acting 'as if' is not an escape in a make-believe world of certainty and control. One has the sincere belief that the other party can be trusted, based on one's assessment of the trusted party. For this reason several authors conclude that trust is in the same category as beliefs. Hardin, for instance, argues that trust is very similar to belief. According to him, 'trust is in the cognitive category with knowledge and belief.' Consequently, to say I trust you 'is to say nothing more than that I know or believe certain things about you – generally things about your incentives or other reasons to live up to my trust, to be trustworthy to me.' A truster does not choose between given alternatives and thus act, but her trust may be grounded in her belief 'in your morality or reciprocity or self-interest' (2002, pp. 7, 10, 13). Pettit also argues that 'trust materializes reliably among people to the extent that they have beliefs about one another' (1995, p. 202), and Gambetta stresses 'the importance of the beliefs we hold about others' (1988b, p. 213). It is obvious that these statements are not about the importance of beliefs in general, but about beliefs with respect to the trustee, the situation and the possible course of interaction. These beliefs can result in a positive expectation towards another agent.

From the perspective of trust as a belief, i.e. a mental conviction, evidence is a central issue. 'What is sensible for a given individual to expect depends heavily on what that individual knows, both about the past and the future of the person or other party to be trusted' (Hardin, 1993, p. 525). According to Hardin trust necessarily entails that we

have some presumption or knowledge about the trustee (1998, p. 11; 1993). We need evidence about the incentives of the trustee to act in the expected way. On the basis of this information I can form a belief about the trustee's competence and motivation. This starts from the idea that trust can be evidence based. In trusting others we sometimes have obvious indications that the person or institution has the adequate competence and motivation that makes him trustworthy. Suppose a consumer trusts science and industry to develop novel food products that are beneficial to health. When she is asked why she trusts these institutions, she can refer to the evidence, e.g. her own experiences or external sources, such as newspapers, the internet or hearsay. This evidence with respect to the competence and motivation of the trusted agents is sufficient reason to formulate positive expectations with respect to new products. However, this does not entail that she knows that the institutions are trustworthy, since evidence can be deceptive, inadequate or wrongly interpreted. Even though her belief is based on the available evidence, it can be false. Nevertheless, even when the evidence is objectively false the truster still can think that her belief is true (cf. Creel, 2001, p. 106).

Moreover, the relation between the available evidence and someone's belief concerning the trustworthiness of the other agent is not always clear. The accessible information does not necessarily lead to trust. Suppose that (a) all the available information indicates that a company is competent and sincerely concerned with its consumers, (b) I recognize this as relevant information and therefore accept it as evidence of the trustworthiness of this company, and (c) I would like to trust this company (e.g. because they sell interesting products). Even then, it is possible that I do not believe them to be trustworthy and thus do not trust them. This situation can be explained in two ways. First, the available information is interpreted in the light of one's personal and subjective question of whether the agent is trustworthy rather than analysed with the aim to come to justified true beliefs. Secondly, it indicates a leap between the acknowledgement of the accessible information as relevant evidence on the trustworthiness of the other agent and my trust in this person or institution. Both explanations are discussed in further detail below.

3.6.1 Evidence and a specific epistemology

Evidence facilitates trust. Without any evidence about the other agent we can hope or gamble that he will act in the favourable way but we cannot trust him. We need information. However, trust as a belief does not yet mean that it is a matter of knowledge. To speak about knowledge implies that four conditions be met: (1) you must think that a specific proposition is true; (2) that proposition must be true; (3) you must have reasons, which prove that your proposition is true; (4) you must understand how those reasons prove that that proposition is true (Creel, 2001, p. 109).

Trust as a belief certainly meets the first condition. When you have the belief that the other agent is trustworthy you think this to be true. This equally holds for the third condition: you have reasons why you think that your belief is true. The second condition, however, often appears to be beyond the focus of a truster. To trust one's beliefs does not necessarily need to meet the criterion that the proposition is in fact true. If they do meet this criterion trust can be strengthened but it never is a necessary condition. Therefore, the positive beliefs of trusting are to be distinguished from knowledge. Consequently, in a case of trust the fourth condition is not relevant. With respect to knowledge you must understand how reasons prove that the proposition *is factually true*, yet in a case of trust the focus is on the reasons that prove why you *think* it is true. It can be sufficient when the truster thinks that his reasons prove the belief to be true. Formulated in this way this fourth condition coincides with the third.

This illustrates that the epistemology that is needed in the case of trust does not and need not to meet the standards of traditional epistemology aimed at justified true beliefs. Trust requires an epistemology that leads to 'knowledge' that provides reasonable evidence about the competence and motivation of the other agent. The reasonableness however is defined by the truster, not by external or universal criteria. When the knowledge that leads to trust can be justified in terms of universal criteria it may be an extra motivation to trust, yet it is not a necessity. According to Hardin, 'the knowledge at issue … is that of the potential truster, not that of the theorist or social scientist who observes or analyzes trust' (2002a, p. 13). From an outsider perspective, it is possible to assess trust based on the criteria of standard epistemology and qualify a specific case as blind or naïve, or distrust as disproportional, yet even in these situations the individual sincerely believes that he has good reasons to keep trusting or distrusting. As Hardin emphasises, trusters are not concerned with the core moment of standard epistemology, defined as the justification of true beliefs based on public criteria. A truster does not primarily search for evidence that is justifiable on public criteria; yet is after an answer to the question whether the other person on whom he has to rely is sufficiently competent and motivated to be trusted. To answer this question he needs evidence. However, a truster wants to 'use' the evidence, 'not to verify or justify it' (Hardin, 2003, p. 17).

This pragmatic focus in the search for evidence demands a specific epistemology. In Hardin's words, it asks for a subjective and pragmatic 'street-level epistemology' rather than an epistemology that focuses on justification, seeks for knowledge and deals with truth claims (1993; 2002a; 2002b; 2003; 2004).[46] A truster certainly will check the value

---

[46] Hardin applies his idea of a street-level epistemology to all fields of epistemology. He claims 'Philosophers might enjoy the enterprise of standard epistemology, the enterprise of the justification of knowledge, but ordinary people can rightly forgo it for most of their own lives' (2004a). I will not go into this broader claim, but agree with Hardin as far as it concerns situations of trust.

of the information he has obtained, yet this check does not need to meet external or public criteria of justification. Hardin stresses that the epistemic search is focused on usefulness rather than justification.[47] If information provides an answer it is relevant. It is possible that some information is accepted by a truster as evidence for the trustworthiness of another agent, since it appears to him as a useful answer to the question of trust, even though it cannot be justified as true knowledge by any of the criteria of a standard epistemology. Obtaining knowledge with respect to trust is 'not simply a matter of analyzing the given information to get to some factual expectation on what will happen or has happened. It rather amounts to evaluating information in a certain way and asking certain questions' (Lahno, 2001, p. 178).

It is evident that there are differences between persons with respect to the amount of knowledge they need and the extent to which they accept knowledge as evidence for the trustworthiness of another. This depends on their assessment of knowledge as a useful answer to the question of trustworthiness. This makes clear why we lack an objective amount or quality of evidence that is both necessary and sufficient for someone to come to trust. And it explains why the same information does not necessarily result in the same level of trust for all trusters. For example, when I trust a company, I sincerely *think* that my beliefs about the trustworthiness of that company are true. I can even have reasons for this conviction based on my assessment of the available information. However, justification of the validity of my belief in terms of public criteria will only be an additional motivation to trust but is never a necessary condition.

Furthermore, the pragmatic focus of the epistemic process explains why providing as much information as possible is not the best way to address problems of trust. O'Neill notes that:

> We are flooded with information about government departments and government policies, about public opinion and public debate, about school, hospital and university league tables…. If making *more* information about *more* public policies and institutions and professionals *more* widely and freely available is the key to building trust, we must be well on the high road towards an ever more trusting society.
> (2002, p. 65-67, original emphasis)

It is obvious that this last conclusion is only rhetorical. In spite of all the available information we can easily access via various kinds of media, the level of public trust in the food sector did not increase in a proportional way. To enable us to trust an agent it is not necessary that this person or institution burdens us with full disclosure of all possible relevant information. Trust does not presuppose that we need to know everything about

---

[47] Hardin defines this epistemology as an 'economic' theory. This implies that the costs and benefits of the individual 'believer' or 'knower' of gaining knowledge is central in this theory rather than the justification of one's belief. Furthermore, he emphasises that, in contrast to rational expectations theory and much of game theory, his economic theory does not presume full knowledge (2002b, p. 215).

the agent to be trusted. We can trust friends, colleagues and neighbours 'whole-heartedly, without any wish, or need, to know everything about their private lives – or to have them know everything about mine' (O'Neill, 2002 p. 69). What we need is information that provides an answer to our questions with respect to the trustworthiness of the other agents. This illustrates that we do not need 'just' information. We place and refuse trust 'not because we have torrents of information … but because we can trace specific bits of information and specific undertakings to particular sources on whose veracity and reliability we can run some checks' (O'Neill, p. 76). Trust requires the ability to check one another's claims, which is more than a matter of checking as many sources of information as possible. For instance, when meetings of a Food Safety Agency are open to the public this does not automatically build trust.[48] This kind of openness and the public character of a meeting are only relevant with respect to trust if they offer the opportunity to check the agency's competence and motivation.

3.6.2 A leap between trust and evidence

Next to the pragmatic character of the epistemology a truster uses to obtain evidence about the trustee, it also has been stressed that evidence is a necessary, but never a sufficient condition for trust. Lewis and Weigert write:

> No matter how much additional knowledge of an object we may gain … such knowledge alone can never cause us to trust …, The cognitive element in trust is characterized by a cognitive 'leap' beyond the expectations that reason and experience alone would warrant – they simply serve as the platform from which the leap is made. (1985, p. 970)

There is a leap between the interpretation of the accessible information on the trustworthiness of the other agent and my trust in this person or institution. As Gambetta noticed, 'if evidence could solve the problem of trust, then trust would not be a problem at all' (1988, p. 233). The problem is more complex. Confronted with information a truster is able to decide whether this is relevant, i.e. whether the available information counts as evidence rather than as mere data. However, there is 'something else' between the recognition of information as relevant for the assessment of trustworthiness and the actual trust in a person or institution. There is more than the evidence at stake. There is an element in the process of coming to trust that 'happens to us' rather than that we decide to adopt a stance of trust.

At the start of the 20th century Georg Simmel was the first sociologist who stressed that there is 'a much weaker link between the identifiable bases of trust and the actual expectations that human beings have when they reach the state of trust' (Möllering,

---

[48] In their empirical research on trust in GM-food products, Poortinga & Pidgeon concluded that involving the public in the debate would not automatically improve the level of trust (2004, p. 1485).

2001, p. 404). He defined the gap between the interpretation of the facts and the expectations that are the result of trust as a process of 'suspension'.[49] It is a suspension of one's hesitation and doubt with regard to another person or institution. Although one can be hesitant to rely on others for many reasons and have serious doubts, it is possible to suspend these doubts and come to 'a hypothesis certain enough for practical conduct' (1950, p. 318) and consequently come to trust. He identified in this leap from interpretation to expectation a 'mysterious further element, a kind of faith that is required to explain trust and to grasp its unique nature' (Möllering, 2001, p. 404).

A reference to religious faith seems appropriate since coming to faith in God equally entails a leap. Like trust, faith is not the logical conclusion from all possible states of affairs. On the other hand, faith is not completely independent of the available information (cf. Muis, 2004, p. 37; Berkhof, 1993). Historical evidence, scientific research, or formal rules of logic certainly influence one's faith in God, but do not fully determine whether or not one comes to faith. This leads, for instance, Kierkegaard to the claim that the reflection on the facts available can be halted only by a leap. Otherwise, the subject 'is made infinitive in reflection, i.e. he does not arrive at a decision' ([1846] 1992, p. 105). This does not imply that one is no longer aware of the uncertainty or one's vulnerability, yet the commitment of authentic religious faith is deeper than one's interest to justify one's attitude based on identifiable reasons only. Hence, faith can leap beyond the evidence (cf. Adams, 1987).

This reference to commitment can also be recognised in the work of Giddens, who describes the additional element of trust as 'a leap to commitment, a quality of "faith", which is irreducible' (1991, p. 19). He argues that the 'leap of faith' is a process of bracketing the lack of knowledge and ignorance (pp. 18-19, p. 224). This underlines that the need for trust remains unchanged: one still is ignorant or confronted with uncertainty; one still lacks control. However, the truster brackets this problem and consequently trusts the other agent, since he has a commitment with respect to the trustee or the object of trust. For instance, from my commitment to friendship, I can trust my friend even beyond evidence. This is not just naïve or blind trust but the result of my commitment. If I would frequently be hesitant to trust him because of the lack of information or the probability of error in my interpretation of the facts, one may wonder whether I am really a friend. Then my core commitment suggests that I am pursuing the ideal of acting on universally justified true beliefs only.

Having a deep commitment however is not the only way to be able to make a leap. Lewis and Weigert identify two other reasons. First, the particular psychological make-up of the individual can enable him to make the leap beyond evidence. This refers to the level of basic trust or generalized trust as discussed at the start of this chapter. If the basic trust does not outweigh basic mistrust, then making the leap beyond evidence

---

[49] 'Suspension' as translation of the German *Aufhebung*.

becomes extremely difficult. Secondly, one is able to make a leap on the assumption that 'others in the social world join in the leap…. Each trusts on the assumption that others trust' (1985, p. 970). This belief with respect to the participation of others illustrates that trust presumes trust, or in other words that we have to trust in trust (Gambetta, 1988; Luhmann, 2000).

This leap aspect in formulating positive expectations explains why trust is possible even if one lacks sufficient information and can thus run counter to the demands of evidentialism.[50] It can also explain why one does not trust although the available evidence suggests that trusting is the most adequate way of dealing with the uncertainty at stake. The agent's commitment to the trustee or the object of trust, the physiological make-up and the assumption with respect to the readiness of others to make the leap of trust deeply influence the step between the interpretation of the information and one's belief with respect to the trustworthiness of the other agent.

3.6.3 Trust: More than a mental conviction

Let us return to the main question of this section, 'Is trust a belief?' The relation with evidence suggests that trust is a belief. Even though it is not and need not be a justified true belief, trust can be defined as the mental conviction with respect to the trustworthiness of the trusted agent. If one trusts another, one has a sincere belief about the motivation and competence of the trusted agent, based on presumed knowledge. Nevertheless, I do not share the conclusion that trust is 'fundamentally a cognitive notion' that fits in 'the cognitive category with knowledge and belief' as Hardin claims (2002a; b). Problems of trust are not only problems of understanding. The leap element of trust illustrates that trusting is informed by, but not exclusively based on evidence. A focus on cognition and evidence too easily ignores that the relationship between trust and evidence is more dynamic and, especially, that trust is more than a cognitive belief.

First, there is a dynamic relation between trust and evidence. Evidence is not only the input in the complex process of coming to trust. The direction is also the other way around: trust appears to be a precondition to obtain knowledge. If one is a layperson who has to rely on the available expert knowledge to judge a situation, acceptance of the evidence 'is grounded in trust rather than reason' (Faulkner, 2003, p. 33). This entails that obtaining evidence as part of the assessment of trustworthiness already implies trust, since only if you trust another, you make her a source of knowledge (2003, p. 31, 35). This raises the problem of why you trust the expert who provides you with the information. If we take trust as a cognitive belief, then trust is equally based upon the available evidence. However, obtaining this underlying evidence would again imply

---

[50] Evidentialism states that a belief is justified if and only if one has adequate evidence for that belief at a certain point in time. Cf. R. Feldman and E. Conee, (1985), 'Evidentialism', *Philosophical Studies*, 48/ 1, pp. 15 – 34.

trust. Consequently, we end up in an endless regression that will not stop as long as trust is considered as cognitive only. Nevertheless, we trust. This can be explained in two ways. On the one hand, the leap aspect of trust illustrates that at a certain point one can accept specific evidence and trust the provider as trustworthy because of a certain commitment that does not have any further justification. On the other hand, trust effects something that beliefs normally cannot do. Beliefs can shape and sharpen other beliefs, but trust colours the value we attach to certain beliefs, make them resistant to change or exclude other beliefs from deliberation.

Secondly, trust has an additional characteristic that shares important conceptual properties of a belief, but is not fundamentally cognitive. This additional characteristic is the emotional element. In practice, emotions can have characteristics in common with beliefs. These characteristics may lead us to mistake them for beliefs. Solomon uses the example of being angry about something. He claims that '"Being angry about…" is very much like "believing that…"' (1980, p. 253). This however does not mean that this feeling or emotion is the same as a belief (cf. Lahno, 2002a). The same is true for trust.

The conclusion with respect to the question whether trust is a belief is more complex than in the case of risk. There it was possible to claim that trusting and taking risks are complementary but fundamentally different. With respect to beliefs, I conclude that trust includes a sincere belief about the trustworthiness of the trusted agent that is informed by the available evidence. However, trust is more than cognitive, more than a mental conviction based on the available evidence. It further includes an emotional component.

3.7 Trust as an attitude: The emotional element

At the start of the above section on trust as a belief I quoted Pettit. However, I only used the first part of the opening sentence of his article. The full sentence runs: 'trust materializes reliably among people to the extent that they have beliefs about one another *that make trust a sensible attitude* to adopt' (1995, p. 202, my emphasis). The description of trust as an attitude can also be recognised in the position of Govier, who states that 'trust is an attitude based on beliefs and feelings and implying expectations and dispositions' (1997, p. 4). The positions taken by Pettit and Govier do not rule out the role of beliefs, but stress that trust does not coincide with these beliefs. When we define an attitude as 'a tendency to evaluate a particular entity (the attitude object) with a certain degree of favour or disfavour' (Frewer et al. 2004, p. 1183), it is obvious that an attitude is different from a belief. Beliefs can shape and sharpen other beliefs, but an attitude has the ability to colour the evaluation of the object of the attitude. This ability of an attitude has also been recognised with respect to trust. Lahno, for instance refers to trust as an attitude, since he argues that it makes us recognise something as a motive (2002a, p. 210). Govier also defines trust in terms of an attitude because trust (and

distrust) 'affect the way we think, the way we feel, and the way we act' (1998, p. 6). This can explain why the evaluation of the available information tends to confirm the pre-existing trust or distrust. The presence or absence of the attitude of trust substantially influences the way we interpret and evaluate information about the person or institution we have to rely on. Trust is informed by evidence, is intimately linked to beliefs, but in addition has an emotional dimension that plays a central part in coming to trust.

The introduction of emotions in an account of trust is not unproblematic however. It appears to make trust a highly intangible concept. The truster is not only using a subjective and pragmatic way of obtaining evidence, but also seems to adopt an attitude that is beyond rational control. Hence, building trust, such as in the agro-food sector, becomes more complicated. Nonetheless, it does not make it unrealistic. The emotional dimension implies additional criteria concerning the policy that aim to build trust, but does not rule out the possibility of building trust. To substantiate this claim, I elaborate on the relation between emotions and rational control, and between emotions and beliefs.

3.7.1 Trust as an emotion: Unreflective and beyond rational control?

Emotions have often been evaluated as feelings or moods that are mainly unwanted disturbances of an otherwise completely rational calculation (Solomon, 1980, p. 251; Nussbaum, 2001). For this reason, the inclusion of emotions in an account of trust would be an unwelcome addition if trust were conceived as mainly a rational belief based on evidence. From the perspective of trust as an evidence-based rational belief, emotions are either irrelevant since trust is essentially a rational belief, or the inclusion of emotions entails that trust becomes fully unreflective and is actually blind or at least naïve. Thus, emotions either do not contribute anything substantial to the account of trust, or the attitude that is based upon emotions is no longer worth being described as trust.

When we take the work by Lahno (2001; 2002a) at face value, it seems to confirm the idea that the inclusion of emotions deteriorates trust to an attitude that is hardly different from hope. Lahno strongly emphasises the affective aspect in trusting. He argues that 'genuine trust is an emotion' and that this is 'quite different from rational belief.' Furthermore, trust is 'in general not subject to direct rational control' (2001, pp. 171-172). Lahno stresses that emotions have a stronger impact on trust than evidence, because emotions intervene in a more immediate way than reasons. Emotions do not require the mediation of reasons (2001, p. 176) and hence Lahno denies that trust has a cognitive basis (2002a, p. 158). This view probably will be a bogy for those who consider rationality in terms of goal-oriented actions and consider trust as a rational belief based upon evidence, as has been standard in rational choice accounts.

However, Lahno's position, as well as the broader idea of trust as an emotion is more sophisticated. Lahno explicitly opposes the idea of trust as a rational belief that is based

on information alone, and the view on the truster as a rational agent who is fundamentally self-regarding.[51] He argues that this view results in 'the impression that trust is essentially only a question of rationally calculated risk' (2001, p. 172). The discussion with authors who advocate this view leads Lahno to make some bold statements on the relation between trust and rationality. However, while his conclusion with respect to the relation between trust and rational risk calculation is straightforward: trust cannot be understood in this way, his conclusion concerning the relation with reason is more subtle. He does not state that trust is fully distinct from reason, but that 'trust is beyond the *direct* control of reason' (p. 185, my emphasis). He even explicitly notes that emotions as such do not necessarily annul reason (p. 175). Emotions are not completely autonomous from rational reflection. Thus, the inclusion of an affective element in trust does not make trust completely disconnected from any form of reflection and deliberation. A truster is still able to reflect on his emotions and the emphasis on the affective aspects of trust does not exclude the impact of critical reflection by other agents. Therefore, accepting the emotional, non-calculative aspects does not make trust automatically blind (Nooteboom, 2002, p. 42).

This indicates that the classification of trust as an emotion does not by definition imply that trust is fully unreflective or beyond any control of reasoning. However, in order to fruitfully incorporate the emotional dimension of trusting in the analysis, I propose two steps. First, an emotion has to be understood as a mental state that is more complex than a feeling or a mood. If we grant that trust entails an emotional element, then this is different from saying that trust is just a feeling of affection or of sympathy with respect to another agent. Feelings can exist without any specific object, trust cannot. The expression 'I trust' is almost empty if it is not specified by describing the object and subject of my trust. If we make a distinction between trust and hope and faith, trust is always directed to specific agents and always has a certain object.

Secondly, the above analysis of trustful beliefs illustrates the importance of analysing the impact of emotions on beliefs. Although beliefs are evidence-based, they are not immune to the impact of emotions.

3.7.2 Emotions as more than feelings

Emotions have been defined in very different ways. We noticed earlier that emotions are sometimes considered as physiological disturbances of rational processes. Others, however, consider emotions as a certain kind of belief. When one considers emotions to be a kind of belief, they still can be understood as inner experiences, feelings or moods.

---

[51] The critique of rational choice accounts is not exclusive for authors who refer to trust in terms of an emotional attitude. Sztompka, for instance defines trust as a bet, yet also argues that 'rational-choice theory, at least in its orthodox version … seems to forget that calculating rational agents are also full-fledged persons, often emotional and irrational as well' (Sztompka, 1999, p. 66).

An exclusive focus on each of the characteristics easily provides a one-sided view of emotions. Therefore, I consider a definition of emotions which deals with both the elements of feelings and beliefs as the most promising for the analysis of the emotional dimension of trust. An example of such a definition is the one proposed by Frijda. He defines emotions as states that comprise feelings, psychological changes, expressive behaviour and that modify action-readiness (1986; 2000). Another example is Aristotle's definition in his *Rhetoric*: 'Emotions are all those feelings that so change men as to affect their judgements, and that are also attended by pain or pleasure' (Lahno, 2001, p. 175).

Both definitions link emotions to feelings, but explicitly emphasise that emotions are broader. Solomon (1980) has proposed an interesting view on emotions that starts with an argument of why emotions do not coincide with feelings. He argues that emotions are directed to an object. Emotions are 'about something'. A feeling of pain or loneliness can de described apart from its object. An emotion cannot. Solomon argues that 'the expression "I am angry" is incomplete – not only in the weak sense that there is more information which may be available … but "I am angry" requires that there *must* be more information available.' But feelings have no such requirements (p. 253). This does not imply that an emotion is always directly caused by its object. The emotion of fear can be directed to my dentist since he has certain tools that can hurt me. However, the actual fear can be caused by something else. The cause might be a specific cinema film or the experience of a friend. Nevertheless, my fear is not about the film or my friend's experience, but directed at the dentist (cf. Solomon, 1980, p. 256).

The argument that emotions are not similar to feelings is for Solomon the first step to define emotions in terms of judgements. My fear or love is a judgement on what is at stake. Traditionally, it has been argued that I cannot have the emotion of fear as along as I do not have the feeling. Solomon, however, argues that the main element is not the presence or lack of a feeling. The main element is whether I hold the judgement that the dentist can hurt me. If I do not judge the dentist in this way, I cannot fear him.[52] This leads the author to the conclusion that 'emotions … are judgments, and so emotions can be rational in the same sense in which judgments can be rational' (p. 262). Although, in contrast to traditional judgements, emotions 'can never be deliberate and carefully considered', they can be rational, because it is possible to make the evidence that underlie the emotional judgments coherent, consistent, and complete (pp. 270, 278-279). We cannot simply switch an emotion on or off. But 'we can open ourselves to argument, persuasion and evidence' (p. 270). If we do so, it is possible to choose our emotions: 'emotions are our choice' (p. 270).

---

[52] It is of course possible to feign an emotion, but this is different from the actual emotion. This has a similarity with the possibility of acting as if one trusts another.

This last step is debatable. It is an improvement to stress the judgement element in emotions and the consequent ability of reflection. Nonetheless, I think that the relation between the feeling involved in having an emotion and the element of control and regulation is more dynamic than Solomon suggests. Even when we open ourselves to argument and evidence, we can be struck by a certain state of affair or event. Then we cannot say that we choose our emotion. Solomon seems to neglect that there are limits to the process of reflection (cf. Frijda, 1986, p. 405).

In spite of this point of critique, Solomon has two interesting arguments that follow from his definition of emotions as a kind of judgement. First, he claims that emotions are interestingly similar to beliefs. Secondly, he argues that they are like beliefs, because 'emotions are judgements – normative and often moral judgments' (1980, p. 257). Below, the relation with belief will be discussed in further detail. Here I want to discuss the latter claim. Solomon argues that emotions are moral judgements. For instance, the emotion of anger can be interpreted as a moral judgement about the belief that you are wronged (1980). Or the emotion of vengeance is directly linked to one's belief about retributive justice (1994). The emotion is the belief that one is wronged, is treated in an unjust way or that one has suffered a loss. This shows that to have an emotion is 'to hold a normative judgment about one's situation' (p. 258). Solomon substantiates his point by elaborating on the example of being angry when John has stolen your car. The object of your emotion is the fact that John stole your car. This is also the belief you probably will have. However, your emotion includes an additional element. It is not only about the fact that John stole your car. Your emotion of anger is also inseparable from your judgement that John has wronged you in steeling the car. Therefore, your anger is your moral judgement that John has wronged you (p. 258). From this perspective, trust as an emotional attitude entails a moral judgement. Consequently, adopting a stance of trust implies that the other agent is considered as trustworthy, but it also includes a moral judgement, for instance the judgement that the trustee is benevolent or honest. Similarly, distrust can coincide with moral judgements. If one expects an institution to take care of food safety, and it turns out that they do not, then one might distrust the institution for reasons of prudence, but one's distrust can equally be seen as a moral judgement, because the trustee has wronged the truster and thus failed on a moral level.

The above discussion has shown that it is possible to provide a coherent and defendable view on emotions that neither deny that emotions are feelings nor end up in a view on emotions as completely beyond control of reason and reflection. Such a view on emotions is not advanced by Solomon alone. Nussbaum, too, stresses that emotions are not just 'unthinking forces that have no connection with our thoughts, evaluations, or plan.' According to her, they are appraisals or value judgments that we ascribe to things and persons that we regard as important for our own flourishing, yet are beyond our full control (Nussbaum, 2001, pp. 26-27, 43, 90). Against the background of this view on emotions as feelings that are directed to a specific object and that include judgements

about that object, it is interesting to analyse the dynamics between emotions and beliefs and the consequences for trust.

3.7.3 The influence of emotion on trustful beliefs

As noted, Solomon holds that emotions are 'interestingly similar to beliefs' (1980, p. 257). Having the emotion about *x* often includes a belief about *x*. Yet the relation between them is not immediately clear. In his account on emotions, Frijda states that emotions and beliefs are two mental states that are to be distinguished, yet mutually influence each other. Beliefs determine emotions. Our emotional judgements are always structured by concepts and models that we already have. The influence, however, is also in the reverse direction: emotions can create and shape beliefs and can make them resistant to change. Emotions do so by guiding our attention and making 'some things appear more salient than others.' They stimulate certain associations and suggest 'certain patterns of interpretation' and 'guide our evaluation of some aspects of the world and motivate our actions.' (Lahno, 2001, p. 175). Emotions have the capacity to 'permeate our experiences' (Solomon, 1994, p. 296), and make us look at the object 'through one's own window' (Nussbaum, 2001, p. 28). This implies that our emotions influence our beliefs and modify action-readiness, because they define the way we perceive the world and our conception of the world (Frijda et al. 2000, p. 3; Solomon, 1994, p. 297). Emotions can influence beliefs, because beliefs are more emotion sensitive than knowledge[53] (Frijda et al. 2000, p. 4).

Thus, our beliefs are directly influenced by emotions, because the perception and interpretation of the world around us, which can serve as the evidential basis for beliefs, are coloured by our emotions. When we refer to emotions in these terms, it is a small step to recognise the emotional dimension of trust. In the literature on trust it has been observed repeatedly that 'trust is a way of seeing that guides our attention, colours our perceptions, and thus gives rise to certain beliefs' (Miller, 2000). The presence of trust makes us interpret another's behaviour and the available information through a 'lens of trust' (Jones, 1996, p. 13), which implies a specific way of evaluating information (cf. Lahno, 2001; 2002a).

This view on trust explains the observed differences between individual trusters in perceiving and interpreting information. Some evidence can be sufficient reason for person A to entrust something to an institution, while person B does not consider the evidence as an adequate reason to trust. For instance, someone who trusts the agro-food sector will probably perceive a large-scale recall of a product by a food company as a confirmation of his trust. While someone who lacks such trust in the sector may well

---

[53] As we have seen above beliefs do not necessary include the reference to truth claims that knowledge does have.

have the idea that she just escaped from another food crisis. The same situation with the same level of available information is perceived completely differently. The inclusion of an emotional element explains this situation, since not only information influences one's stance of trust, but also the presence or absence of trust colours the perception of the information and the truster's beliefs. In other words, 'trust itself affects the evidence we are looking for' (Gambetta, 1988, p. 233) and consequently trust tends to confirm our beliefs. In literature on communication and trust, this is called the 'confirmatory bias' (see e.g. Poortinga & Pidgeon, 2004).[54]

However, the idea that the presence or absence of trust colours the way we evaluate the available evidence does not necessarily require the inclusion of an emotional element. Hardin for example can explain this point without considering trust as an emotion. He states that 'my estimation of the risk is merely my degree of trust in you' (2002, p. 119). This implies that one's perception of the relevant matter is coloured by the presence or absence of trust. Hardin emphasises that what appears to a person as sensible to expect, is inherently subjective (2002, p. 130). For Hardin the subjective element remains a matter of cognition without any reference to emotion. It is subjective because trust is based on the personal assessment of what is sensible to expect of the agent to whom trust is extended. However, trust is still is a matter of cognition, because the assessment is based upon the individual's knowledge about the agent. From this perspective, it is not surprising that Hardin can deal with the differences in perception, because the beliefs we have also influence other beliefs, and thus the way we perceive the world. For instance, the pre-modern belief that the world is flat directly influenced the way people perceived that world for ages and explains the problems that occurred when it was discovered that we inhabit a round planet.

Trust as an emotion, however, implies one further step. It does not only modify our beliefs, but can directly influence them in a stronger way than knowledge can. This influence can be so strong that we can hold an attitude of trust or distrust although we lack the evidence, and even do not have the belief that it is justified. Suppose that the producer of a useful novel food product has been recommended by your local GP, who is a close friend. You believe that the producer is trustworthy. Nonetheless, you do not buy the product since you are unable to help regarding the company with suspicion, because you look through a 'lens of distrust' at this kind of companies. Consequently, you see them as untrustworthy partners (cf. Jones, 1996, p. 24). If trust were to be a belief based on justified knowledge your attitude would be unjustified, however, your suspicion is

---

[54] Communication is further complicated by another bias, i.e. the 'negativity bias' (Siegrist & Cvetkovich, 2001) which means that sources of bad news tend to be considered as more credible than sources of good news. In combination with the 'asymmetry principle' that states that there is an asymmetry in trust-building versus trust-destroying processes, this easily explains why trust is much easier to destroy than to be established (Slovic, 1999, p. 698), and why Baier describes trust as a 'notoriously vulnerable good, easily wounded and not at all easily healed' (1994b, p. 130).

sincere even though it is hardly possible to articulate the reason for it. Ascribing this emotional dimension to trust can explain why we can remain hesitant to trust even when others or our own beliefs tell us to do otherwise, and conversely can clarify why we can remain trustful even if the evidence tells us not to trust.

Furthermore, the influence of emotions on our beliefs surfaces at the level of motivation. On the one hand, emotions can immediately motivate to act. For instance, we can act out of anger or fear, i.e. rash judgements which do not need the intervention of reasoning. On the other hand, emotions can motivate us, since they make us perceive certain beliefs or states of affairs as more attractive than others. Since they can change our beliefs, they have an influence on our action-readiness and serve as a motivation to act. This again is the result of the characteristic of emotion as a phenomenon that colours our perception and influence the process of reasoning.

In sum, our beliefs structure our emotions. However, our emotions, i.e. our emotional judgements, influence our beliefs in an even stronger way by guiding our perception and evaluation of specific aspects of the world.

3.7.4 Conclusion

In this section, I have argued that a definition of trust as an emotional attitude is preferable to proposals that define trust only in terms of cognition and beliefs. This does not imply that beliefs play no part. On the contrary, our perception is coloured by emotions, but at the same time, 'anything we perceive is already structured by concepts' (Lahno 2001, p. 176). Consequently, it would be impossible, but also unnecessary for my position to deny this. This holds also for the relevance of cognition. The discussion of emotions as judgements has illustrated that defining trust in terms of an emotional attitude entails that trust has a cognitive element, but it is not fundamentally cognitive, as Hardin and others insist it is.

If we define trust as an attitude with a demonstrably emotional dimension, then we can explain the complex influence of evidence and of communication on building and maintaining trust. From my position, it is possible to discuss these points without reducing them to problems caused by the truster's irrationality or ignorance. If the emotional element is explicated, then it is possible to show that not (only) ignorance leads to different evaluations of the trustworthiness of an agent, but that the individual's emotional judgements steer the perception of the available evidence. First, this illustrates that building trust requires more information – and it requires time. Time to change, and to revise or adapt one's view on the other agent. Partly this change of view will be based on cognitive elements, but it will also be based on the behaviour of the agent who wants to be trusted. Second, allowing emotions a more prominent position in the analysis of trust helps us to see that the differences in evaluation of information are not just an indication of the irrational truster. In a situation of a lack of trust, the suspicion is sincere

even if it is hardly possible to articulate the reason for it. Consequently, this emotion has to be taken seriously rather than disregarded as irrational. We should do so, first, because it can be a marker of an implicit issue with respect to either the competence or motivation of the trustee. Secondly, because it is possible to ask the individual to reflect on his emotion, since it is directed to a specific trustee and includes a judgement about that agent.

3.8 Summary and working definition

Against the background of the discussion of trust as an action, as a belief and as an attitude, I return to the outline of trust that has been presented in Section 3.3.5. At that point I stated that trust (a) is relevant in cases in which agents lack control and are uncertain about future actions and events, (b) is related to the sincere belief that the other party can be trusted. This belief is based on one's assessment of the trusted party. (c) Trust is an orientation that we have towards individual or groups of human agents rather than to machines or natural phenomena and (d) trust presumes a relationship that includes social interaction between the truster and the trustee.

These four aspects can be elaborated based on the above discussions. Trust results in actions, but does not coincide with its results. Furthermore, trust is relevant in cases of uncertainty, yet has to be distinguished from the action of risk taking. Both risk calculation and trust are ways of dealing with uncertain situations. In that sense, they are complementary. However, risk taking and trusting are fundamentally different.

The characteristic of trust as a sincere belief about the trusted party has been the basis for the discussion of trust as a belief. The conclusion was that trust is intimately linked with beliefs, but that it is not to be identified with belief. From the analysis of trust as a belief two difficulties remain. First, the so-called leap element of trust illustrates that trusting includes more than a belief that is exclusively based on evidence. Information facilitates trust, but it is impossible to define a sufficient level of evidence to arrive at trust. There is an element in trust that 'happens to us', rather than that we decide to adopt a stance of trust. We can be struck by some information or some behaviour that makes us trust a specific agent while in another situation the same information will perhaps not lead us to trust. Secondly, the dynamic relation between trust and evidence remains unanswered if we consider trust as a cognitive belief. Evidence is not the only input in the process of coming to trust. The direction is also the other way around: trust appears to be a precondition to obtain knowledge.

This illustrates that trust has an ability that beliefs normally do not have: it can colour the value we attach to certain beliefs, make them resistant to change or exclude other beliefs from deliberation. To deal with these features of trust the emotional element of this concept has to be taken seriously. Explicating the emotional element shows that

emotional judgements steer the perception of the available evidence. This is not to say that trust is only a feeling that appears fully independent of evidence. It is possible to ask the individual to reflect on his emotions. Given this possibility of reflection, differences in evaluation of information can turn out to be a marker of an implicit problem with respect to the competence or motivation of the trustee rather than the result of irrationality or ignorance.

As a conclusion, I define trust as:

> an attitude towards (collective) humans that enables an agent to cope with situations of uncertainty and lack of control, by formulating a positive expectation towards another agent, based on the assessment of the trustworthiness of the trusted agent.

This view on trust underlies the next chapters that are the building blocks for and the elaboration of the moral condition of trustworthiness. I will argue that any attempt to be trustworthy should start from respect for the autonomy of the truster, i.e. in the recognition of him as a moral agent and as a moral equal. The next chapter will function as the basis for this moral condition of trustworthiness and deals with the elements of agency and freedom. Trust as an attitude towards (collective) humans starts from the acknowledgement of the other as a free and autonomous agent. Chapter 4 will show the background, the necessity and implications of this assumption.

CHAPTER 4

FREEDOM, AGENCY AND THE POSSIBILITY OF TRUST

If your attitude towards someone is wholly objective,
then though you may fight him, you cannot quarrel with him,
and though you may talk to him, even negotiate with him,
you cannot reason with him.
You can at most pretend to quarrel, or to reason, with him.
P.F Strawson, 1974, p. 9

4.1 Introduction

Speaking about trust implies several assumptions. This chapter aims to trace and analyse the assumptions that are the necessary conditions to define an attitude as 'trust' and a response as a proof of 'trustworthiness'.

A first basic requirement is the recognition of the individual as someone who operates within a social and moral context but cannot be reduced to being a member of a society or a community. If individuals are seen as indistinguishable from their society or community, then one is reliable as long as one represents that community or performs a role that is determined by it. This is something other than trust however, because in this situation one's personal competence and motivation are irrelevant. To show this point, I start with a discussion of the historical process of recognising the individual as distinct from his social and moral context. This discussion shows that the recognition of the individual automatically confronts us with questions of trustworthiness (Sections 4.2–4.3). If the acts and behaviour of the trusted are not determined by his membership of a community or his religious beliefs, it raises the question of what motivates him to respond to your trust. When will an individual trust and respond to trust? Hobbes was one of the first to address this problem, especially the issue whether it is possible that individuals cooperate on a public level. By discussing his account in its historical and theological context, I aim to show its relevance for the current debate on trust, but also that Hobbes's answer is problematic. His proposal to establish public cooperation and coordination by introducing coercion and sanctions may establish order but does not establish trust (Section 4.4). Coercion or sanctions as tools to reduce the individual's freedom or restrict his agency block trust or at least lead to serious problems.

This observation from the discussion of Hobbes shows a second requirement to define an attitude as trust and a response as trustworthy. *Freedom and intentionality* as constitutive elements of agency are conceptual requirements of trust (Section 4.5). These requirements restrict the number of possible answers to the question of when an individual will trust and respond to trust. Any valid answer has to respect the freedom and agency of the individual; trust is a way of 'handling the freedom of other human agents or agencies' (Dunn, 1988, p. 73).

Rational choice theories (which play an important role in the literature on trust and cooperation) present some answers to questions of trust and trustworthiness without reductions at the level of freedom and agency. Therefore, I discuss their answers and the theoretical background at some length. Rational choice theories take the individual as a free agent seriously, but also have great difficulty with including trust in the account. Because of the specific view on rationality as a self-interested means–end calculation, trust is not prima facie excluded, but it is difficult to explain why someone comes to trust or start to act in a trust-responsive way. The difficulty can be resolved if what motivates to act is not restricted to sheer self-interest, but broadened to other-regarding interests as well. Moreover, trust can better be explained if rationality is not considered to conflict with reasonableness. With these changes at the level of motivation and rationality, the possible answer to the question of when an individual will trust and respond to trust is no longer restricted to 'if and only if it is in the self-interest of the individual.' We can also trust for other reasons and can be expected to respond to trust even if it runs counter to our self-interests (Sections 4.6 and 4.7).

This illustrates the need a third requirement. Trust presupposes a participant stance towards the other agent. As long as both agents consider each other as free, but not as participants in an interaction, trust remains problematic. The other agent is not a source or tool in the process of achieving one's aims in the most efficient way. To have a full answer to the question of when a free agent will trust and respond to trust, the truster and trustee should recognise each other as full participants in a trusting relationship. Both should be considered as autonomous persons (Section 4.8).

4.2 Faith and trust in a free God

Trust is not a new phenomenon. From the collection of papers edited by Frevert (2003), it is clear that trust has a long history. Because we have referred to trust as an attitude, which is relevant if one is confronted with uncertainty and with a situation of a lack of control, it should not come as a surprise that trust has such a tradition. Even though pre-modern societies were far less complex than our 21st-century society there was as much uncertainty and vulnerability. For instance, travelling or doing business were real

ventures, and far more risk related than at present. Hence, the need to rely on others was for a medieval merchant as strong as it is within the current food sector.

Traditionally, trust is most frequently mentioned in the realm of religion. The Latin concepts for trust and faith show this close link. Latin *fides* has the same root as the word used for trust: *fiducia* (cf. Pagden, 1988; Frevert, 2003). A 17th-century Puritan for instance defined faith as '*vera et propria fiducia*', true and proper trust (Ames, 1628, I, 3, 13). From this perspective, faith in God includes trust in God. On the one hand, this trust is of a unique kind, since it is the reliance of a creature on his creator. On the other hand, trust is seen as the source of and paradigm for all other forms of trust (Frevert, 2003, p. 15).[55] Trust in other humans was a reflection of the trust one has in God.

Even though the relevance of an analysis of trust in God for the understanding of trust in humans has been declining after the 18th century, a focus on trust in God shows a first indication of the relevance of the concepts of agency and freedom. Already the medieval discussion on the attributes of God touches upon the relation between the freedom of God and his trustworthiness. In the debate on the question whether God can make things otherwise than he has ordered them to be made, the issue of trustworthiness immediately surfaces. If God did not have the freedom to act otherwise than in line with the order of his own creation, the question is whether he can be trusted (Guleserian, 2001). The answer seems negative. Such a god certainly is predictable, but we have seen before that trust implies more. If God could only act according to a predictable pattern that is determined by his so-called ordained power (*potentia ordinata*), we could only rely on him as we rely on machines which have the power to do *x*, but not to deliberately refrain from it. Thus if one says that God is trustworthy, it means that God is not just predictable, but worthy of trust even though he has the power to change his mind (cf. Guleserian, 2001). On the other hand, if God through his absolute power (*potentia absoluta*) could have acted otherwise than He factually did, this freedom raises a problem too. 'The reliability of God's character and the reliability of created reality would no longer be guaranteed' (Veldhuis, 2000, p. 222). The freedom of God to act otherwise than the order he created suggests jeopardising the trustworthiness of God, since it suggests that he has an arbitrary and unpredictable character and that the structure of creation is unreliable (p. 229). The discussion is even more complicated by the question whether a god who is free and can be trusted has the capacity to disregard or even violate trust. If he did not have this capacity since he cannot do a wrong act it has been argued that this 'moral immutability of God renders trust in God impossible' (Guleserian, 2001, p. 294). On the other hand, if He has the freedom to violate trust it raises new problems of trustworthiness.

---

[55] Frevert focuses her analysis on the use of trust in German language (Vertrauen). Her results, however, also apply *mutatis mutandis* to the broader use of trust (e.g. Pagden, 1988).

This paradoxical function of freedom as a necessary condition for trust on the one hand, and as the source for questions of trustworthiness on the other hand has been addressed in various ways. For instance, the medieval theologian Duns Scotus argued that God has the freedom to act otherwise than the order he created, but that this does not make Him arbitrary and unreliable, since 'God's acts cannot be in conflict with his essence' (Veldhuis, 2000, p. 230). Whatever the precise answer is to the problem of whether God can be both free and trustworthy, it is evident that the problem cannot be 'solved' by disregarding one of the aspects. If the aspect of freedom is denied then the trustee, in this case God, is no longer a proper subject of trust. If the trustworthiness of God were doubted then the attitude of trust would be out of place. Trust requires a free agent *and* the assessment of the agent as competent and properly motivated. This holds for trust in humans as it holds for trust in God. On top of this, some arguments, which show that God can be both free and trustworthy, do not directly apply to humans. For instance, Scotus's answer that God is always worthy of trust since 'God's acting is structured by his essence' does not hold for humans to the same extent (Veldhuis, p. 225). This illustrates that the question what incentive a trustee has to respond to what is entrusted becomes even more prominent with respect to trust between humans than in the case of trust in God. This point will be elaborated in Sections 4.4 and 4.7. The sections hereafter deal with agency and freedom as essential features for trust and illustrate that trust surfaces as a result of a double emancipation: the emancipation of the individual from the structure of society and the emancipation of morality from social and divine sanctions.

4.3 A double emancipation and the consequences for trust and trustworthiness

It would be a caricature to argue that individuals played no independent role in societies up until the Renaissance. The Western emphasis on the individual as independent agent within social and religious structures has its roots in classical antiquity. However, since the Renaissance and the Protestant Reformation of the 16th century the individual as an independent agent within a community has been more and more recognised in arts, politics and morality. Locke for instance remarks that 'truth and keeping of faith belong to men, as men, and not as members of society' (1960 [1689-90], Section 14). The recognition of the individual as discernible from the community results in a double emancipation. First, the emancipation of individual agency from the structures of society. This implies the acknowledgment that an individual operates within a community, but that his acts and behaviour are not fully determined by being a member of the community. Second, the emancipation of morality from social and divine

sanctions. These two developments have direct consequences at the level of trust and trustworthiness.

4.3.1 Three consequences of the recognition of individual agency

First, the emancipation of the individual is an important condition for speaking about trust in contrast to mere reliance. Trust – in terms of the assessment of someone's competence and motivation – requires the recognition of the individual as a free agent who is embedded in a group or community, but whose acts and behaviour do not coincide with this group or community. Given this acknowledgment, reliance can turn into trust. As long as the acts and behaviour of someone are (almost) fully community structured, one can rely on her as a mere part of that specific community. This reliance is not yet trust in her as an individual. The reliance is mainly grounded on an assessment of the patterns and routines that structure the group or community, rather than on an assessment of individual competence and motivation. If I note that someone's acts and behaviour are influenced more by community norms, patterns and routines than by competence and motivation I need not refer to 'trust'. This does not imply that trust is always between atomistic individuals who are fully independent of any group or community. Trust does not rule out that the acts and behaviour of the trusted person are embedded in and formed by the community she lives in or the organisation she works for. Trust however requires that the acts and behaviour of the individual can be distinguished from those of the community, even though they cannot be fully separated from the community.

Second, the recognition of the individual implies that the number of trustful relations grows enormously. Trust is no longer based on a group or on group membership, but trust is in the individual. Consequently, if one is confronted with uncertainty it is not sufficient to rely on the community only, trust in its individual members is necessary too. This confronts the truster with an increasing number of known and unknown agents on whom he has to rely in order to pursue his aims and goals. This situation is intensified by the fact that in our differentiated world both the truster and the trustee are no longer members of a single community. We all participate in different groups that have different aims and in which we have different responsibilities. The individual can be a father, a member of the choir, an employee, a citizen, secretary of the local social club etc. In each of these roles, the truster has to rely on other agents who are embedded in a social, cultural, or religious context, yet are not interchangeable with other members of the same community. This leads to a third consequence.

The recognition of the individual agent does not only influence the number of trusting relations, it also raises questions of trustworthiness. The competence and motivation of a member of a community may be structured by this membership. However, an individual has beliefs, preferences and desires that differ from those of his

community. Since his acts and behaviour are not determined by the group membership, he has the opportunity to act in a way which is different from or even contrary to that which is common in the group. Hence, trust in one member of a certain community does not necessarily give enough reason to trust all of them. The trustworthiness is not determined by the fact that one is member of a certain group, but has to be assessed on the individual level.

In short, it is a precondition for speaking about trust to recognise the trusted individual as an agent whose acts, beliefs and desires are not fully steered by the community in which he lives and works. This precondition has two implications. First, being a member of a certain community can be a sufficient proof of reliance, but not of trustworthiness. Trustworthiness always relates to the individual competence and motivation of the trusted agent. Being a member of a certain community can be constitutive for both aspects, but can never replace them. Second, the precondition implies an increase in the number of situations in which one has to rely, because the truster has to enter trustful relationships with *individuals* rather than with the *community* represented by the agents.

4.3.2 Individuality, morality and trustworthiness

The consequences for trustworthiness are amplified by a second emancipation, which is on the moral level. This also starts in the recognition of the autonomous position of the individual agent. The individual receives a more prominent position in moral deliberation. Especially the Protestant Reformation created space for an individualisation process on the moral level by emphasising the importance of the individual conscience in moral reasoning. The Scriptures only gave general principles to keep the believers responsible for precise application of the moral rules. The important change is that the mediating structures of the church came to be considered as less important than the unmediated access of the believer to God (cf. Seligman, 1998). For instance, the Puritans of the 16th and 17th century argued that every person should be his own confessor.[56] Individual conscience became central to morality.

Initially this liberation did not lead to the detachment of morality from divine authority. God was still considered as the law-giver and the Bible as the only authority for both faith and discipline. Faith in God still provided a clear answer to the question 'why be moral?' For instance, Bishop Butler (1692-1752) considered our conscience to be framed by God. The focus on individual responsibility still started from a law-based obligation since God is the law-giver. The motivation to act morally is the obedience to

---

[56] According to Goeters this kind of development towards individualisation (or subjectivism) is one of the general characteristics of the Puritan movement. Goeters, W. (1911), *Die Vorbereitung des Pietismus*, Leipzig, pp. 59-60.

God rather than one's own ideas about the compatibility of the law with one's idea of the good. With respect to trustworthiness, this is important for two reasons.

First, one's belief in God can be a prudential reason to rely on another. As a truster, one enters, by definition, an asymmetric relationship in which one is vulnerable and dependent on the trustee. In the context of a Divine Command Theory, the truster has a clear reason to enter this relationship in spite of the entailed vulnerability, because the trustee is in a dependent position too. The trustee shares the predicament of dependency with the truster, because both depend on God. This dependency on God can make a trustee reliable if it is combined with a threat of divine sanctions if the believer does act in line with the (moral) requirements that result from one's faith. Consequently, I can rely on the other person, since his fear for the sanctions gives him a prudential reason to act in the favourable way. However, the validity of this incentive to trust is restricted to the group that shares the faith in God and the fear of sanction. Therefore, Locke argues that atheists cannot be trusted because they do not fear the ultimate retribution of God (Hardin, 2002, p. 208).[57]

Second, the trustee's motivation to act in the expected way is based on the recognition of the moral authority that is shared by the truster. This motivation is the divine law that has authority since it is provided by God, who as the Creator of all has legitimate authority over us. This is more than a value similarity. It even implies that we share the same justification for a certain moral expectation. The trustworthiness of the other is a reflection of God's trustworthiness, and the nature of God serves as sufficient justification of the moral expectations. This view entails that all others who have faith in God and recognise his moral authority are initially motivated to act in a trustworthy manner.

However, for all who are not so motivated or even deny that moral obligations are divinely imposed, these two elements no longer give a reason to trust one another. This situation surfaced when the Enlightenment project took a further step in the emancipation process of the individual and rejected the Divine Command Theory as the starting point for morality. The command or will of God was no longer considered as sufficient for an action to be morally right, or obligatory. For example, according to Kant, morality should not be based on theological conceptions which derive morality from a Divine all-perfect will. For this position, Kant has two reasons. First, an epistemic argument: we have no intuition of the divine perfection and can only derive it from our

---

[57] Advocates of a (sophisticated) Divine Command Theory agree that moral motivation based solely on prudential arguments such as reward and punishment is inadequate. Nevertheless, they emphasise that the motivational power is one of the positive features of a Divine Command Theory. Van den Beld, for instance, argues that this theory 'is able to account not only for the reality but also for the inescapability of costly moral obligations. Nobody can get off the hook of a particular costly moral obligation by reminding him or herself … of the limits of moral obligations in general, and by taking distance from a particular one. For moral obligations are ultimately not self-imposed … but divinely imposed' (2001, p. 397).

concepts, the most important of which is that of morality. Consequently, we would end up in a vicious circle if we infer our moral requirements from God's will. Second, if we want to avoid this circular explanation, there remains, according to Kant, only one concept of divine will. This is the concept that is made up of the attributes of desire for glory and dominion combined with the awful representations of might and vengeance that 'would have to be the foundation for a system of morals that would be directly opposed to morality' (Kant, 1996, p. 91). From another theoretical perspective, Hume also defines morality in contrast to the divine-law account as a natural fact. 'Our situation in nature ... render[s] particular virtues useful or necessary for us and society' (quoted by Rawls, 2000, p. 58). Its naturalness is defined as a part of a natural law, but it is natural, not because it stems from God but because it is fully consistent with human psychology.

To abandon the Divine Command Theory does not imply that everything is permitted, or that there is no morality left. It implies, however, that an additional motivation in human nature to act morally is required 'distinct from the sense of its morality' (cf. Hume, 2007 [1739-40], 3, II, I). Even if God is still recognised as the creator of all, the divine element is not considered as a sufficient argument for the justification of moral obligations and as a motivation to act morally. For trust this implies that both the fear of 'the ultimate retribution of God' and a shared religiously based justification of a trustful expectation fail as reason to trust someone else. This illustrates that a new question of trustworthiness surfaces: 'What motivates an individual to act in a trustworthy manner when the Divine Will is neither the source of prudential reasons to act in a reliable manner, nor the basis of shared moral understandings that make someone trustworthy? Hobbes offered an early – influential – answer.

4.4 Individuals, order and sanctions: A Hobbesian contract

Trust in others, who are recognised as individuals with beliefs, preferences and desires that are not fully determined by the social, cultural or religious community they belong to, faces us with the question of what incentives these others have for responding to one's trust. This highlights the problem of trust between individuals within a society. Can we trust each other as individuals with respect to public issues such as safety and security? As one of the first, Hobbes started out from the individual in answering the question of cooperation on a social level. His answer is still relevant and interesting. It is relevant since it is still influential in the current debate on public trust and social cooperation. It is interesting because some of the flaws ascribed to Hobbes's answer can show some blind alleys in the discussion on trust while other alleged problems of the Hobbesian approach are due to misunderstanding of the historical and theological context of his writings (Meijboom, 2008).

4.4.1 A real war as the cause

Hobbes's *Leviathan* was not written as just a proof of scholarly competence. As for many of his contemporaries,[58] the 17th-century Civil War is for Hobbes the main reason for writing his political theory. His is a direct response to the social and religious collapse during and after this civil war. Hobbes claims that society is so seriously shattered that no authority is in a position to enable cooperation. We are left with individuals only. There is no unifying doctrine, no central power, nor do we have enough knowledge of God's attributes that may bring us together and enable us to cooperate (cf. Weil, 1987, p. 774). Hence, it is understandable that he argues that such a society is not of any religion at all. A society allows religions of individuals, but there is no "public worship" (1651*, XXXI, p. 225).

As a further complication for cooperation, Hobbes realises that individual rational choice may yield all kinds of serious problems of cooperation and irrational outcomes (Hollis, 1998, pp. 34-35). Hobbes considers humans as fundamentally equal in power, but also equal in their self-interestedness. Hence, in cases of competition or scarcity, we initially will not care for others unless our interests are closely related to those of the other. 'The condition of man' he says, 'is a condition of war of every one against every one, in which case every one is governed by his own reason, and there is nothing he can make use of, that may not be a help unto him in preserving his life against his enemies' (1651, XIV, p. 80). It is obvious that in this situation of 'war of every one against every one' destruction and violence are unavoidable and cooperation pointless (1651, XIII, pp. 78-79). This picture has serious consequences for trust: it becomes impossible (Hardin, 1993). The 'condition of man' applies to all; hence you can never be sure that the other will perform in the expected way or will do what you entrust the other to do (1651, XIV, p. 84). In *The Elements* Hobbes calls this a state of 'perpetual diffidence' (1640, I, Ch. 14, p. 11). This situation is not easily overcome, since we simply lack any reliable information, or authoritative person or institution to serve as the foundation for our trust. The individual is locked into a situation in which he has to rely on another in order to act, but at the same time cannot count on the reliability of the other.

4.4.2 Hobbes and the view on human nature

The view on human life as 'solitary, poor, nasty, brutish, and short' is often seen as the core of a Hobbesian anthropology. His view on cooperation and human motivation,

---

[58] Weil (1987) notes that the political situation in the mid-17th Century was an incentive for many authors to draft their political theories and statements (p. 761, with references to work of Q. Skinner).

however, is more complex.[59] He was well aware of the extreme character of the picture he provided, but he asks his readers to consider their own experiences, for instance, when they travel or when they go to sleep. One will protect oneself by using arms and locking the doors. Evidently Hobbes was prepared to be 'realistic' about human nature. More important, however, for the understanding of Hobbes's view on human nature is the fact that the state of nature has a much more complex function than picturing the original situation that we have to escape.[60] It is used as a rhetorical device in his writings especially for those involved in the Civil War: it is a possible and unpreferable destination (Pasquino, 2001, pp. 407-408). From this perspective, the state of nature does not merely reflect who we are and what motivates us by nature as a necessity, but serves as an example of where we will end up without a central power that makes us cooperate.

The state of nature has an evident dialectic character. It has three functions in Hobbes's work. First, the original situation is the reason for forming a commonwealth. Second, the state of nature is the situation that still underlies a society and forms the permanent reason for attributing exclusive power to the sovereign. However, thirdly, it is the continuous threat to a society, since it is a realistic scenario. Thus, the state of nature is both cause and threat to Hobbes's society. It is both inside and outside the society. The state of nature serves as a hermeneutical device to focus on and analyse society, not to providing a complete anthropology. It only shows that human beings have a tendency to be essentially self-regarding if they are not bound by a society, i.e. a sovereign. In a society the state of nature is still present, but constrained by social agreements and commitments to others, especially to the sovereign.

4.4.3 A Sovereign and the ability to cooperate

In order to make social cooperation possible, Hobbes emphasises the need of a sovereign. It is not in the power of the individual to go beyond the state of nature. To do so would be similar to the adventure of Baron von Munchhausen, who pulled himself out of a swamp by his own hair. It is crucial therefore that one transfers one's rights to a central authority. Otherwise, 'every man has right to everything and consequently, no action can be unjust' (1651, XV, p. 88). The only possible state is a civic state formed by individual citizens who wish (and by natural law are obliged) to surrender their power to one central authority in order to live in peace and order. Hobbes claims that this is a necessity, because the diversity of wills and voices and the lack of a *finis ultimus* or

---

[59] See Weil, F.D., 1987, 'The stranger, prudence, and trust in Hobbes's theory', *Theory and Society*, 15, pp. 759-788 and Skinner, Q., 2005, 'Hobbes on Representation', *European Journal of Philosophy*, 13/2, pp. 155–184.
[60] As Strauss has argued, the state of nature is the actual starting point of Hobbes's political philosophy, Strauss, L., 1961, *The political philosophy of Hobbes, its basis and its genesis*, Chicago: University of Chicago Press.

summum bonum leaves individuals with no option other than to 'confer all their power and strength upon one man, or upon one assembly of men, that may reduce all their wills, by plurality of voices, unto one will' (1651, XI, p. 60; XVII, p. 106). We simply can no longer use biblical or ecclesiastical traditions as a fall-back position.

In establishing and obeying the central authority, there is a clear resemblance with the way God has created man, and the reason man has for obeying God. First, the sovereign needed to facilitate cooperation is a self-created sovereign. In spite of all the negative views on human nature, the first sentence in the introduction of *Leviathan* states that we are able to imitate nature and make an 'artificial animal' (1651, p. 7). This imitation of nature is closely related to God's way of creating man. Hobbes states in the introductory sentences that nature is the art whereby God made and governs the world. As humans, we are able to imitate this nature. Creating a sovereign is a kind of *imitatio Dei*.

Second, this imitation exceeds the level of reproducing what God has done. Miller shows that Hobbes has a higher aim. We are capable of and need to be imitators of God in the sense that we use the art in order to produce a Commonwealth (1999, pp. 164-168). It is the creation of a man as we, but with 'greater stature and strength' than in the natural condition. Thus, it is *imitatio Dei* in the sense of creating like God, but it is also *imago Dei* in the sense that we create after our own image like God has done. The Leviathan is man's creation after his own projected image. This imitation is crucial for Hobbes, not as a contradictory statement on human nature, but as the only way to go beyond the perpetual state of distrust, since he[61] imitates God in establishing order. Only this figure of greater stature and strength can enable us to progress from the natural state to an artificial, civil state. In Hobbes's words, 'that mortal god to which we owe, under the immortal God, our peace and defence' (1651, XVII, p. 106).

Third, the sovereign has this crucial function within a Hobbesian society because he is also an imitation of God with respect to the reason for obedience. Like God, the sovereign of a society has to be obeyed because of his power (XXXI, p. 219). Nevertheless, there is a fundamental difference. A sovereign will not replace God, since as mortals we cannot make something immortal (XXX, p. 197). Yet, the fear of God that is 'a confession of His power'[62] (XXXI, p. 224) has a parallel on earth in the fear of the sovereign. The emphasis on the notion of fear is crucial in order to provide a commonwealth with stability. As long as there is unclarity about the motivation to cooperate, i.e. when there is a 'fear of non-performance on either part' it frustrates any form of cooperation. 'there must be some coercive power to compel men equally to the performance of their covenants, by the terror of some punishment greater than the benefit they expect by the breach of their covenant' (XV, p. 88).In this context, 'fear is

---

[61] Hobbes emphasises that the central power does not refer to one (natural) person, but to 'the disembodied and fictional Person whose generic name is the State' (Skinner, 2005, p. 178).
[62] This fear is not only fear in the strict sense; it includes awe and respect.

the key to trust' (Hollis, 1998 p. 31). Only the combination of fear and power opens the way to cooperation, since then it is in our interest to forfeit 'our own chance to do the other down, in exchange for protection from the other agent's aggression' (Hollis, 1998 p. 31). The reason to rely on one another is prudential rather than morally or altruistically motivated. The agents in a Hobbesian contract act honestly, but are not necessarily honest. They only act honestly out of fear and thus may change their behaviour immediately whenever it is safe to do so. Therefore, Hollis concludes, 'reliance' in Hobbes's model is merely prudential and Hobbes can only include a moral obligation as the basis of trust by bringing into play God 'as a joker, by invoking eternal laws and divinely ordained principles of justice' (1998, p. 35). I think Hollis is right in so far as the Leviathan idea needs God to deliver a genuine moral article. However, there is a far more fundamental problem: the Leviathan does not establish trust at all.

4.4.4 The recognition of the individual a warrant for order, not necessarily for trust

From the discussion of Hobbes's answer to the question what incentives individuals may have to cooperate, it is possible to draw some conclusions. First, Hobbes's account clearly illustrates that the recognition of the individual as discernible from his or her social and religious community is a precondition to speak about trust, but also shows the profound question of trustworthiness. The recognition of individuality is a necessary, yet not a sufficient precondition to speak about trust.

Second, Hobbes has made clear that the issue of trustworthiness is intensified by the specific nature of the individual. His recognition of the individual as self-interested has found many followers. Hume for example writes: 'nothing is more certain, than that men are … govern'd by interest, and that even when they extend their concern beyond themselves, `t is not to any great distance; nor is it usual for them, in common life, to look farther than their nearest friends and acquaintances' (2007, [1739-40], 3, II, VII). More recently, rational choice theories start with a view on human interest that is close to that of Hobbes. As we saw in Chapter 3, Section 3.5, Coleman (1990) starts from the individual as atomistic, essentially self-regarding agent.

A third conclusion is that Hobbes's approach has some serious flaws as well. The identification of these points can inform the current discussion on trust and trustworthiness. One of the problems is that trust is more than mere order. Hobbes aims to enable individuals to cooperate by establishing a contract that results in a clear order. However, building contracts to establish order is different from cooperation based on trust. Order in a society can make others reliable in the sense that they are predictable, yet it does not give an indication of their trustworthiness. Trustworthiness requires that the trustee is not only predictable, but that he is competent to respond to what is entrusted and that he is committed to the truster or the object of trust. Order as such is silent about the aspect of commitment and motivation.

Hobbes deals with this by introducing the aspect of fear of sanctions. However, this introduces another problem. Punishment and guilt as artificial mechanisms may lead to cooperation and social stability, but are not likely to result in trust (Johnson, 1993, p. 77). Trustful cooperation cannot rest on external sanctions, as Hobbes would have it. Forced reliance rules out both trust and trustworthiness. If one is forced by whatever reason to act in a specific way it makes this person reliable in the sense that he is predictable, yet it does not make him trustworthy, since his own intentions are irrelevant in this case. It equally rules out trust, because trust presupposes a certain extent of freedom to abstain from trusting. For example, if I force someone by using a gun to rely on me, he may do so, but he will not trust me.

In sum: the necessary acknowledgement of the individual raises the questions of when an individual will trust and why a trusted person would act trustworthily. However, the specific answer given by Hobbes shows that, for matters of trust, this issue cannot be 'solved' with the introduction of coercion by the creation of a sovereign who demands the surrender of all power. This also holds for proposals by other authors who modify the idea of the sovereign, but still aim to build trustful behaviour by introducing sanctions. As long as the issue is restricted to the question of how individuals can cooperate, the introduction of coercion and sanctions might be a legitimate answer. Yet as an answer to the question of how individuals can trust each other, the introduction of fear will not do because it disregards an element that should be added to the recognition of individuality: agency in terms of freedom and intentionality as conceptual requirements of trust. This is the topic of the next section.

4.5 The conceptual requirements of freedom and intentionality

The Hobbesian contract shows that recognition of the other as an individual who can act independently of the social, political, or religious community is essential, yet is no more than a start. His emphasis on coercion and sanctions can be efficient tools if the aim is public cooperation, they are not if the aim is to establish trust. The reason why coercion and the threat of sanctions fail to contribute to trust is the fact that it limits the agency of an individual or institution. By forcing someone to act in a certain way it will be easier to rely on him. His freedom is restricted and consequently he will act predictably. In the introduction of this chapter however I mentioned that trust is a way of 'handling the freedom of other human agents or agencies' (Dunn, 1988, p. 73). The use of coercion and the introduction of sanctions can make agents and agencies more reliable, but not trustworthy. We can speak about trust only if the truster and the trustee consider each other as free agents who act intentionally. To substantiate this claim, the aspects of freedom and intentionality as constitutive elements of agency are discussed in further detail.

4.5.1 Agency as a necessary condition for trust

Recognition of the other person or institution as capable of acting is crucial for trusting and being trustworthy. If things would merely happen to people or if they would only undergo events passively it would complicate both trust and trustworthiness. If we provisionally define an act as behaviour that is directed to a certain goal and produces expected outcomes, it follows that entities that lack the capacity of agency can neither trust nor be trusted. For instance, a bridge cannot act. It's purpose or goal is evident but it does not show any purposeful behaviour. The fact that it does not collapse is not because the bridge is aware of its purpose or its goodwill. It is only an indication of the competence of its engineers and constructors. Bridges lack the ability to trust others and cannot act 'as if' certain states of affairs would not occur. One can rely on the bridge, but cannot trust it. If one relies on the bridge it is because of the expectation that the bridge will perform in the way that it has been constructed, not because one considers the bridge as competent and properly motivated. Thus, we can conclude that trust requires that both the truster and the trustee have the ability to show purposeful behaviour.

We need a more specific view on agency. Let us move on to a consideration of animals. Compared to bridges, animals do fulfil the condition of having the ability to show purposeful behaviour. Most animals are directed to a certain aim and goal and show behaviour that is directed towards this goal. They move their legs or wings in such a way that they walk or fly in order to fulfil a goal (cf. Frankfurt, 1978). The animal itself shows activity and is directed to a goal that is not directly enforced by others. In this way, they are much closer to action than the bridge. Although we can define this behaviour as activity it is not yet a signal of the agency we need in the case of trust. The diversity between the different species of animals prevents us from making general statements on trust and animals. Nevertheless, we probably would be hesitant to say that an insect can be a truster or that it can be trustworthy even though it has the ability to show purposeful behaviour. To actually trust someone we must necessarily attribute two additional characteristics next to the ability to show purposeful behaviour.

These characteristics are (a) the ability and freedom to choose the goal related to one's action among alternatives, and (b) to knowingly and willingly do or omit to do *x* (cf. Düwell, 2002, p. 156).

4.5.2 Acting in freedom

Let me start with the first characteristic. Trust presupposes both the ability and the freedom to choose a goal and to choose it among alternatives. This criterion excludes inanimate objects; a bridge cannot be truster or trustee. A bridge can perform in a way that is counter to what is expected, but it does not choose to adopt this alternative. It does not have the ability to choose at all. As a result, the bridge can neither trust nor be trusted. This equally holds for situations in which an agent has the ability, but lacks the

opportunity to act in freedom. For instance, a hostage is severely restricted in choosing alternatives with respect to several matters. Consequently, with respect to these issues trust is not possible, because he is forced to act in the expected way. His trustworthiness has become irrelevant in deciding to rely on him.

The ability to choose one's goal among alternatives does not imply that the agent should have full knowledge of all possible states of affairs. If one is confronted with uncertainty this is beyond reach. Nevertheless, the capability to choose one's goal among others is still available. For example, suppose that a consumer is confronted with uncertainty regarding the quality of a specific food product. To address this issue she has to rely on the information of experts. However, this dependence on external information does not change her ability to choose the goal related to her action among alternatives. Nor does it affect the freedom to act in a way that achieves the goal that she has adopted, based on an overall judgement of the options and opportunities. Both the ability and the freedom to choose among alternatives are essential for the truster and the trustee. If a truster is not in the position to choose among alternatives, his assessment of the trustworthiness of the other agent does not play a constitutive role. The only option he has is to rely on the other agent, whatever his idea of the competence or motivation of this agent or agency.

The significance of freedom also holds for the trustee. If a trustee lacks freedom there is no need to trust him and no possibility for him to act trustworthily. First, there is no need to trust if we do not consider the other to be free. If external forces coerce an individual to act in one specific way, trust is not necessary because we are not confronted with uncertainty. We know or can calculate based on the available information how someone will act if he is compelled to do so. Second, no agent can act trustworthily if he lacks freedom. If he can only act in the expected way his competence and motivation make no difference. The way he acts is fully as determined as the operation of a machine, on which one can rely but which one need not, in fact cannot, trust.

Thus, as Johnson accurately claims, 'to speak then, of the origins of trust is to describe the variety of ways in which agents become conscious of the freedom of others' (Johnson, 1993, p. 79).

4.5.3 Acting intentionally

The second characteristic of action highlights the intentional element of action as a conceptual requirement for trust. Creel makes this element central to his definition of action. He refers to an action as 'a unit of intentional behaviour that produces expected outcomes' (Creel, 2001, p. 160). To be an agent one needs a certain ability of reflection and deliberation that underlies one's activities. If this element is entirely absent it is difficult or impossible to speak of trust. If a person does something unintentionally, for

instance, when a quality assurance officer forgets to complete a safety check during the production process of a food product, this is not an act inspired by trust in his colleagues to make good his failing. Forgetting is not an intentional action: it is not a decision of will. The presence or absence of trust in his colleagues has nothing to do with it. It would be different if, trusting his colleagues to perform the next check properly, he deliberately neglected his responsibility. Then he would genuinely trust, since his activity is intentional. The same condition holds for the trustee. To be a trustworthy trustee presupposes the ability to respond to another's trust. A trustee can have different reasons to act in line with the expectation of the truster. Trustworthiness, however, presupposes that the trustee responds in his behaviour to the expectations of the truster. This response is by definition the result of an intentional action. Therefore, Frankfurt's claim that 'a person is active when it is by his own will that he does what he does' (1982, p. 271) holds for both the truster and the trusted person when we speak about trust.

In short, it is a conceptual requirement that both the truster and the trustee are agents in the sense that they show behaviour that is intentionally directed to a certain goal and produces expected outcomes, and have the ability and freedom to choose among alternatives. Consequently, every proposal to restrict the freedom or intentionality of agents may help to manage uncertainty, but will never contribute to trust.

4.6 Rational agents and the motivation to trust

Now we can return to the question that led us to analyse the Hobbesian approach: What motivates an individual to trust or to act in a trustworthy manner? The additional criteria of the ability and freedom to choose alternative goals and of intentionality illustrate some flaws in Hobbes's approach and restrict the number of relationships that we can define as trusting relationships, yet they do not yet indicate why these free agents would be motivated to trust one another, especially since trust implies vulnerability. In other words, why would a truster make himself vulnerable by trusting another, and what is the motivation to trust or to respond to trust?

Rational choice theories provide one of the most influential answers to this question. Roughly, the answer is that one will trust another if and only if it is in one's interest to do so, i.e. when trusting enables the truster to achieve some pursued goal. Trust enables an agent to act in line with personal preferences. For that reason it can be worthwhile to accept the vulnerability implied by trusting. Because of the significance of this account in the current literature on trust and cooperation, I discuss it at some length in this section.

4.6.1 Some background on rational choice theories

To understand the answer given by the theory of rational choice it is important to pay attention to the common background of current rational choice theories. To begin with, the individual is considered as 'the elementary unit of social life' (Elster, 1989, p. 13). Actions are always considered from the perspective of the single individual,[63] and consequently rationality is considered as a property of individual action. If we want to understand what is rational to do and to expect, we have to start with the analysis of individual choices.

This leads to a second shared characteristic. Individual action is considered as instrumental and goal oriented. It is instrumental in the sense that individual choices aim at the realisation of one's preferences and goals, and are structured by the available resources and possibilities. Consequently, an action is described as rational if and only if it is 'the product of the agent's belief and desires, and is directed to securing what the agent most wants, all things considered' (Hollis, 1998, p. 45). Furthermore, the rationality is instrumental to the pursued end in the sense that the rationality of the goal is not questioned (cf. Hollis, 1994, p. 118). This definition of rationality is primarily descriptive and positive. It is descriptive, because it aims to describe human behaviour. The definition is positive, because the aim is to explain human behaviour as it occurs rather than to steer individual behaviour in order to fit in the theory's concept of rationality. It does not have prescriptive or even normative aspirations. The direction is the other way around. The theory assumes that a rational individual shows goal-oriented behaviour, and acts in a way that suits his preferences in the best way given the available options for action. The individual is considered as inventive in the sense that he will not automatically act in line with social conventions, norms or values. The main considerations that play a role in deciding how to act are steered by the individual's belief and desires. Hence, theories of rational choice commonly use a view of human behaviour as completely autonomous: there are no incentives for action other than the internal deliberation on beliefs and desires given the available opportunities (cf. Van Hees & Vromen, 2002, p. 9).

Third, this interpretation of rationality as goal oriented is in most versions of rational choice theory related to the view of the individual as 'a maximizer, who will settle for nothing less than the best' (Simon, 1959). The rational individual always aims to maximise the pay-off of an action and to minimise any possible losses given his own

---

[63] In the social sciences this assumption is linked to the discussion on methodological individualism and ontological individualism. Methodological individualism argues that for collective actions and processes only an individualistic explanation is possible. Ontological individualism makes the same claim. However, in addition it argues that only human beings exist and that society is only a product of individual human action. Consequently, the epistemological claim is that 'all knowledge about society derives from knowledge about individuals' (Udehn, 2002, pp. 485-486).

preferences. Acting rationally implies that one chooses the action that leads to the optimal realisation of one's preferred target. Every decision is taken in the light of the goals that one pursues,[64] and the available ways to achieve these goals. In this process, ideally, the individual has full information, fully ordered preferences, and a perfect internal mechanism to compute the risks and benefits of each option of action (cf. Hollis, 1994, p. 116). Consequently, it is possible to formalise the relation between the preferences, the specific context of alternatives, and the pay-offs and losses of the available actions in mathematical models. Such models can be found in many versions of this decision theory.

Fourth, from the above characteristics it is derived that the individual actor is essentially self-regarding. An actor who is confronted with a dilemma of action will act such that his own interests are promoted in the most favourable way. Preferences and goals of others play a role in his deliberation only as far as they influence the agent's possibilities to realise or maximise his own goal. Every available option of action is evaluated from the perspective of one's own interests. Nevertheless, it has to be noted that this does not inevitably lead to the conclusion that maximising preferences implies a monopoly of self-interest as motivator for actions. Mansbridge shows that within rational choice theory, it is possible 'to abandon the claim that self-interest is the sole operative motive', and that it is possible 'to work with any motive, provided only that the decision maker maximize and be consistent' (1990, pp. 20-21). Nonetheless, the emphasis on self-interest as the main incentive for actions can easily be recognized in many versions of this theory.

In short, we have indicated four characteristics of rational choice theory: (a) individual decisions are taken autonomously, i.e. decisions are based on individual deliberation only, (b) these individual considerations are determined by one's own beliefs and desires that lead to preferences and goals, and (c) the individual acts rationally if and only if he acts in a way that he maximises these preferences and goals, and (d) individuals are mainly self-regarding. With this background, we can analyse situations of cooperating under uncertainty starting out from the question of why an individual agent would make himself vulnerable.

4.6.2 Dilemmas of cooperation in onetime events: Not a matter of trust

Rational choice theory can provide a clear answer to the question of what is the most rational thing to do when someone is confronted with a situation in which he has to rely on one or more agents. Even when confronted with uncertainty, the individual acts rationally if and only if he acts in such a way that he maximises his preferences and

---

[64] This does not necessarily imply a complete self-centred focus. Acting upon one's own preferences may also entail that one reckons with preferences of other agents, e.g. the preference to care for others.

achieves his goals. Thus, the aim is to search for ways of action that enable the individual to maximise his preferences and achieve his goals. In finding these ways of action, game theory is often used as an analytical device.

Game theory deals with the calculations we have to make if we are confronted with a range of choices for action. It models problems of cooperation in terms of a game with two or more agents who are interdependent and have to decide under uncertainty. One of these models is the well-known Prisoner's Dilemma (PD).[65] This dilemma is often considered to be a good example of problems regarding cooperation in a situation of uncertainty. It contains both the elements of non-simultaneous or sequential exchange[66] and a lack of information. In short, the PD can be outlined as follows. Two persons have committed a serious crime (Tucker, 1980 [1950]). The police have serious suspicions against the two men, yet the evidence is not sufficient proof to have them convicted. Consequently, they are held separately by the police so they cannot consult each other. In this situation, they are offered a deal:

(1) If one cooperates with the police and confesses his crime, then he will be given a reward of one unit (+1) and the other suspect will be fined with two units (-2).
(2) If both confess, each will be fined with one unit (-1).

Additionally, both have good reasons to believe that if neither of them confesses both will be released, because then the police does not have sufficient evidence against them (0).

This situation has been analysed as a game by two parties in which either of the two agents has two options: (a) to cooperate with the other prisoner, and refuse to confess or (b) to defect, and thus confess against the other. This results in four possible outcomes:

A. Both prisoners (P and Q) confess. In this case, they both are convicted with a moderate sentence.
B. P confesses, but Q does not. If this happens, then P is released, since he has collaborated with the police, and Q is convicted, and gets the maximum sentence.
C. Q confesses, but P does not. This results in the same situation as above, except that now Q is released and P receives the maximum sentence.
D. Both P and Q do not say a word and do not confess. In this case, they cooperate with each other. There is not enough evidence for a conviction, so they will be released.

---

[65] A.W. Tucker introduced the name of this dilemma in 1950 in a paper document from Stanford University (1950/1980, pp. 101-103). Merrill Flood and Melvin Drescher, however, first proposed the dilemma in this game that demonstrates the difficulties of cooperation, in 1950 (cf. Flood, 1958, p. 12). Howard Raiffa independently conducted experiments with the Prisoner's Dilemma at the same time. His findings, however, remained unpublished until 1992 (cf. Holt & Roth, 2004, p. 4000).
[66] Non-simultaneous or sequential exchange is the situation in which the involved agents do not act at the same time, but act one after another. The temporal order is that the second agent acts *after* the first agent has made his move (cf. Ross, 2006).

These four actions lead to the following pay-off matrix:

		Q	
		confession	no confession
P	confession	-1, -1 (A)	+1, -2 (B)
	no confession	-2, +1 (C)	0, 0 (D)

From the perspective of P the most desirable outcome is (+1, -2) and the least is (-2, +1). So the preference ordering for P is B > D > A > C. Game theory, however, teaches us that P's choice is not merely depending on his own preferences, but also on his expectations towards the behaviour of Q. In order to be free, he has to trust that Q refuses to confess. Yet, why would Q do so? P knows that the preference order of Q is C > D > A > B. Thus, Q will initially choose for confession, because this is necessary for option (C). If P does not know anything about Q and did not make any agreements, the most rational thing to do is to confess, since this strategy dominates the strategy of not-confessing. This implies that he does not cooperate with the other prisoner. For Q we can draw a similar conclusion. He initially prefers not to confess, but he cannot rely on the cooperation of P, consequently he also ends up at option (A). Hence, the outcome tends to a situation in which both choose to confess, although it is for neither of them the best outcome. However, this is the only part of the preference order that they have in common. For both hold that D > A is the dominant strategy. They only can predict that confession is the strategy they will both choose, unless they made a previous agreement to act otherwise.

This outcome illustrates the remarkable result of being 'rational'. Both agents act strictly rationally, yet they do not manage to maximise their interests, i.e. to choose the most profitable strategy option (D).

From the perspective of trust we can draw another conclusion: there is no room in this model for trusting each other. Even though both agents are confronted with uncertainty, the preference orders of both agents tell them not to cooperate. P arrives at his evaluation of the most rational strategy of action completely independent of any trustful expectation towards his partner in crime. Hence, the PD is not a problem of trust. For a rational prisoner, trust is completely irrelevant. Neither prisoner has reason to rely on the other. Horsburgh concludes, 'the weight of self-interest inclines us to distrust rather than to trust not only if there is evidence of untrustworthiness but also if there is little or no evidence for or against trustworthiness. Thus, trust usually has to make its way against an initial current of distrust' (1960, p. 351). This seems a rather tragic situation: for both prisoners, as it is for all other rational agents in similar situations; it is relevant to act on trust, yet it is prudent to be hesitant to confidently rely on other agents.

This conclusion also holds for the proposals in which the PD is applied as a model for collective action. For instance, if it is used as a model to approach multi-participant

problems in the food sector, such as problems of cooperation with respect to the production and consumption of sustainable food products. The problem is that although empirical research has shown substantial support in society for more animal-friendly and environment-friendly products, the market for sustainable products like organic and animal-friendly foods is rather small. This can be explained in terms of a problem of cooperation. Consumers have to cooperate with each other, because as many consumers as possible have to buy products from sustainable production systems to maximise the contribution to sustainability. This situation can be framed in terms of a large-number Prisoners Dilemma.[67] An actor either can buy a sustainable product and thus choose to cooperate with others to contribute to sustainable ways of production, or he can defect and buy another, non-sustainable product. Again, this results in four options:

A'. P and all other consumers (Q-Z) cooperate and buy sustainable products. This contributes optimally to the overall aim.

B'. P cooperates and buys sustainable products. He contributes to the overall aim, but agents Q-Z do not buy the products. Consequently, the result of the action of P is hardly noticeable.

C'. P defects and does not buy the sustainable products. However, most of the other consumers do. P will save money in this case, while the negative effect of his defection on the total result is only very marginal.

D'. Neither P nor Q-Z buy sustainable food products. They all choose the defect option. All save money, but the pursued goal is not reached.

It is possible to draw the following preference order. For P the order is C' > A' > D' > B'. The most rational strategy for action for P is not to invest his money in the expensive sustainable products if all others take care for the overall aim. This is the so-called free-rider position: if a majority takes care of public issues such as sustainability or safety, then the input of one individual does not really make a difference, for neither good nor ill. However, the problem is that Q-Z is not one collective actor, but the total of all individual actors. For instance, agent Q, R or S has a preference order in which P and all other agents buy the expensive sustainable products, while she still buys the cheaper nutrient. Hence, for each of them the most rational strategy is not to buy the sustainable product as long as they cannot trust the other consumers to do the same (cf. Diederen, 2003). And they cannot trust them to do so since it is for them not the most rational thing to do either. This results in a vicious circle in which all may agree on the importance of a more sustainable way of food production, but no one will buy the actual products. Again, we end up in a sub-optimal situation. This looks very much in line with the traditional Prisoner's Dilemma.

We can conclude that the model of the Prisoner's Dilemma, combined with the rational choice ideas on agency, does not provide a real answer to the question of why someone would make himself vulnerable by trusting another. Even before trust becomes a topic, rationality tells the individual not to rely on another at all.

---

[67] Cf. the environmental dilemma in Pellikaan & Van der Veen, (2002), Chapter 1.

4.6.3 Dilemmas of cooperation in iterated events

Models such as the Prisoner's Dilemma help to structure problems of cooperation, but they have limitations. The PD only gives information on cooperation in terms of one-time events. However, we often cooperate with others in situations in which it is likely that we will meet each other again, or in which it is likely that we have met before.

This context underlies the so-called iterated games. In such games, not only your own preferences and those of the other agent play a role, but you also reckon with the fact that you will meet each other again or that you have a history with each other. In those situations conditional cooperation may arise if both will profit from cooperation. If agent P knows that cooperation contributes to mutual advantage, he may have a reason to cooperate, because future engagements are clearly an incentive to formulate expectations towards the other party. For instance, suppose I buy bread at the local baker's and he sells me bread that is no longer fresh. I will not notice it the first day, but on the second day it becomes mouldy and I have to throw away half a loaf of bread. As a single event, the strategy of the baker is a rather effective one. In this way, he can maximise his profits. However, it is unlikely that I, as a customer, will return. In a single game, the fact that I will no longer be one of his customers is no problem, because the long-term relationship is beyond the scope of the model. However, in practice the baker also has an interest in keeping his customers so he had better sell fresh bread. He would gain more from a default position in a one-time game, but in the end cooperation is more profitable. Because both the baker and I know that we will meet again, I have a reason to rely on him and to buy his products. The aspects of history and future enable me to form expectations. Consequently, it becomes possible to formulate positive expectations with regard to the competence and motivation of this baker, in other words, trust becomes possible.

In the context of iterated events, the aspect of future engagements is pivotal for the possibility of trust. However, it also is a vulnerable aspect. The trustee will only continue to do what is expected of him as long as there is an ongoing relationship. When the cooperative relationship ends, the last time that both agents meet is like a single game. This shows the weakness of this situation, especially when the end of the relation has been fixed. For instance, if P and Q know that case 20 is the last one, the situation in case 20 is the same as a one-time event of cooperation. Thus, it is rational for Q to defect. However, since P knows this too, it is rational for him to defect in case 19. It does not require much insight to see that this will result in a serious regression and that it shows in the end that it is rational not to start at all.

Hence, there has to be something to halt this problem of regression. That is also the point that surfaces in Hume's well-known problem of the farmers. Hume (2007, [1739-40], 3, II, V) writes:

Your corn is ripe today; mine will be so tomorrow. It is profitable for us both that I should labour with you today, and that you should aid me tomorrow. I have no kindness for you, and know that you have as little for me. I will not, therefore, take any pains upon your account; and should I labour with you upon my own account, in expectation of a return, I know that I should be disappointed, and that I should in vain depend upon your gratitude. Here then I leave you to labour alone; you treat me in the same manner. The seasons change; and both of us lose our harvests for want of mutual confidence and security.

This example can be perfectly analysed as a trust game in which there are two competing options: trust or distrust (cf. Hollis, 1998; Lahno, 2002a). The problem in this situation is that the farmer who helps his colleague has to work in advance. He may expect the other farmer to help him, but he has in no way a guarantee. The only reason is that *if* both cooperate it will be profitable for both of them, but this is only a rational expectation in an iterated situation. Only in the context of a common future the first farmer will be motivated to help his colleague in return; otherwise it is not rational for him to do so, since it does not suit his preferences in the best way. However, he does not have an assurance with regard to future interactions. There is no convincing argument why the first farmer should take the risk of helping the other farmer. The only rational position is not to adopt a cooperating attitude even though it would be in their mutual interest. Thus, as long as there is uncertainty about the iterative character of the cooperation, we are back at the problem that we also face in the Prisoner's Dilemma. The problem is that we need additional arguments to start up cooperation. Hume tries to approach this problem by underlining that we are not completely self-regarding, but that sympathy is also part of our human character. This makes cooperation possible; however, it does not solve the problem entirely. Sympathy will change one's initial motivation to cooperate, yet it still does not change the problem of the rationality of cooperation. One cannot put sympathy on the forefront in all situations. Hume – quite aware that all lay load on a willing horse – introduced artificial virtues like justice, which entail that we do not only conceive a situation from our own perspective, but also take an impartial point of view. However, the introduction of such artificial values and the emphasis on the element of sympathy in human behaviour implies that one leaves the field of strict rational choice theory in which it is argued that we are motivated by no other motivators than our own preferences, which are primarily self-regarding.

Thus, we can conclude that the elements of previous encounters and a common future in iterated events of cooperation can motivate agents to trust and to act in a trust-responsive way, but are not a necessary condition.

4.6.4 Can trust be rational?

The theory of rational choice can provide an at first sight clear answer to the question what is the most rational thing to do if we are confronted with a situation in which we have to rely on other agents. The above analysis of deciding under uncertainty, however,

complicates the answer, because it does not immediately include trust. Only in the case of iterated games and in the context of possible future engagements trust is not excluded as a rational option.

This raises the question of whether trust fits within the view on rationality as operative in rational choice theory. Hollis (1998) presents two answers to this question. On the one hand, he states: 'trust grows fragile when people become too rational' (p. 2). If rationality is defined in terms of performing that action which best suits one's preferences, the rational agent will act like the agents in the Prisoner's Dilemma. With no other master than rationality, the individual can only be expected to act in a way that maximises his interests. Consequently, if trustful expectations are not identical to what the trusted person would have done anyway, given his rational aim of maximising his interests, trust is no more than an irrational belief in the goodness of the other party. On the other hand, Hollis states that 'trust grows fragile when people are not rational enough' (p. 2), because also from a rational choice perspective it is clear that a cooperative attitude may contribute to mutual advantage, and that it can be perfectly rational to trust as strategy that best suits your interest.

These two possible answers to the question of whether trust fits within the view on rationality of rational choice theory lead me to the statement that the theory neither *excludes* nor *includes* trust. Let me start with the claim that it does not exclude trust.

From the discussion of the Prisoner's Dilemma, the suggestion may arise that trust is not compatible with instrumental rationality. This view seems to be confirmed by the literature. Bacharach (2003, p.1) for example notes that they even seem mutually exclusive. To the question as to what makes it rational to trust or to respond to trust the theory of rational choice answers 'nothing'. Vaarkamp (2001) and Lahno (2002a), however, argue that this would be too hasty a conclusion. First, Vaarkamp shows that within game-theoretical models some kinds of cooperation remain unexplained if we exclude trust from the deliberation. Traditionally, it is argued that only direct reciprocity is the sufficient condition for cooperation. However, it has been demonstrated that indirect reciprocity could do the job as well (Nowak & Sigmund, 1998). This means that cooperation also emerges in cases where the accompanying reciprocity may not be directly clear at the very moment of the decision to cooperate or when it is not even directed towards you personally. The explanation of the results of Nowak & Sigmund with respect to the relation between indirect reciprocity and cooperation, however, was hardly self-evident and after a hundredfold repetition the fact that indirect reciprocity leads to cooperation turned out to be just one of the possible outcomes of the simulation model (Vaarkamp, 2001, Ch. 5). This suggests that the traditional view – that only direct reciprocity is the sufficient condition for cooperation – is the only position that can be defended. Vaarkamp, however, modified the 'indirect-reciprocity' simulation model by introducing the concept of trust, which he defines as someone's belief in the trustworthiness of another (p. 61). Indirect reciprocity can be a condition to make

cooperation emerge, provided that trust is included in the model. The personal experience with trust has a profound influence on the willingness to cooperate and, to a lesser extent, even the experience interaction with others (Vaarkamp, 2001, pp. 23, 63). He argues that trust is a *sine qua non* for cooperation that, even if rationality is modelled in a strict sense, cannot be reduced and is a crucial part of human life (p. 26). In Vaarkamp's view trust cannot be restricted to the field of belief and hope, but is part of the rational answer to the central question of cooperation (pp. 26, 69).

This argument illustrates that trust can be explained rationally. Instrumental rationality, then, does not exclude trust. In an ongoing relationship a truster may have very good, i.e. rational reason to trust an individual or institution. Lahno shows that it is possible to extend the trust game as used in game theory in a way that reckons with fairness and altruistic motivation, but still remains within the scheme of instrumental rationality (2002, pp. 62-67). If the agents are motivated by reasons of morality the choice for a trustful action can be perfectly rational according to the standards of instrumental rationality. If P knows that Q is fully motivated by reasons of altruism, then this has a direct influence on what is the most rational thing to do. For instance, if P is confronted with uncertainty about the aims and goals pursued by Q, he also knows that the advantages and disadvantages of either of the options of actions for him will be taken into account in the reaction by Q. Thus, the choice to act trustfully can be rational within the scheme of the theory of rational choice. Lahno, however, remarks accurately that this does not say anything about whether a trustful action is also an action *based on* rational beliefs (p. 72). It nevertheless shows that the element of trust is not incompatible with the definition of rationality in terms of maximising one's preferences. Trust can be a rational attitude in given circumstances.

From this conclusion, however, we cannot infer that instrumental rationality includes trust. Trust can be rationally explained when someone already trusts. Nevertheless, if we ask whether it can be rational to enter a trusting relationship Bacharach (2003) is correct: nothing makes trust rational. The fact that a trustful attitude can rationally be explained does not yet explain why someone takes upon himself the vulnerability involved in starting to trust. Therefore, I conclude that a rational choice perspective does not include trust. It can be rational to keep trusting, but to consider the start of a trusting relationship as rational we need additional criteria.

Hardin (1993; 2002; 2006) has presented an additional criterion which shows that it can be rational to adopt an attitude of trust. He takes the assessment of the interests of the other agents as central in coming to trust. He argues that trust is rational if the truster knows that it is in the interest of the trusted agent to take the interests of the truster into account. This idea of trust as 'encapsulated interest' is in line with rational choice theory, but can go beyond the problems of cooperation as presented in the Prisoner's Dilemma and the trust game. In the next section, Hardin's account is presented and discussed in further detail.

4.7 From the assessment of interests as the motivation to trust to the demand of reasonableness

What motivates agents to trust and to respond to trust? The answer provided by Hardin can be summarised in one term: 'interest'. 'Interest is one of the best and most useful of internal motivations' (Hardin, 2002, p. 52). Emphasis on the importance of interests 'fits a centrally important class of all trust relationships' (Hardin, 2002, p. 1, also Coleman, 1990; Gambetta, 1988b). With this answer, his view is largely in line with the fundamental assumptions of rational choice theory.

4.7.1 Trust as 'encapsulated interest'

Hardin starts his account of trust in line with the rational choice view that human action is driven by the rational pursuit of interests. 'for many … trusting relationships, the whole point is likely to be interests' (2002, p. 4). From this starting point, he conceives problems of trust as merely epistemic problems by defining them in terms of the truster's assessment of the interests of the trustee. Knowledge about the other agent's interests tells you something about the trustee's incentives to act trust responsively (2002, pp. 129-131). However, this assessment itself says nothing about whether the other agent is trustworthy. Someone can be trustworthy if and only if it is in his interest to take your interest seriously. From the assumption that human behaviour is motivated by interests you only have a reason to rely trustfully on the other person if it is in his interest to act in the expected way. Otherwise, it would be no more than a guess. Thus, according to Hardin, we can speak about trust, if your interests are 'encapsulated' in the interest of the trustee (1993; 2002; 2006). We can trust another agent if 'I think it is in your interest to take my interests in the relevant matter seriously' (2002, p. 1).

The position of encapsulated interests does not imply that both persons have the same interests. It only assumes that it is in the other person's interest to consider the truster's interests, e.g. because one has good reasons to remain colleagues or business partners (Hardin, 2004, pp. 6-7). This shows that Hardin uses interest in a rather broad way. He defines it as merely 'a proxy for all that you might take into account on my behalf' (2002, p. 24).

The advantages of this account can be illustrated when we return to Hume's example of the two farmers, who are faced with the dilemma of whether it is rational to trust each other and to cooperate in order to harvest theircrop. From Hardin's perspective, it can be perfectly rational for Hume's farmers to trust and cooperate. The problem in Hume's example is that the farmer who helps his colleague has to work in advance. He has no guarantee that the other farmer will help him. He lacks a convincing argument why the first farmer would take the risk of helping the other farmer. The only rational position is not to adopt a cooperating attitude even though it is in their mutual interest. Hardin can

take a step further, because of his start at the mutual interest. If it is in their mutual interest to cooperate, then it is likely that the trusted farmer will take the interests of his fellow farmer into account. For the truster this is a good reason to confidently rely even though he still may run a risk in trusting. Thus, Hardin convincingly shows that entering a trusting relationship can be rational from a rational choice perspective.

4.7.2 Trustworthiness and interests as the main motivator to act

The central position of interests as the only motivator to act in the theory of rational choice is also essential in Hardin's account. The empirical claim that personal interests are strong motivators can hardly be denied. Can a person who is motivated only by his personal interests be a genuinely trustworthy agent, i.e. someone who is competent and motivated to act in a trust-responsive way? Essentially self-regarding people are to that extent reliable. I can formulate expectations with respect to their future behaviour. But this does not make an agent a trustworthy partner; his motivation to act is not influenced by the cares and concerns of the truster.

In answer to this problem, Hardin can show that in his account the rational, interest-driven agent is genuinely trustworthy. His argument consists of three steps. First, rational choice tells us that an action is only rational if the action is motivated by self-interest. Thus, to rely on an agent is only rational if it is in his interest to act in the expected way. For that reason, one starts with interests. However, interests as motivation to act in the expected way do not yet make someone trustworthy. A second step is needed here. One should not only be motivated by interests, but it should be in the interest of the trusted agent to take your interest into account. My trust is rational if it is in the other's interest to consider your interests. However, this still includes the option that the other agent is not motivated by *your* interest, but that there is only a similarity of incentives and interests. Someone can be thought reliable because in view of our mutual interests I can reasonably predict how he will act. Trustworthiness is something more. Hardin explicitly, states, 'our merely having the same interests with respect to some matter does not meet the condition of trust as encapsulated interest' (2002, p. 4). If one is genuinely trustworthy, he deliberately takes your interests into account because they are *yours* (Hardin, 2002, p. 11). Trustworthiness is not only about sharing an interest, it also includes that there is a relationship between the truster and the trustee and that the trustee values the continuation of the relationship with the truster. Only if the trusted is concerned with the truster's interests, then, according to Hardin, one can speak about trust.

From a rational choice perspective, Hardin's account is a clear step forward. It shows that it is possible to act rationally and to trust or act trustworthily at the same time. Nonetheless, Hardin's focus on interests as the only motivator for action raises the question of whether it can be rational to expect an agent to respond to trust if it is not in

his interest to do so. This is not a hypothetical question. We entrust objects to agents and agencies, and we expect them to take care for these objects even if it is not in their self-interest to do so. For instance, I trust the government to regulate that the health claims of pharmaceutical products are scientifically sound. I expect them to take care of this, and entrust them with this matter even if it were in their interest to do otherwise. I do so, because I believe that they should be motivated by my interest and by my vulnerable position, and that I am entitled to expect this of the government. The question is whether this expectation can be rational or whether we are left with the 'fragile gum of mutual self-interest' only (Hollis, 1998, p. 109).

4.7.3 Trustworthy in spite of self-interest; The ability to be trustworthy when it crosses one's interests

In trusting we do not merely expect *that* another agent will act in a certain way, it also includes that we sometimes expect something *of* another (Hollis, 1998). These expectations are not based on existing routines or predictable patterns, but on beliefs about how the trusted agent should be motivated. Living up to these expectations sometimes implies that acting trustworthily 'crosses one's interests' in the sense that the trustee will be required to forego particular advantages of his position as a trustee (cf. Rawls, 1958).[68] Given the emphasis in the theory of rational choice on interests as the only motivator for action and man as mainly self-regarding, one may wonder whether it can be rationally expected of an agent or agency to act in a trustworthy manner if this implies that it crosses one's interests.

My answer is positive, but implies a critique on the theory of rational choice. First, I do not agree that all personal aims are self-interested (cf. Scanlon, 1998, pp. 132-133; Blackburn, 1998b). We can be motivated to care for other people, not only because it is in our personal interest to do so, but also because we consider it a good thing to do so (cf. Van den Hoven, 2006, pp. 127-28). Hardin also states that it is 'a mistake to suppose that my well-being is merely selfish … my well-being will often depend on my sharing intentions with you to do things with or for you' (2002, p. 23). We have altruistic or other selfless motives to trust and to act in a trustworthy manner that cannot be reduced to self-interest without losing the essence of the motivation that underlies this behaviour. This shows the possibility that someone can be motivated to act in a trust-responsive way even if it is not directly in his self-interest to do so.

Moreover, the emphasis on the seemingly self-regarding character of our aims presumes that there is a clear line between what counts as self-centred motivation, and

---

[68] Rawls uses the expression 'crosses one's interests' in the context of justice. He argues that justice does not require of anyone that he sacrifices his interests, but that 'the duty of fair play will often cross his interests in the sense that he will be required to forego particular advantages which the peculiarities of his circumstances might permit him to take' (Rawls, 1958, p. 181).

120

what can be considered as altruistic behaviour. This line, however, is very difficult to draw. We are often confronted with a mixture of motivations. Williams (1988), for instance, has proposed a distinction between four sorts of motivation. He distinguishes (a) egoistic micro-motivation, (b) egoistic macro-motivation, (c) altruistic micro-motivation, and (d) altruistic macro-motivation. With this scheme, it is possible to differentiate various motives for depending on others. The motives in a Hobbesian account of cooperation for example are categorised as egoistic macro-motivation. 'The motivation to cooperation is egoistic, since it lies essentially in the fear of the sanctions of the sovereign' (1988, p. 10). This scheme is elucidating as an analytical device, yet what counts as macro, and what as micro, what as egoistic, and what as altruistic is still subject of discussion.[69] A food company may act trustworthily because it genuinely cares for the health and well-being of its consumers or because it considers it as part of its corporate social responsibility. Nevertheless, it certainly is in the company's interest too, which may also be a motivation to act trust responsively. In this case, it would be inappropriate to ask for the 'real' operative motive of the company, since both motivational elements play a role. It is not an either/or situation. It does not help us much to insert partitions in the mixture of motives, even if we could. Fortunately, says Baier,

> we do not need to wait until we have expert insight into human motivation, and can recognize 'altruistic' motivation should we encounter it, before we can design schemes of cooperation that will encourage both trustworthiness and trust, and can judge the comparative success of different schemes. (Baier, 1991, p. 156)

Finally, if rationality includes reasonableness it can be rational to expect an agent to act in a trustworthy manner even if this implies that it crosses his interests. One is reasonable if one is 'ready to propose principles and standards as fair terms of cooperation and to abide by them willingly, given the assurance that others will likewise do so' (Rawls, 1993, p. 49; also Scanlon, 1998). Trust among rational agents is possible, but they 'lack the particular form of moral sensibility that underlies the desire to engage in fair cooperation as such, and to do so on terms that others as equals might reasonably be expected to endorse' (Rawls, 1993, p. 51). From this perspective, reasonableness has something to add because it consists of the ability to take the interests of others into account, and to respect the rights of other persons. This, however, is often considered to conflict with rationality regarded as a self-interested means–end calculation. The ability to take different perspectives in a situation and the capability of caring about the other

---

[69] On top of this, Williams concludes that none of the possible combinations of sorts of motivation provides rational persons with motivations to become a 'dependent party' or to respond to trust (p. 13). In the same volume in which Williams's article was published, Bateson stresses the problems of a distinction between altruism and egoism and the usefulness of the concept of altruism. He, however, addresses the issue from a sociobiological approach to cooperation in evolutionary terms of explanation (1988, pp. 14-30).

agents' interests are considered incompatible with the idea of instrumental rationality that is held to characterise efficient, self-interested action.

Gewirth (1983) denies that rationality and reasonableness necessarily conflict, and argues that it can be rational to be reasonable. He does not deny that self-interest may conflict with the interests of others, but he starts with the common background of rationality and reasonableness in reason (p. 227). His article shows that being unreasonable, i.e. not respecting the rights of all (including my own), implies self-contradiction (p. 228). He defines freedom and well-being as necessary conditions to act in order to achieve intended aims. Thus, these are necessary conditions to act rationally, even if rationality is understood in line with the rational choice theory. From the argument that (1) the agent must have freedom and well-being, he deduces the claim that (2) the individual agent has rights to freedom and well-being. If this claim right is denied, then it does not hold that (3) all others should not interfere with his freedom and well-being. If this is the case, then (4) it can be permissible that other persons interfere with his freedom and well-being. If this would be permissible, then (5) it may be permissible that he has no freedom and well-being. Accepting this last claim (5) implies a contradiction of the first (1), and because (5) follows from the denial of (2) 'every agent must reject that denial' and accept the rights to freedom and well-being (pp. 230-231). This *Principle of Generic Consistency* (PGC) shows that it can be perfectly rational to act reasonably and to accept the rights of others even if this would run counter to one's self-interest. In fact, one must necessarily accept these rights for reasons of rationality on pain of self-contradiction (pp. 232-233).

In summary: humans can be motivated by reasons other than self-interest, and rationality does not rule out reasonableness. On the basis of these two arguments it can be rationally expected of an agent or an agency to act in a trustworthy manner even if this crosses the own interests. This is a crucial step. Trustworthiness does not necessarily imply that individual interests are sacrificed. After all, in many cases it can be fully justifiable that a trustee benefits from responding to trust. Nevertheless, acting trustworthily conflicts with the view on the truster as an instrument to achieve one's aims. The truster has to be recognised as a full participant in the relationship. Therefore, we have to add an additional element to our discussion of trust. This is the element of adopting a 'participant attitude' towards the other agent.

4.8 A participant attitude

The discussion in this chapter has shown freedom and intentionality as a conceptual *sine qua non* for trusting one another. Moreover, the last section illustrates the need of one additional criterion. The truster and trustee should take a specific attitude towards one

another. This attitude is very close to Strawson's notion of 'an attitude of involvement or participation in a human relationship.'

4.8.1 Participant attitude

In 'Freedom and Resentment' (1974) Strawson presents the 'participant attitude' as distinguished from an 'objective attitude'. He argues that though we normally have feelings of resentment as a reaction to injury or indifference, we do not always react that way. In those cases, we take an objective reactive attitude towards the other agent. If you adopt such an objective attitude towards another human being, then it implies that you see the other 'as an object of social policy; as a subject for what, in a wide range of sense, might be called treatment; as something certainly to be taken account, perhaps precautionary account, of; to be managed or handled or cured or trained; perhaps simply to be avoided' (p. 9). You do not see him fully as an individual who has a freedom to act. The same applies to individuals who are inhibited by abnormalities or by immaturity, like psychiatric patients or children. Strawson claims that some essential features of interpersonal relationships are lost if the objective attitude is adopted. With these persons, you cannot reason. 'You can at most pretend to quarrel, or to reason, with him' (p. 9).

This also holds for the ability to trust (cf. Holton, 1994; Lahno, 2001; 2002a). We have seen that trust requires the recognition of individuality and of the agency and freedom of a specific individual. But if one takes an objective attitude towards this individual agent, then trust is still beyond reach. Taking an objective attitude we pretend to trust another or act as if we respond. For instance, with young children you can adopt an objective attitude because of their immaturity. This implies that you can rely on them but do not genuinely trust them. Consequently, you do not feel betrayed by your one-year-old son when he harmed your trust. You may probably blame yourself, but you are not likely to resent your son. The boy does not yet possess all aspects related to full agency as described above. Nevertheless, we sometimes entrust certain objects to the care of young children. The reason to do so is not that we consider them as full participants, but that we recognise them as potential participants. In Chapter 3 it was noted that this kind of entrusting has primarily an educational purpose: we trust to evoke trust, i.e. to teach the content of trusting and of taking care (Sztompka, 1999) or, in relation to adults we use it for therapeutic reasons (Horsburgh, 1960). If we do not trust for educational or therapeutic reasons trust implies a participant attitude towards the trusted agent. We consider the other as an active participant with whom we interact. When you bet on the wrong horse you feel disappointed, but when trust is harmed you feel betrayed. You trusted because you considered the agent competent and adequately motivated. In line with Strawson it may be said that the emotion that arises when the

trustee does not act as expected is a direct sign of someone's participant stance in the interaction with the trustee.

The emphasis on the participant attitude does not imply that the truster needs special cognitive capacities. Trust is, in principle, accessible to everyone. The argument that trusting a child may not always involve a participant stance does not imply that the child cannot trust. A child does not need special cognitive abilities to trust his parents. In the case of a child, trust can be referred to in terms of the result of an (implicit) assessment of the trustee's competence and goodwill. Whether this assessment can be justified from a third-person perspective is irrelevant. It is genuine trust if the child believes his parents to be trustworthy and if they are considered as participants. Thus, he will feel betrayed if his trust is harmed.

A participant attitude, then, is an essential and necessary element of trust. Trust prevents that we see the other as merely a free individual whom we can treat as a source or tool to overcome our problems of uncertainty or lack of control. Trust requires that we consider the other agent as an autonomous person.

4.8.2 Autonomous persons

The requirement that one sees the other agent as autonomous touches upon an extensive debate on autonomy and trust. To indicate how important it is for trusting that the other is seen as autonomous I should make two preliminary remarks. First, autonomy should not be mistaken for mere independence. To think of autonomy as individual independence is to consider the actions of the individual agent as independent of the interests, beliefs and desires of others. The expectations or interests of other agents play no substantial role in deciding what is, all things considered, the best thing to do. If autonomy were defined in this way there would be a clear tension between trust and autonomy (O'Neill, 2002a). In fact, it rules out trust.

Second, autonomy should be considered as broader than the freedom aspect that has been discussed above. Autonomy is usually understood in terms of liberty, and in this form is attributed to John Stuart Mill (O'Neill, 2002a, p. 30). However, in *On Liberty* Mill uses autonomy or what he calls 'individuality' and 'character', in a much broader sense than freedom or liberty. His position is that 'the free development of individuality is one of the leading essentials of well-being', since it 'brings human beings themselves nearer to the best thing they can be' (1999, p. 102, 110). Such individuality is not just a matter of freedom; it refers to the situation in which one's desires and impulses are one's own. If this is not the case one is not really autonomous. 'One whose desires and impulses are not his own, has no character, no more than a steam engine has a character' (1999, p. 106). Here we see the importance of autonomy for trust. An autonomous person can be a truster and can be trusted. A steam engine lacks that capacity. Consequently, one can only rely on it, but will not have any trustful expectations with

respect to the engine. Liberty as emphasised by Mill is a freedom that protects individuality just like the forest allows room for its trees to grow.[70] It is obvious that this is different from mere individual independence or freedom that automatically entails the rejection of the moral demands of others or the liberation 'from all bonds', as suggested by O'Neill (2002a, p. 83). However, there is a problem with Mill's interpretation of autonomy that makes it less interesting with respect to trust. This is the utilitarian character of his theory. Because of this focus, autonomy as individuality is not considered as intrinsically good, but has to be judged according to the principle of utility. Accordingly, it suggests that autonomy, but also respect for autonomy can be overruled by the utility principle. This is problematic if we take the recognition of agents as autonomous as a requirement for trust.[71]

Now that we have distinguished autonomy from independence and freedom, we can turn to the Kantian notion of autonomy. O'Neill's discussion on trust and autonomy strongly relies on this Kantian view, which she calls 'principled autonomy' (2002a; 2002b). For her autonomy is not a matter of independence or sheer freedom; autonomy presupposes the moral equality of all agents. Mill's perspective has room for this. But in addition to Mill's idea of individuality, autonomy in the Kantian sense includes the notion of self-legislation. One acts autonomously if and only if one acts on 'principles that can be principles for all of us' such that 'we do not treat others as lesser mortals … whose abilities to share our principles we are at liberty to undercut' (2002b, p. 96). The element of self-legislation broadens the idea of autonomy beyond the idea of individuality. O'Neill emphasises that Kant does not speak about autonomous individuals, but about the 'autonomy of reason', the 'autonomy of ethics', the 'autonomy of principles', and the 'autonomy of willing' (2002a, p. 83). Autonomy understood as self-legislation does not require full independence or mere freedom, but implies a specific freedom. The freedom to act upon principles of obligation which are neither enforced by others nor derived from authorities other than the self and which can be adopted by all others.

It hardly will be a surprise that this view on autonomy as self-legislating is disputed too (cf. Blackburn, 1998a).[72] I will not go into the discussion whether ideal autonomy understood as the situation in which desires and inclinations are completely overruled by

---

[70] This metaphor is from a 1859 review on '*On Liberty*' and was quoted in the 1999 edition of this work (1999, p. 189).

[71] This touches upon the extensive literature on the question whether utilitarians can be friends and to what extent the personal perspective fits within this moral theory (e.g. Brink 1986; Mason 1998).

[72] Blackburn attacks the idea of the agent who is fully self-legislating in the sense that desires and inclinations are completely overruled by pure practical reason. He uses the metaphor of a captain (the autonomous agent) who has authority of the crew (the desires). According to Blackburn, such a captain is 'a peculiar figure, a dream – or a nightmare – of pure, authentic self-control' (1998a, p. 247).

pure practical reason is realistic or just a utopia. The relevance of the interpretation of autonomy in terms of self-government does not depend on one's position in this debate. The relevance of this view of autonomy for trust lies in the ability to see the other as a 'person', i.e. as someone who has the capacity to choose among alternative ways of living, to set goals for herself and choose values. This recognition of the other as a person is still close to Mill's idea of 'individuality' and 'character', but is not necessarily related to the value of well-being and cannot be overruled by a reference to the utility principle.

In summary: from the requirement that one should take a participant stance towards the other if one trusts another, it follows that we should recognise the other as an autonomous agent, i.e. as a person who is able to personally choose goals and values. This focus on the autonomy of the individual does not imply that an agent is fully separated from her social and historical context. But, in the recognition of the other as a person she cannot be reduced to that context. With respect to trust and trustworthiness this has some consequences that can improve trusting relationships.

4.8.3 The recognition of the autonomous person and its consequences for a truster

Two conclusions follow from the recognition of the truster and trustee as persons. First, we can conclude that trust is incompatible with exploitation. Trusting or responding to trust can never be reduced to a function to achieve one's aims in the most efficient way. Second, if we take individuals as autonomous persons with their own goals and values, the moral dimension in trusting and responding to trust surfaces. These two conclusions need some elaboration.

Let me start with the consequences of the first conclusion for the truster's attitude towards the trustee. If one sees the trustee as an autonomous person who chooses his own goals and values one has a justifying argument for the claim that trust excludes the exploitation of the truster. But the recognition of the other as an individual and as a free agent does not necessarily imply that exploitation of the trustee is problematic. Dealing with the freedom of an agent in a case of uncertainty can be done in various ways. For instance, criminals have to deal with the freedom of others and sometimes explicitly rely on agents in situations of uncertainty. Suppose that a burglar plans to burgle a house. In making this plan, the burglar relies on the working ethos of the owner of the house. He expects that the commitment to his work will prevent the owner of the house from coming home early. However, he does not trust the houseowner. The burglar considers him as an individual, who is a free agent, yet he does not perceive the owner of the house as a person with goals and values which should be respected. He exploits the house owner's goals and commitments for his own interests, and observes the owner's acts only in order to realise his own aim as efficiently as possible. In this way, he disregards him as a person because he uses him as a means to his private ends (cf. Lahno, 2001, pp.

178-179, 182). Trust is a way of dealing with situations of uncertainty and lack of control, but with the autonomy requirement, this function can never result in a situation in which we judge other's actions only by how effectively we can use them to pursue our own goals. The trustee should always be seen as an agent whose goals and values have to be respected.

The second conclusion has consequences for the truster's attitude towards the trustee. As an autonomous person the trustee is also a moral agent who has her own goals and values (cf. Horsburgh, p. 354). The trusted person has to be conceived as an agent who can act morally or immorally, can have duties and responsibilities and can be held accountable for what she does. Consequently, the truster cannot only rely on predictable patterns, but can make himself vulnerable, because he shares aims and values with the trustee. This provides trust with an extra dimension. If trust is based on predictable patterns, then one may feel fooled by the trustee if trust is harmed. Yet, if trust is based on the perception of mutual goals and values, one even may feel wronged by the trustee, because the truster considers it an obligation of the trustee to respond in the expected way. The trustee not only acted counter to the expectation, but also failed on a moral level. O'Neill shows that this last claim is a direct consequence of the recognition of autonomy. She argues that autonomy implies that '"no competent person, and none of the institutions that human beings construct, is exempt from fundamental duties' (2002b, p. 33). Autonomy is not a matter of lawless independence; it has to do with the capacity of giving reasons that others can follow and act likewise (2002a, p. 93). This expands trust from mere expectations with regard to the competence and motivation of a trustee to expectations about what the trustee is obliged to do. In other words, what one may expect *of* another agent (Hollis, 1998).

4.8.4 The consequences for a trustee

The two conclusions from the recognition of an agent as a person have consequences for the trustee as well. The conclusion with respect to the incompatibility of trust with exploitation also holds for the trustee and is in practice even more important. Because the truster is by definition vulnerable, the trustee can easily use her position in favour of her own aims. However, the truster is also a person, one whose aims and values should be respected. This entails that the trustee should take this vulnerability seriously. The question, however, is what this implies. Various authors have indicated that agents can be or even should be motivated to act in a trust-responsive way if others are counting on them (Jones, 1996), or argue that being entrusted always brings about a 'tacit demand' (Løgstrup, 1997 [1956]). The next chapter will elaborate on the consequences of the recognition of the truster as a person. What can be expected of a trustee if he is confronted with a vulnerable agent?

Secondly, we have seen that the recognition of someone as an autonomous person highlights the moral values and ideals of that person. Here again, this also holds for the truster. Not all objects of trust are moral ones, but trusters are moral agents and consequently entrust moral issues to others. This means that where trust is damaged the truster may well be wronged. From this perspective, the following quote by Lagerspetz can be understood. 'Calling someone's attitude "trust" is never just making a neutral, empirical point about her mental states, behavior, or the like. It is to claim that we must respect the expectations she has of us…. To "discover" it is to see a human relation in a light that requires a moral response" (Lagerspetz, 1998, pp. 161-162). As a moral agent, the truster is not just struggling with the practicalities of uncertainty. The truster also has ideas about how the trustee should react to his uncertainty and his vulnerable position. For the trustee this requires an awareness of this moral dimension and demands a specific response. The question is of course what this response consists of. This will be the topic in the next chapter.

4.9 Summary

I began this chapter with the definition of trust in terms of an attitude that enables an agent to deal with uncertainty or lack of control. This definition includes dealing with a broad range of uncertainty. More specifically, we have seen that trust is a way of dealing with a specific kind of uncertainty. It is an attitude that enables to deal with autonomous agents in situations of uncertainty, rather than with uncertainty as such.

This implies certain constraints to the tools and methods to build and maintain trust. If trust were about dealing with uncertainty as such, then power, coercion, or controlling behaviour would be relevant methods, which would help to establish trust. However, if we take freedom, agency and a participant attitude as constitutive for trust, these options are incompatible. The discussion of the Hobbesian contract has shown that coercion, power and sanctions may be relevant to establishing public order, yet not to establishing public trust. Genuine trust starts from a participant attitude and from the recognition of the other as an autonomous agent.

Fortunately, the demand of taking the other seriously as an autonomous agent does not merely complicate trust by putting constraints on the possible methods to deal with uncertainty, but it also opens new opportunities. To recognize the other agent as autonomous is also to recognize him as a moral agent. Consequently, if the truster shares aims and values with the trustee he can rely on predictable patterns, but he can also make himself vulnerable. This provides the opening to formulate expectations with a moral dimension. Trust is not only the result of a descriptive assessment of competence and motivation, but also entails a moral judgement. One believes to be entitled to expect something of another agent and to have a legitimate moral claim regarding her.

CHAPTER 5

MAKING TRUSTWORTHINESS OPERATIONAL: RESPECT FOR AUTONOMY AND THE PROBLEM OF MORAL PLURALISM

Mutual respect, in a pluralistic world,
urges us to acknowledge that we are all embedded in cultural contexts
that unavoidably limit our understanding, skew our judgment,
but do not preclude our responsibility to confront and diminish
our prejudices in wider cross-cultural communication.
Th. E. Hill, jr., 1994, p. 38

5.1 Making trustworthiness operational

Trust is an attitude that enables an agent to deal with autonomous agents in situations of uncertainty. Consequently, genuine trust, in contrast to mere reliance starts from the mutual respect of persons who are capable of autonomous agency, i.e. as persons who have the capacity to choose their goals and values personally. In the former chapter, we have seen that this definition directly results in constraints on the tools and methods available to build and maintain trust. Trustworthiness presupposes a participant stance towards the other agent, that is to say, an attitude whereby the trustee considers the other as a participant in a relationship, i.e. as an agent who is able to express trustful expectations.

This illustrates that trusting and responding to trust are not morally neutral. The present chapter defends the claim that if trust grows fragile, it can be a signal of – often implicit – moral judgements about (a) the value of the autonomy of oneself and of the other participant, and about (b) what one can reasonably expect of one another in a situation of uncertainty. In the course of this chapter, I elaborate on the implications of this claim at the level of trustworthiness. First, I introduce problems of trust as a signal of problems related to respect for autonomy and conflicting mutual expectations (Section 5.2). Subsequently I focus on the problems related to the respect for autonomy. Section 5.3 discusses the argument that underlies the duty to respect the autonomy of a person. Next, I present two constraints on what counts as trustworthy behaviour. Both constraints follow directly from the duty to show respect for the autonomy of a truster.

They show the importance of this duty of respect as a necessary, though not sufficient condition for being trustworthy (Section 5.4).

In Section 5.5, the focus shifts to problems of trust as a signal of conflicting moral views on what one can reasonably expect of each other in a trusting relationship. These problems arise because trust is not based on predictable patterns only. Trust can also rest on a judgement about what one may reasonably expect of another agent. As a moral agent, the truster has moral beliefs and ideas that can be recognised in trusting relationships. The duty to respect autonomy entails that if one wants to be trustworthy, the trustee should take these moral beliefs and ideas seriously. However, the trustee is a moral agent too. He also has moral beliefs and ideas about what can be reasonably expected of him. This complicates the demand of respect. In principle, the expectations of both the truster and the trustee are of equal moral weight, because both are autonomous persons. When these views conflict moral deliberation is needed. Such deliberation can at times be quite complex due to moral pluralism. We often lack an overall good or overarching principle that guides us in situations of conflict. Under conditions of moral pluralism trustworthiness is always a matter of balance or compromise between acknowledged moral values and principles adhered to by either party. This raises the question of the feasibility of being trustworthy in a context of moral pluralism. To answer this question, I propose a balance between accommodation and integrity (Section 5.6). To reach this balance it seems unavoidable to compromise on moral matters and is therefore problematic (Section 5.7). Finally, I conclude that trustworthiness that starts from the mutual respect for autonomy and that strives for a balance between accommodation and integrity is feasible in the context of moral pluralism (Section 5.8).

5.2 Problems of trust as a signal of moral questions

A problem of trust can be interpreted as a signal of moral questions. Before elaborating on this claim, it is crucial to formulate two disclaimers. First, loss or lack of trust in itself is not a moral problem. There can be very legitimate reasons not to trust others. It can even be *morally desirable* not to trust some agents or agencies – or to distrust them. For instance, the hesitance to trust the UK government during the BSE crises in the mid-1990s turned out to be a legitimate attitude. Explicit distrust regarding a known criminal may be morally desirable. The claim is that moral questions underlie the problems of trust.

This leads to the second disclaimer. The claim is not that all problems of trust can be reduced to moral questions. It only holds that the reasons that underlie problems of trust may include moral questions and that a neglect of this dimension will result in ongoing problems of trust.

130

Having said this, we can elaborate on the claim that a problem of trust can be interpreted as a signal of one or more moral questions. I see two general contexts in which moral questions underlie the loss of trust. First, a decline of trust can be a signal that one of the agents does not act autonomously or disregards another's autonomy. Trusting or withdrawing trust illustrates how a person judges his own autonomy and that of the other agent. For instance, someone who rashly trusts everyone without any deliberation is easily accused of 'blindness'. The problem of blind trust is not a flawed risk calculation; it signals disrespect of one's own autonomy. The truster is at fault because he did not exercise his capacity of autonomous agency. Similarly, certain responses of a trustee can raise feelings of resentment with the truster and can result in reproaches towards the trustee. These feelings of resentment and reproach are not merely signals of disappointment about a mistaken evaluation of the risks at stake. They emerge from the belief that one is wronged as a person. The trustee's reaction is considered as a response that is incompatible with an attitude of mutual respect as participants in a trusting relationship.

Secondly, trust can become fragile if there is disagreement about mutual expectations. Entrusting and responding to trust reflect ideas on what one considers as legitimate moral expectations. Earlier we saw that trusting is not only a matter of prediction and anticipation. Trusting is a matter of judgement. This includes moral judgement, because the truster and the trustee are moral agents. As a truster, one is not merely struggling with the practicalities of uncertainty, one also has ideas about how the trustee should react given the uncertainty and given the truster's vulnerable position. On the other hand, the trustee has also ideas and beliefs about what is morally right and wrong and about what can reasonably be expected of him. Therefore, trusting and responding to trust includes moral communication. Trusting can be considered as a way of expressing moral claims and responding to trust as accepting the claims.

This relation between problems of trust and the underlying moral judgements has direct consequences for the content of trustworthiness. First, making trustworthiness operational presupposes respect for one's own and the truster's autonomy. Secondly, to be trustworthy one needs to reflect on what this respect implies when one is confronted with conflicting expectations expressed in a trusting relationship. In the next sections these conditions for being trustworthy are elaborated.

5.3 Trustworthiness and due respect for autonomy

Respect for autonomy is essential in trusting relationships. To be trustworthy one must respect one's own autonomy and that of the truster. When the trustee disregards his own capacity of autonomous agency or that of the truster, trust will fade rather than flourish. This section focuses on the moral reasons that make respect for autonomy so important.

The next section elaborates on the implications of the duty of due respect for autonomous agency on the level of trustworthiness (Section 5.4).

To say that someone's autonomy should be respected is not a claim likely to raise much controversy. In very different contexts, including public health, medicine, education, and food, it is generally agreed that one has to respect the autonomy of a patient, student or consumer. Mostly, such a reference to autonomy implies that a patient should be treated in a way that he is able to exercise his right of informed consent and to choose his own treatment. Or that a student is not forced to start a study, and that he is well informed about the curriculum. In the agro-food sector, autonomy is the argument that underlies the emphasis on the consumer's freedom of choice that, among other things, results in the labelling of many products. Despite the fact that the importance of autonomy is broadly shared in society, it is relevant to ask why it is so important that it provides ethical arguments for respecting it in a trusting relationship. To answer this question, autonomy has to be discussed in terms of (a) the ability of self-government and (b) the resulting moral equality of all agents.

5.3.1 Autonomy and moral agency

To take the step from recognising someone as autonomous to the duty to respect autonomy, autonomy has to be understood in terms of self-government. In the chapter above we concluded that autonomy rests on individual freedom and agency. One is autonomous if one has the capacity to act freely in accordance with a self-chosen plan or aim. Autonomy is not mere independence. If it were, proving yourself autonomous would imply that your actions were independent of the interests, beliefs, and desires of others. For trust this would have far-reaching consequences. Trust and autonomy would be incompatible. From this perspective, you either are autonomous and do not need to rely on others, or you are in a dependent position and have to rely on others, but then you are no longer autonomous. In short, you are either a truster or an autonomous person. However, a self-governing person is different from a fully independent person. Autonomy starts in the capacities to act intentionally and to initiate one's own actions, rather than in independence. We can perfectly well act autonomously and concede our past, present and future dependence on others. Baier's analysis of a person also holds for an autonomous person. She agues, 'a person is best seen as one who was long enough dependent upon other persons to acquire the essential arts of personhood. Persons essentially are second persons, who grow up with other persons' (Baier, 1985, p. 84). An autonomous person need not be independent of other agents or agencies. To be autonomous one needs the capacity to 'govern and live [one's] life in an active sense' (Scanlon, 1998, p. 105). Accordingly, Hill argues that autonomy does not just imply independence, but also 'the freedom to choose one's personal ends as opposed to "discovering" them in nature or social practices' (1992a, p. 753). The core of autonomy

is the ability to determine one's own will or, in the line of Mill, a person is autonomous if her desires and impulses are her own. Therefore, a truster can be an autonomous person even if she has to rely on other agents or agencies.

As such, the compatibility of trust and autonomy does not necessarily imply that the autonomy of the truster ought to be respected. Why do we owe respect for a truster as an autonomous person? The Kantian tradition has provided a clear answer to this question. Autonomy is not just one capacity among others. As autonomous persons, we have the ability of self-legislation. That means that we can determine our own duties. Autonomy implies the ability of moral agency. Unlike a patient passively subject to the moral actions of others, autonomous agents can initiate moral action by self-legislation. This capacity does not result in a situation in which 'every man does that which is right in his own eyes.' In determining one's own duties the autonomous person takes into consideration whether this duty can be a duty for all of us. The agent's reasons to adhere to a principle are only to be understood from his reflections on the possibility of ethical justification of this principle for all other agents. This point of view is compatible with the claim that what we morally can expect of others and which duties they have towards us cannot be understood independently of our relationships and of our communities. On the other hand, what an autonomous person is obliged to do cannot be reduced to these relations or communities. The reason to act according to duty is not founded in the obedience to external authorities. It always starts in the capacity to reflect on the question whether this act can be willed by other autonomous agents. This is the reason for treating an autonomous agent as a moral agent of whom we can have moral expectations. Moreover, it is the reason why we should treat him with respect. The fact that he has the capacity to determine his own duty and that he does so from the perspective of the justification of such a duty for all other autonomous persons provides him the value as an end in himself. The capacity to obligate 'is an argument that identifies a certain kind of value – being an end-in-oneself' (Korsgaard, 2004, p. 92). This is the essence of Kant's second formulation of the Categorical Imperative: 'So act that you treat humanity, whether in your own person or in the person of any other, always at the same time as an end, never merely as a means' (*Groundwork,* 1999 [1785], 4: 429). From this imperative it follows that, confronted with an autonomous being, our only appropriate response is to respect this person as an end in itself.[73]

5.3.2 Autonomy and moral equality

The acknowledgement of another as autonomous does not only result in the recognition of the capacity of moral agency. It also leads to the recognition of the moral equality of

---

[73] This is not to say that we cannot have other ethical arguments for respecting human beings. The point here is that the capacity of autonomy results in the moral requirement to respect the autonomous person.

all autonomous agents. This is the second reason that underlies the duty to show respect for the autonomy of the truster.

Moral equality follows directly from the acknowledgement of another as an autonomous person. In spite of the apparent differences between humans, and in spite of the fact that humans have to be respected in their differences, they are equals as autonomous agents. This has implications for the respect that is due to the autonomy of the truster. Even in his vulnerability, the truster is the trustee's moral equal as an autonomous person. His autonomy provides him with an inalienable worth.

To elaborate on the meaning of this worth, Hobbes's view on human worth can serve as a contrast. According to Hobbes, 'the value or worth of a man is, as of all other things, his price; that is to say, so much as would be given for the use of his power, and therefore is not absolute, but a thing dependent on the need and judgement of another' (1651, X, pp. 54-55). The value of a man is here related to his role and function, and the context in which he operates: 'an able conductor of soldiers is of great price in time of war present or imminent, but in peace not so' (Hobbes, 1651, X, p. 55). As the ratio between demand and supply determines the price of many goods, so it also determines the value of man. From this perspective, the truster's value depends on the value that is attached to him by others. If his value is great, like that of the war-time general, he commands much respect even if as truster he is in a vulnerable position. If the value accorded to him is low it is less clear how a trustee should respond.

From the perspective of autonomy this view is untenable. We do not deny that there is a greater need for soldiers during wartime than during peace, but this does not change a human being's worth. In terms of respect for autonomy, the value of a human is in no sense relative to his role, his function or the specific context in which he lives. His value lies in his capacity to 'govern and live his life in an active sense', regardless his position or context. Therefore, he 'is exalted above any price; for as a person ... he is not to be valued merely as a means to the ends of others or even to his own ends, but as an end in himself' (Kant, *The Metaphysics of Morals*, 1996 [1797]**,** p. 186). When we take the autonomy of a person seriously, we cannot but admit that the worth of a man is 'an unconditioned and incomparable worth' (Kant, *Groundwork,* 1999 [1785], 4:436). As an autonomous individual, one person is equal to others, but cannot be simply substituted by another one. The loss or gain with respect to one's autonomy can never be compensated by the loss or gain of that of another person. His worth is something that 'cannot be compared to, traded off against, or compensated for or replaced by any other value' (Wood, 1999, p. 115).

This shows that disrespect for the autonomy of the truster is more than just the disregard of a capacity. It is the disrespect of another as a moral agent and as a moral equal. This explains the feelings of resentment if someone's autonomy is not respected. Being disregarded as a moral agent or as a moral equal raises feelings different from those that arise when is confronted with a social offence. In the case of a lack of respect

for you as a moral agent or as a moral equal, you feel morally wronged. However, in the context sketched by Hobbes, this feeling would be out of place. 'If persons thought of themselves as having only instrumental value, they would not feel such indignation (although they might feel anger, frustration, or irritation)' (Pritchard, 1972, p. 310). In the Hobbesian view there is no component of moral wrongdoing. In contrast, acknowledgement of a human being as a moral equal who has intrinsic worth makes it possible to speak about moral demands for just treatment, and of accusations of wrongdoing if one's autonomy is disrespected. Consequently, we can formulate a moral duty to respect an autonomous person in his capacity as a moral agent and as a moral equal. This implies that another ought to be treated as an end in itself and never as mere means to an end. This holds for every context of human life, trust included. Trustworthiness starts out from an attitude of 'due respect' for one's own autonomy and for that of the truster.

Before elaborating on the implications of this duty of 'due respect' for trustworthiness, we must deal with two questions. The first question is whether the emphasis on autonomy implies that only Kantians will recognise that a person ought to be respected. In other words, if this is so important for trustworthiness, can only a Kantian be trustworthy? Obviously not. The worth of a human being and his capacity to initiate his own actions in freedom commands respect on many other grounds. For instance, intuitionism as presented by Audi emphasises the inherent worth of a human being as a worth that ought to be respected, and that to intuit this special and intrinsic worth one does not necessarily have to see a human being as a rational moral agent in the strict Kantian way (2004, pp. 143-144). Also Løgstrup (1997 [1956]) presents arguments why a trustee should react to a truster and pay him respect, though he does not appeal to acknowledgment of autonomy. The recognition of the ethical demand that is, according to him, inherent in trusting belongs to the 'spontaneous modes of response', which we know by intuition, rather than as results of practical reason. Thus, the importance of respect for the truster is not restricted to those who operate in a Kantian ethical framework. The emphasis on autonomy does have a clear advantage however. It makes the demand to pay due respect to another as a moral agent and as a moral equal binding even for those who do not intuit this demand or for whom this respect is not included in their spontaneous modes of response.

A second question is related to the interpretation of moral equality. Moral equality can easily be mistaken for the view that we respect each other as 'fundamentally equal in power' as Hobbes formulates it, or it can be confused with the claim that we are equal as strictly rational and independent individuals, as some Kantian traditions hold. Both interpretations are problematic in many contexts, but especially in the case of trust. Trust is not among equals in power or among well-informed rational individuals. Trust is by definition asymmetric, and includes differences in knowledge and power. The truster invariably depends on the trustee. This vulnerable status is constitutive for trust. Without

this vulnerable position, there would be no need to trust. Consequently, a symmetric and power-free trusting relationship is impossible. Nonetheless, the truster and the trustee are equals on a moral level. In spite of the vulnerable and depending position of the truster and despite his imperfect knowledge with regard to the object of trust, he still is an autonomous agent. He has the potential capacity to be self-governing even if he does not have the actual opportunity to express this capacity. For that reason he is worthy of respect. On the other hand, the level of power and knowledge of the trustee does not change her moral quality. Her position as a trustee does not make her morally superior to the truster. Both truster and trustee as participants in a trusting relationship have the potential capacity of autonomy. This provides them the status of moral equals in spite of their differences in power and knowledge. For that reason, the trustee has a strong moral reason to treat the truster as a moral equal. That means that he has a duty to respect the autonomy of a person. This is different from the acknowledgement that we are all part of different webs of trusting relationships, which may show that it can be in the trustee's interest to take the truster seriously, because in another context the roles may be reversed. This element of the recognition of the equality in the potential vulnerability can be a strong motivation to take the position of the truster seriously, but it is not the result of an ethical demand. The moral equality based on the potential capacity of autonomy does result in such an ethical demand. Therefore, the fact that the truster's capacity of autonomy is under pressure should not be a reason for the trustee to treat him as non-autonomous, but to enable him to act as autonomously as possible.

5.4 The duty to show due respect for autonomy and the implications for trustworthiness

Basing ourselves on the moral duty to respect an autonomous person in his capacity of a moral agent and as a moral equal, we have concluded that trustworthiness has to start from the duty of due respect for one's own autonomy and for that of the truster. This conclusion introduces constraints on what can count as trustworthy behaviour. In general, it implies that a truster ought not to be considered as just someone in a vulnerable position, but should be treated as a person who has the capacity to choose her goals and values personally. If one takes this attitude as a start, it explains why the expression 'I trust you' involves 'an appeal which will either be ignored or which will bring about moral changes' (Horsburgh, 1960, p. 349). As a trustee, you cannot continue as you did before. It is of course possible to go on with what you were doing, but there is a strong claim to respond (cf. Korsgaard, 1996, p. 140). If you decide not to respond when you are aware that you are trusted, this decision often comes with internal arguments meant to 'justify' your failure to act. Even then we are prompted by awareness of the truster as a moral agent and as a moral equal who is worthy of respect.

Moreover, this process of deliberation also shows that the trustee takes her own autonomy seriously. The duty to show due respect for autonomy does not result in the demand always to respond in the expected and favourable way. Even if the trustee takes the autonomy of the truster seriously, she may have legitimate reasons not to respond to what is entrusted.

When this last aspect is conceded, the duty to pay due respect results in constraints on what can count as trustworthy behaviour. First, any abuse or exploitation of the vulnerable position of the truster rules out trustworthiness. Secondly, manipulation and trustworthiness are incompatible.

5.4.1 Trustworthiness and exploitation

A first constraint on what counts as trustworthy behaviour is the exclusion of any kind of abuse of the vulnerable position of the truster. The duty to pay due respect to the truster as a moral agent and as a moral equal cannot coexist with an abuse of this person. It runs counter to the ethical imperative never to use a person as mere means to an end. Exploiting someone and treating him as an end in himself exclude each other. Consequently, such treatment of the truster can never make an agent worthy of trust.

Nonetheless, the literature provides some examples that suggest that is possible to speak about exploitation and abuse without doubting the genuine character of the trust. A first example can be found in an article by Horsburgh (1960). He sketches a situation in which a wife's trust is repeatedly abused by her husband. He admits that such abuses can result in the end of the trusting relationship. Nonetheless, he argues that one still 'can continue [one's]efforts indefinitely secured from premature discouragement' (p. 349). This suggests that trust and abuse or exploitation can go together, and even can result in the continuation of a trusting relationship. This sounds odd. Why would one consider someone as worthy of trust and accept abuses of trust at the same time? The answer in Horsburgh's example is relatively simple. He speaks about a different and very specific kind of trust, which he defines as 'therapeutic trust'. This kind of trust is not so much an attitude that has its origin in the belief about the trustee's competence and motivation, but rather in an attitude that is close to the 'due respect' condition at the level of the truster. According to Horsburgh, it is possible and even morally required to continue to treat the husband as capable of being trustworthy, because we 'value him as a moral agent' rather than as 'a husband, an employee, or a friend' (p. 349). The case of therapeutic trust and the emphasis on the trustee as a moral agent does not show that trustworthiness is compatible with the exploitation of the truster or the abuse of trust. It only means that the recognition of the trustee as a moral agent comes with a moral claim to 'trust someone, in the therapeutic sense, regardless of the strength of the temptation he is under to abuse one's trust,' although one normally would 'take up an attitude of inveterate and systematic distrust' (p. 355).

The second example is not about this specific kind of trust. In discussing whether trust is a social or moral good, Baier (1991) emphasises that in a climate of trust, exploitation is still possible. One of the two cases she mentions is the 'sexist history of marriage as an institution aiming at providing children with proper parental care.' According to her, this case should convince us that 'mutual trust and mutual trustworthiness in a good cause can coexist with the oppression and exploitation of at least half the trusting and trusted partners' (p. 110). Another case is the practice of certain business firms who exploit their workers in a way that 'is sugarcoated by a paternalistic show of concern for them and the maintenance of a cozy familial atmosphere of mutual trust' (p. 110).

In these cases trust and exploitation indeed seem to go hand in hand. The women and the workers really trusted the others, i.e. they considered them to be competent and willing. Nonetheless, these trusters were exploited. From these examples, we should not conclude that exploitation is compatible with trustworthiness. They only show that the duty of 'due respect' is constitutive for trustworthiness. According to Baier, trust is a belief in the trustee's competence and goodwill (1986, 1991). However, to be trustworthy the condition of goodwill needs to be interpreted in a specific way. The goodwill of the trustee should start out from respect for the truster as a moral agent and as a moral equal. The problem in Baier's example is that the trustees do not pay this respect to the trusters. Those who arranged the marriages were probably competent and may even have had some goodwill towards the women. Also, the companies were likely to be competent and were to a certain extent concerned about the well-being of their workers. However, they all have in common that they did not consider the trusters as autonomous agents who have an inherent worth that ought to be respected. The women were obviously not seen as full partners in the relationship. This equally holds for the workers in the example. In the end, they are used as mere means to commercial ends. If we restrict the condition of goodwill to certain levels of attention to the well-being of the truster, the trustees in the examples could be said to be trustworthy; by adding the duty of 'due respect' for the autonomy of the truster, it turns out that none of the trustees in these examples are trustworthy in the moral sense of the word.

Thus, the two examples by Horsburgh and Baier do not deny the claim that trustworthiness rules out the exploitation of the truster. Horsburgh's case only shows that we may have reasons to treat someone *as if* he were trustworthy, even if he abuses our trust. Baier's examples illustrate what may happen if the duty of the due respect for the truster's autonomy is not taken seriously. As long as the trustee considers the truster as a possible object of exploitation or abuse, he may be trusted by that person, but he will not in fact be worthy of trust. The duty of 'due respect' for the autonomy of the truster, then, is a necessary though not a sufficient condition for trustworthiness.

5.4.2 Due respect for autonomy and the problem of manipulation

A second consequence of the duty to show due respect concerns situations of manipulation. I define manipulation as explicitly distinct from exploitation although they sometimes can go together. I see exploitation as the act of making unfair use of persons. It is deliberately abusing the vulnerable position of truster for one's own ends. Manipulation is a somewhat less direct way of taking advantage. It is *turning* a situation to advantage, by presenting data or using devices in a specific way. The aim is not necessarily to abuse others, but to cleverly but also unscrupulously control or influence a situation.

Intuitively it is often clear that manipulation and trustworthiness mutually exclude each other. However, in practice it is not always that easy to determine the wrong of presenting facts and figures in a specific way that suits you best in order to be trusted by others. For instance, what is the wrong of presenting a product as low in fat, without explicating that it is high in sugar? This way of presenting the quality of the product does not include any flawed claims. But by advertising the product as low in fat the suggestion is that it is beneficial to our health, while the high concentration of sugars may lead to serious doubts about the health impact of the product. Nonetheless, this way of presenting the product may lead consumers to 'trust' it as healthy. The form of presentation can induce the belief that the producer is competent enough with respect to health issues and, at least partly, motivated by health concerns in society with respect to problems such as obesity. The problem here is of course not any flawed claim, but the creation of false beliefs. In this sense, it is a way of manipulation. But why is this incompatible with the duty to respect a person's autonomy?

One argument focuses on the consequences of manipulation for autonomy. From this perspective, the problem of manipulation is that it interferes with someone's ability to exercise autonomous choice. If we, in accordance with Beauchamp & Childress (1994, p. 123) see an autonomous person as someone who acts 'without controlling influences', the problem of manipulation is precisely the problem of interference. Manipulation directly interferes in the ability to act freely in accordance with a self-chosen plan. This claim, however, is too stringent. We all are in one way or another subject to many influences that interfere with our decisions and acts. At least, past states of affairs or events continue to influence our decisions. But is not the same as saying that they undermine an agent's autonomy (Buss, 2002, p. 212). The central element is the interpretation of the word 'controlling'. Interferences in our decisions as such cannot be avoided. To be autonomous it is not even necessary to avoid them. However, if these influences really control our choices, we may question whether we are still autonomous. Buss (2002), however, argues that we overestimate the influence of the manipulator if we define his power in terms of controlling the autonomous person. It is not a matter of control, but the manipulator's will and his actions become 'one of the factors that causes [the truster] to believe – for reasons she herself can defend "all the way down" – that the

choice is, in fact, justified' (p. 213). Admittedly the manipulator influences her decision, but it does not rule out that the agent is still autonomous in the sense that she has governed her own action 'in light of what she did know' (p. 213). Therefore, it is possible to hold the truster responsible for the 'thoughts and actions in which the manipulation results' (Frankfurt, 2002, p. 61). Thus, the consequences for the capacity to act autonomously are not the core of what makes manipulation problematic.

A second type of argument focuses on the moral wrong of manipulation by arguing that an autonomous person cannot reasonably endorse being manipulated (Buss, 2002, p. 217). This is more convincing. If you knew that your decisions were seriously manipulated by the behaviour of someone else, you probably would not have consented to this influence on your action. However, in spite of its attractiveness, Buss claims that this is not a sufficient argument to show that manipulation is wrong because of its impact on autonomy. She states, 'it rests on an exaggerated sense of the desirability of being in control of one's own choices' (2002, p. 218), and that 'our daily lives are filled with situations in which we do not really believe that manipulating and/or deceiving someone is incompatible with treating her as an end in herself' (p. 224). This is not something that happens to us; it is a situation that we, as reasonable people, can permit 'in cases where there are no countervailing moral pressures' (p. 223). There are at least three reasons why Buss's argument is unconvincing. On the one hand, to show that an autonomous person can sometimes consent to the influence of the manipulator does not make the practice as such morally acceptable. It is similar to lying. In some situations, it can be morally desirable to lie, but this does not make lying as such morally acceptable. Second, to show that an autonomous person can consent to and even value manipulation Buss has to stretch the concept of autonomy to the kinds of interference we normally do not define as manipulation. This is the case when she argues that one can consent that others try to influence him in order to 'captivate or distract [him] so as to elicit from him the sort of response she wants – whether this response is an emotion, like love, or an act of forgiveness, a determination not to give up hope, a donation to Oxfam, a job offer, an agreement to consult with a doctor' (pp. 223-224). Finally, her argument shows that the real problem of manipulation is not whether the 'victim' can accept the influence that comes with manipulation.

The real problem of manipulation is neither its consequences for the capacity of autonomy nor is it related to the ability of the 'victim' to accept the influence that comes with manipulation. The wrong lies in the attitude of the manipulator. Manipulation and respect for autonomy mutually exclude each other, because manipulation is a sign that the manipulator fails to recognise another's autonomy as sufficient reason not to manipulate. Not the consequences of manipulation, but the attitude that underlies it makes it a genuine moral wrong. This shift from consequences to attitude is also essential in the context of trust. If manipulation is evaluated in terms of its consequences only, it might appear to be an attractive strategy to evoke trust. There are many ways to

make others believe that you can be trusted (cf. Pettit, 1995). Manipulation can be one of them. However, this all starts from the strategic question how one can be trusted, rather than from the question how one can be trustworthy. If one aims to be trustworthy, manipulation is not acceptable, because manipulation in a trusting relationship is in general a clear sign of a lack of respect for the truster's autonomy. This argument gains the more weight when we take the asymmetry of any trusting relationship into account. Trust is not only about an asymmetry in power, but also in knowledge. Trust is relevant in cases of uncertainty. The truster is in a vulnerable position and from the perspective of respect for autonomy the way information is presented is not neutral. The truster is often not in the position to check or critically question a trustee's presentation of a state of affairs. Manipulation quite often takes advantage of this inability. It treats the vulnerable position of the truster as an opportunity to achieve one's goal more easily, rather than as a reason to help him to act as autonomously as possible.

Thus, if one aims to be trustworthy the moral problem of manipulation is not its consequences nor is it the truster's possibility to consent. Trustworthiness rules out the option of manipulation, because manipulation is a sign of a lack of respect for the autonomy of the truster, while this respect is a necessary condition for trustworthiness.

5.4.3. Respect for autonomy as a necessary condition for trustworthiness

Disrespect for autonomy results in problems of trust. First, one will be confronted with a loss of trust or the inability to build trust if the value of the truster's autonomy is ignored. Exploiting the vulnerable position of the truster is incompatible with taking the truster seriously. It rules out the idea of a trusting relationship between moral equals. The truster is only seen as an opportunity to take advantage of. The central element is the vulnerable position of the truster, rather than what he entrusts the trustee.

Second, trust is undermined when the truster is being manipulated. Even if the impact on the truster's autonomy is relatively small in comparison to other influences such as past experiences and the current state of affairs, manipulation signals an attitude of disrespect for the truster's autonomy. Only if a truster could accept (perhaps retrospectively) this influence as welcome would it fit within the scope of trustworthy behaviour. In that case though, it seems questionable to call this manipulation. In contrast to exploitation, manipulation does not rule out trust as such. At times it can even be an efficient strategy to *gain* trust. However, it rules out trustworthiness and is a very thin basis for trust. If the supposedly trustworthy behaviour turns out to be merely a self-interested strategy, trust is easily harmed and likely to disappear.

The conclusion is that the duty of 'due respect' for autonomy is a necessary, though insufficient condition for trustworthiness. Any attempt to be trustworthy should start from respect for the autonomy of the truster, i.e. in the recognition of him as a moral agent and as a moral equal.

5.5 The moral dimension of expectations in a trusting relationship

The moral dimension of trustworthiness is not fully covered by a focus on autonomy. Problems of trust also can be considered as a signal of differences in moral expectations between the truster and the trustee. Trust is not based on predictable patterns exclusively. Trust also can be grounded in a judgement about what one can reasonably expect of the other agent. Recognition of the truster as an autonomous agent implies that we are not only capable of so-called 'anticipatory trust' (Sztompka, 1999, pp. 27-29) or 'predictive trust' (Hollis, 1998, pp. 10-11), in which the relationship is based on expectations regarding normal patterns and routines. We also have trustful expectations based on more profound beliefs and ideals. In these cases, we trust another to do what is right rather than what is usual. 'Responsive trust' (Sztompka, 1999) or 'normative trust' (Hollis, 1998) shows that trust is not restricted to situations in which we merely expect *that* another will act in a certain way. Trust also includes situations in which we expect something *of* another (Hollis, 1998). These latter expectations can be based on moral beliefs about how the other agent should be motivated and why he should act in the expected way. However, the truster is not the only one who has such moral beliefs. The trustee also has beliefs and ideas about what can be reasonably expected of him. If the beliefs of the truster and the trustee do not coincide, problems of trust arise. This additional moral dimension can be illustrated with the help of the so-called 'consumer concerns', which are regularly expressed in the agro-food sector.

5.5.1 The moral dimension: Trust in food as an example

To trace the moral dimension in trusting relationships, it is interesting to take as an example the so-called consumer concerns which are regularly expressed in the agricultural and food sector.[74] These concerns are relevant in the context of trust for two reasons: like trust, the ability to formulate consumer concerns starts in autonomous agency that includes moral agency; and there is an obvious similarity between the contexts in which these concerns are formulated and in which one has to rely on others.

'Consumer concern' is used as a generic term for concerns that are formulated by consumers and which reflect public uneasiness with regard to a whole range of issues such as food safety, the environment, animal welfare, the acceptability of food production systems, and intrinsic moral objections against certain technologies. These concerns are thought to be a signal of the vulnerable position of the consumer, and his need to rely on others. People formulate concerns if they consider their interests, preferences, and values as endangered. For instance, as long as I am convinced that the

---

[74] The concept of consumer concerns, its moral dimension and the role it plays in the agro-food sector have been extensively discussed in the literature: e.g. Beekman, 2008; Brom, 2000; Brom et al, 2006; Korthals, 2004; Lang, 1999; Meijboom & Brom, 2003; RLG, 1998.

safety of my food is guaranteed, I will not formulate any specific concern on food safety. However, if I have reason to doubt the safety of food because of certain developments in the market, say, the introduction of a new technology, I may formulate a concern in order to stress and explicate my vulnerable position. Thus, the situations in which consumer concerns are formulated are very much like those in which one has to rely on others, i.e. in which trust is relevant. The link between consumer concerns and the need to trust has led some to the claim that these consumer concerns can be seen as 'signs of loss of trust' (Brom, 2000, p. 127; Beekman & Brom, 2007, p. 5). This is not to say that the rise of consumer concerns equals a loss of trust. The expression of a consumer concern is often a signal of the need to rely on others, but does not necessarily indicate the level of trust or distrust.

Given the similarity between the context in which a consumer has to rely on agents and agencies and the situations in which she formulates consumer concerns, it is interesting to see that consumer concerns often have an evident moral dimension. Consumer concerns are not sheer expressions of contingent preferences. Quite often they can be seen as expressions of personal values, norms and ideals (cf. Meijboom & Brom, 2003). For instance, formulating a concern about animal welfare in the food supply chain expresses one's moral view on how livestock should be treated. If one were to consider these animals as mere instruments of meat production, there would be no reason to express any concern about this matter. This example is not rare. Beekman (2008) distinguishes six categories of consumer concerns on issues related to the agricultural and food sector, 'that count as expressions of reasonable and non-superficial food values.' These are concerns about (a) impacts on public and personal health; (b) impacts of genetic modification; (c) impacts on animal welfare; (d) impacts on the natural environment; (e) impacts on international justice and (f) the preservation of regional foods (p. 67).[75] Even though the consumer is in a vulnerable position and in spite of the fact that she often is not able to change matters in the food production personally, she still has the ability to give voice to her moral beliefs with respect to food production and consumption.

On the basis of the similarity between the context in which consumer concerns are formulated and the one in which one needs to rely on others, I treat the ability to give voice as a clear indication that morality plays an important role in trusting too. I distinguish two places where trust has a clear moral dimension.

First, consumer concerns show that the object of trust can have a moral value. As a moral agent, the consumer is able to attach moral value to certain goods and states of

---

[75] Beekman classifies these concerns as 'substantive or end concerns' in contrast to 'procedural or instrumental concerns', i.e. concerns about issues like 'transparency', 'authenticity' or 'trustworthiness'. Beekman accurately emphasises that the latter concerns are procedural, because people mostly have them because they are substantively concerned about one of the six issues mentioned above (2008, p. 67).

affairs. If she has to rely on others with respect to such a good, and thus becomes a truster, she has to entrust others with an object that has a moral dimension. For example, for someone who attaches moral value to botanical species (as part of a natural ecosystem with moral standing), trusting the government to take measures in order to save a certain rare species from extinction has a clear moral character.

Second, the expectations formulated in a trusting relationship can signal the moral beliefs of a truster. Someone's moral beliefs can directly result in expectations that are formulated in a trusting relationship. As a moral agent, the truster has beliefs about duties and responsibilities of the trustee. These beliefs can result in specific expectations in a trusting relationship. For instance, I do not only anticipate that the government is likely to ensure an adequate clinical trial system for the introduction of new medicines. I expect it *of* the Government/ I believe that it has a moral duty to do so, and that I am entitled to expect this even if there were no law or other agreement compelling it.

As said above, this emphasis on the moral dimensions of trust does not entail that trust itself is a moral concept or even a moral good. It only claims that the capacity of moral agency and the position as moral equals have direct implications for the character of the object of trust and for the reasons that can underlie trust and trust-responsiveness.

5.5.2 Problematic responses to the moral dimension

The above-sketched moral dimension in trusting relationships raises the question of how a trustee should deal with this moral dimension if he aims to be trustworthy. To this question, two problematic answers have been formulated.

First, the most rigid answer is the denial of the moral dimension in the trusting relationship. The history of food safety policy offers a clear example.[76] In European countries prior to the late-1990s, governments traditionally presented problems of food safety as technical matters that lack any moral dimension. Policy was presented as based on 'sound science'. Science was considered as value-free, objective, and capable of providing an indisputable input into policy. Consequently, scientists identified the threshold for the safety of food products. They were the only ones to deal with risks and uncertainties. On policy level, risk was conceived as a matter of scientific research, and uncertainties were not explicitly acknowledged. Policy was considered to be based on the indisputable results of science and was presented as if it did not include an evaluative element. Consequently, it was thought that judgments of acceptability did not arise, and that the issue of the acceptability of uncertainties did not need to be explicitly addressed. This absolved the policy maker of the responsibility for making explicit social and evaluative judgments about whether or not the risks were acceptable (Millstone, 2007, p. 491). However, the evaluative element in defining certain risks and uncertainties as

---

[76] For a discussion on food safety policy see Sections 2.7 and 2.8.

acceptable can only be denied at the cost of some serious problems of trust. It is possible to be trusted if matters of food safety are presented as if they were value free, but one cannot be trustworthy. There are two main problems in masking the evaluative moment in assessing risks and uncertainty, and in denying the moral dimension in trusting relationships in general. On the one hand, it clearly does not acknowledge the truster as a moral agent capable of making moral judgements. The assumption is that it is enough to assure the trusting consumer, and that he has no interest in making informed choices. On the other hand, once the truster becomes aware that the trustee downplays or denies that defining a risk as acceptable involves evaluation he will no longer trust the trustee on this point. As the truster sees it, the trustee evidently has neither the competence nor the motivation to deal with this issue. In practice, this denial has resulted in trust-related problems with respect to a whole range of matters, such as the use of genetic modification in food production, the control and prevention of animal diseases, the use of nuclear power and so on. Therefore, the apparent clarity that results from the denial of the moral dimension in the object of trust will in the end only result in serious questions of competence and motivation. Especially when trust is crucial, as during a crisis or in the face of sudden or radical changes, it is often impossible for a trustee to mask the moral dimension of the object of trust. Consequently, problems of trust arise at the moment that trust is most needed. Thus, denial is not a suitable strategy to deal with the problems related to dealing with the moral dimension of trusting.

A second type of answer does not deny the moral dimension as such. It aims to show that problems of trust that appear to have a moral character actually do not include a *moral* disagreement but arise in disagreement on facts. 'The evaluative disagreement evaporates when factual disagreement is removed' (Wolf, 1992, p. 787). For instance, it is likely that only very few agents in the food sector really aim to deceive consumers. Most agree on the duty not to harm and on the moral relevance of informed choice. Nonetheless, there is a substantial debate on how novel food products with health claims should be labelled. This discussion has a clear moral character, but also has its origin in disagreement about non-evaluative facts. Thus, a clarification of facts can be useful.

A further complication is that some disagreements are not really conflicts on values and principles; they are due to a different perception of the (moral) problem. Carol Gilligan's discussion of two reactions to the 'Heinz dilemma'[77] is illustrative in this context (Gilligan, 1993, pp. 25-32). On the one hand, she presents Jake, an eleven-year-old boy, who perceives Heinz's problem as a conflict between the values of property and life. He 'discerns the logical priority of life' (p. 26). Accordingly, he argues that Heinz

---

[77] The Heinz Dilemma is about a man who has to decide whether or not to steal a drug which he cannot afford to buy in order to save the life of his wife. The druggist refuses to lower the price. Consequently, the question is whether Heinz should steal the drug. (see: Kohlberg, L., 1973, 'The claim to moral adequacy of a highest stage of moral judgment', *The Journal of Philosophy*, 70/18, pp. 638-639).

should steal the drug to safe the life of his wife. On the other hand, there is Amy, an eleven-year-old girl. Her response is in terms of long-term impact of the theft on the relationship between Heinz and his wife. She argues that stealing the drug implies the risk of being caught. If that happens, 'his wife might get sicker again' (p. 28). Thus, the situation for the wife could be worse in the long run. Therefore, Amy concludes that 'they should really just talk it out and find some other way to make the money' (p. 28). It is clear that even though both children are confronted with the same moral question they see 'two very different moral problems – Jake a conflict between life and property that can be resolved by logical deduction. Amy a fracture of human relationship that must be mended with its own thread' (p. 31). In trusting relationships moral conflicts can have a similar origin. It can turn out that the initial problem is not about moral disagreement on values and principles, but lies in the interpretation of what the moral problem is about.

It is obvious that this second strategy can be very helpful. A better understanding of the relevant facts and mutual clarity on what is considered the essence of the moral problem is essential for a trusting relationship and certainly helps to reduce the number of situations in which we are confronted with moral conflicts. However, it does not suffice as the only response to the moral dimension in trusting relationships. It begs one important question: How can we deal with genuine moral conflicts in a trusting relationship? These conflicts cannot be reduced to disagreements on factual issues or to a misunderstanding on the moral issue at stake.

5.5.3 Trustworthiness and moral conflicts

From Section 5.4 it follows that the answer to the question of how a trustee has to deal with moral conflicts should start out from respect for the autonomy of both truster and trustee. The recognition of the truster as an autonomous agent results in the duty to take his claims seriously. In other words, the recognition of others as moral agents 'involves the recognition that each of us has the ability, given his or her deliberative capacities, to gain insight in what is morally required' (Van Willigenburg, 2000, p. 401). At face value the statement that in cases of conflict the moral views of the truster should be respected is not very controversial. According to some, it even directly leads to moral requirements for the trustee. Hertzberg for example suggests that 'the grammar of trust involves a perspective of justice: trust can only concern that which one person can rightfully demand of another.' Consequently, he argues that when someone's trust was misplaced 'responsibility rests with the person who failed the trust' (1988, p. 319). However, this claim only holds if we accepted Hertzberg's specific distinction between reliance and

trust.[78] If not, his claim is at least in two ways problematic from the perspective of due respect for autonomy.

First, according to the duty of due respect for autonomy, the trustee should not only respect the moral views of the truster; as an autonomous agent, he should take his own autonomy seriously too. In other words, the trustee should take his own and the other's moral expectations seriously. When these expectations conflict one is confronted with a serious problem that cannot be easily solved. In principle, the expectations of the truster and the trustee are of equal moral weight, because both are autonomous persons. Moreover, there is no easy priority rule that tells the trustee whose expectations are overriding.

A second problem with the above claim is that it rather naively speaks about what 'one person can rightfully demand of another'. If this were clear to all of us, trust would be much easier. However, mostly this is not case. Quite often we are confronted with a plurality of moral views without the availability of an overall good or overarching principle that guides us in situations of conflict. Therefore, respect for autonomy as the answer to the question of how a trustworthy person should deal with this moral dimension automatically leads to the question of how one can be trustworthy in the context of moral pluralism.

5.6 Respect and moral pluralism: Between accommodation and integrity

In Western societies there is a striking plurality of moral views. Pluralism in this sense is mainly a descriptive term. It can be seen as a fact of life. There clearly are a great number of different moral viewpoints within one society. However, with respect to trustworthiness, we are not interested in the fact of the plurality of viewpoints with respect to moral issues. We obviously have tools and ways to address this plurality, e.g. improving communication, increasing transparency or enhancing the level of reflection. From the perspective of trustworthiness, the real problem lies in the plurality that remains even if we are transparent about our moral viewpoints, reason well and are open to further reflection. As Rawls puts it, the main problem for trustworthiness is not that we are confronted with the fact of pluralism as such, but that we are faced with 'reasonable pluralism' (Rawls, 1993, p. 36). The former pluralism can be the result of ignorance, selfishness or bad reasoning. Reasonable pluralism on the other hand is the

---

[78] Hertzberg describes reliance as conditional. It remains subject to one's own judgement about the trustworthiness of others. Genuine trust, on the other hand, does not have limits in advance about 'how far or in what respects I shall trust him' (1988, p. 314). Hertzberg defines the difference in the following metaphor: 'In relying on someone I as it were look down at him from above. I exercise my command of the world. I remain the judge of his actions. In trusting someone I look up from below. I learn from the other what the world is about. I let him be the judge of my actions' (1988, p. 315).

pluralism that remains if we reason well and are well informed. This is a pluralism of inherently incompatible, 'good and important' values and principles (Benjamin, 2001, p. 27) that cannot be reduced by further deliberation. Wolf describes this pluralism as the 'irreducible plurality of values or principles that are relevant to moral judgment' (1992, p. 785). Let us call the pluralism that survives extensive reflection and deliberation 'reflected moral pluralism'.

Reflected moral pluralism is a serious problem for trusting. Because due respect for autonomy is a necessary condition for trustworthiness, being trustworthy implies respect for both one's own and the truster's autonomy. However, it is not clear beforehand what this respect means if the moral viewpoints of the truster and the trustee seriously conflict with each other. The duty of due respect for autonomy only tells *that* we should respect these moral viewpoints as the result of autonomous deliberation, not *how* we should do so when faced with a moral conflict. If one is confronted with conflicting moral viewpoints, there is a 'lack of decisive arguments or reasons in support of one's own position' (Wolf, 1992, p. 786). At the same time there is a lack of such arguments for the position of the truster. The duty to pay due respect to autonomy does not result in an answer to the question whose opinion should prevail in this situation. Confronted with conflicting moral viewpoints in a trusting relationship we have no procedural way to deal with it.

Procedural strategies, such as compensation by money or letting the majority decide what moral point of view is the right one, do not suffice. Financial compensation is incompatible with trust and trustworthiness as such. A truster wants to be assured of what he entrusts another, rather than being compensated for the uncertainty he is faced with. If compensation would settle his problem of uncertainty, there was no real need for him to trust in the first place. On the other hand, a trustee who acts trustworthily for reasons of money, while he does not consider the claim of the truster as legitimate can be accountable, but he is not trustworthy. The motivation with respect to the truster or to the object of trust that is constitutive for trustworthiness becomes irrelevant in this situation.

Similarly, the majority rule is not very helpful. First, this procedural approach is irrelevant in relationships of only two partners. If a truster and a trustee disagree, there is a possibility of building a majority for one position. However, the reason to accept either one of the positions is not the belief that this position is the view of the majority, but because truster and trustee agree on the desirability of this position. Consequently, the disagreement is dissolved and the majority rule as a procedural approach is no longer needed. Moreover, if we broaden the scope of who may form the majority, we are faced with two even bigger problems. On the one hand, there is the question of who should be included. How can those who are no partners in the trusting relationship legitimately decide whose claim should overrule the other? Moreover, even if there are agents or agencies who are in a position to determine whose opinion should prevail, it does not automatically lead to trust. For instance, the voice of parliament or the inference of

courts may help addressing conflicts that result from reflected moral pluralism. Nonetheless, this asks for trust in these institutions, rather than trust in the specific trustee whom one is confronted with. It helps in addressing the moral conflict, but it does not make the trustee trustworthy. It only illustrates the limits of his possible acts and positions.

This leads me to the conclusion that in the case of trust, there is no easy non-moral way of dealing with the moral conflicts that result from reflected moral pluralism. Still, moral pluralism needs not paralyse the possibility to act trustworthily. Respect for autonomy does not only result in a problem. It also provides an opening to address moral disagreement. Rather than the mere observation that one is faced with a moral disagreement, respect for each other requires 'a favourable attitude towards, and constructive interaction with, the persons with whom one disagrees' (Gutmann & Thompson, 1990, p. 76). Respect for autonomy rules out the option of simply ignoring the moral position of the truster. It requires that the trustee acknowledges the moral expectations he opposes as based on moral principles about which one may reasonably disagree. At the same time the duty of due respect prevents that the trustee has to put aside his own moral outlook by definition. His moral position, too, has the status of a view about which one may reasonably disagree. Therefore, I propose a balance between accommodation and integrity.

5.6.1 Two principles of accommodation

To deal with reflected moral pluralism, trustworthiness requires a certain level of accommodation of the trustee.

At a minimal level, accommodation includes an attitude of openness towards counter evidence concerning one's own view and the preparedness to enter the discussion with the truster. This, however, is already a necessary condition in the move from awareness of the fact of pluralism to the level of reflected pluralism. Specific for accommodation is the requirement that the trustee keep open the possibility that one changes one's view in line with the truster's expectation (cf. Gutmann & Thompson, 1990). However, accommodation includes more than a disposition towards openness (cf. Wong, 1992). It incorporates a stance of respect that results in a preparedness to make 'special efforts to "think outside the box" to find a way to accommodate [to the other's] moral views in the particular concrete case that confronts them' (Postow, 2007, p. 202).

This combination of openness, preparedness to change and to actively search for new ways to deal with the conflict means that accommodation is more than just a strategy. In the words of Wong, 'the rationale is not that accommodation is a modus vivendi, acceptable only because the alternatives are in nonmoral terms worse' (1992, p. 774). The trustee has not only prudential, but also moral reasons to accommodate to the view of the truster. A first moral reason directly follows from the element of respect for

autonomy. Because the disagreement is about matters of profound moral importance for the truster, dealing with it in a respectful way requires more than a strategic or procedural approach (Wong, p. 766). Second, the trustee can value the relationship with the truster as an end in itself. If preserving this relationship implies accommodation to the truster's view, the trustee has a moral reason to accommodate. In cases of family and friendship, but also on a public level, one can attach intrinsic value to relationships which are not merely based on 'some degree of agreement in moral belief but also on ties of affection or loyalty, or on a limited set of common goals that may be educational, artistic, political, or economic in nature' (Wong, p. 773). Finally, the trustee can have a moral reason for accommodation because it is often essential to 'promote common moral ends' (pp. 773-774). This is a reason that is especially relevant in the case of trust. Promoting common moral ends requires cooperation – which often does not arise without trust. Therefore, if moral disagreement gets in the way of trust, accommodation is more than a strategy to deal with a problem. As far as the common moral ends are valuable, accommodation is an important step towards their promotion.

Considering these three contexts, accommodation can be seen as morally valuable. Accommodation reflects the respect for the truster, the positive evaluation of the relationship with the truster, or the intention to promote common moral ends. If we grant this moral value of accommodation, it is possible to formulate two principles that are based on our capacity of 'approximating standpoints other than our own, including comparatively external or detached ones' (Benjamin, 1990, p. 76).

The first principle of accommodation follows from the moral demand of respect for autonomy and holds that in accommodating one should defend and act on one's moral beliefs in a way that minimises harm to the truster in his position as moral agent with moral beliefs. As an autonomous agent, the truster is a moral equal; by definition he is also in a vulnerable position. Although the moral beliefs of the truster and of the trustee are of equal value, the trustee can easily overrule the truster's moral beliefs. This vulnerable position of the truster provides the trustee who would defend and act upon his own moral beliefs with a strong moral reason to take additional care not to harm the truster in *his* position as moral agent with moral beliefs. This principle is similar in structure to Rawls's "Difference Principle" (1972; 1993):[79] starting from the moral equality, we have good reasons to take the position of the least advantaged in the relationship, i.e. the truster, into account and act in a way that protects him from harm to his position as a moral agent. The respect for autonomy is the main moral motivation that underlies this principle, just as it is for the principle to treat another as a moral equal.

---

[79] The original Difference Principle is about the fair and equal distribution of basic rights and liberties. The principle states that 'Social and economic inequalities are to satisfy two conditions: (a) they are to be attached to positions and offices open to all under conditions of fair equality of opportunity; and (b), they are to be to the greatest benefit of the least advantaged members of society' (Rawls, 1993, pp. 5-6).

In addition we can formulate a second principle of accommodation. This principle follows from the moral relevance of the relationship. Above, I argued that we can attach value to the trusting relationship as an end in itself, and value the relationship because it is a constitutive element in promoting common moral ends. Consequently, the second principle of accommodation holds that the trustee should act in a way that minimises damage to the relationship between truster and trustee in a case of moral disagreement. An additional, positive effect is that acting according to this principle makes mutual respect a more feasible demand in the context of moral pluralism. This principle 'makes it more likely that opposed sides will be able to regard each other as people to whom it is possible to stand in some positive moral relationship of respect and nonmanipulation rather than simply opponents in a contest to see whose position becomes enforced by our common institutions' (Wong, 1992, p. 777).

These two principles should be action guiding for the trustee if he aims to be trustworthy.

5.6.2 Integrity

In spite of the importance of the above-mentioned two principles of accommodation, they cannot be the full story of trustworthiness in the context of moral pluralism. Acting along these two principles leads to moral costs with respect to the trustee's moral beliefs and ideals. In some cases the trustee considers these costs too high, because some of her ends and values are too important to compromise on. They are so closely related to who the trustee is and what she values in life that compromising would imply an infringement of her moral integrity. If this is the case, two problems arise. First, we may wonder whether the trustee is at all willing to negotiate on the conflicting moral claims. Second, even if she is prepared to do so, the potential truster may wonder whether he can trust people who think that such important moral matters are negotiable (Van Willigenburg, 2000, p. 386). This shows the importance of moral integrity as part of personal integrity for acting in a trustworthy manner.

Personal integrity often refers to the wholeness or intactness of a person's life. In a person of integrity, the various 'parts' of his personality are integrated. However, integrity includes more than the aspect of integration (cf. Musschenga, 2004). Calhoun defines the different aspects of integrity in terms of three 'pictures' of integrity. First, integrity can be defined in terms of the integrated self. In this picture, integrity is about 'the integration of "parts" of oneself … into a whole.' Integrity results, for instance, in the integration of speech and act, but also in the integration of a person's desires and commitments. The second picture links integrity to identity. In this view, 'integrity means fidelity to those projects and principles which are constitutive of one's core identity.' Finally, there is the 'clean-hands' picture of integrity. Integrity is closely related to the way a person deals with conflicts and disagreement. It is especially related

to maintaining the purity of one's own agency in 'dirty-hands situations' (Calhoun, 1995, p. 235). The aspects of integration, identity and coping with conflicts are all relevant for trustworthiness. Having integrity is essential for being trustworthy. Calhoun states, 'We admire and trust those who have integrity' (Calhoun, 1995, p. 235, also Musschenga, 2004, pp. 123-130).

To elaborate on the importance of integrity for trustworthiness, a brief analysis of persons who fall short of integrity, especially of moral integrity, is illuminating. Benjamin (1990) defines four personalities who all lack integrity (pp. 47-52). First, there is the opportunist or the 'moral chameleon' (pp. 8, 47). This is the person who easily changes moral positions and principles when she encounters opposition or if she can gain more from adopting another moral outlook. From the view on integrity as the integration of various 'parts' of one's personality, she cannot be referred to as a person of integrity. This is not to say that she cannot be reliable. If one knows in advance that one is confronted with an opportunist, it is possible to count on this. However, it is evident that this element of predictability does not make her trustworthy. The reason to rely on her is the pattern that appears in her behaviour, rather than the assessment of her competence and motivation. When the element of motivation is included in the assessment and she is found to lack integrity we have reason not to trust her. We expect of those who want to be trusted that they 'espouse their moral positions independently of the circumstances in which they speak' (Gutmann & Thompson, 1990, p. 78). Accordingly, consistency, for instance in speech, is an important sign of trustworthiness. It indicates the moral principles a trustee adheres to and it illustrates that the response to trust is not only prompted by reasons of personal, commercial or political advantage.

Second, we can be confronted with a hypocrite. This is the person who lacks integrity from an external perspective, though he still may have internal integrity. He pretends a certain consistency that is independent of the attached advantage of that position. However, he does not act accordingly. This seriously affects the moral seriousness of a trustee. Integrity without consistency between one's action and the positions one holds is impossible. This especially holds for the consistency between speech and action (Gutmann & Thompson, p. 78).

Consistency is also a problem for the third type of a person who betrays his integrity: the weak-willed person. Although not arising from insincerity, the problem here is similar to that of the hypocrite. Weak-willed persons do not pretend and their position is one of reasonable coherence. They 'lack the courage of their convictions or fail to make conscientious efforts to act in accord with them' (Benjamin, 1990, p. 48). In practice, they can act accordingly to the positions they hold. However, the reach of this consistency is limited. In some cases they are consistent while in other they lack the courage to act on their position. It is reasonable, then, to question their personal and moral integrity because 'though speaking and acting consistently on a particular position,

[they] refuse to recognize and act on its consequences for other related issues' (Gutmann & Thompson, p. 78).

For trustworthiness the lack of consistency is a serious problem. Both in the cases of the hypocrite trustee and the weak-willed trustee it frustrates the assessment of competence and motivation. Even for reliability it is problematic, because the lack of constancy makes anticipating on existing patterns and routines difficult. In the case of the hypocrite problems arise at the aspect of motivation and benevolence. It remains unclear whether this trustee is genuinely motivated by any set of values and principles. His disposition towards you as a truster might be benign, but one never knows whether he is also motivated to act on this position. Confronted with a weak-willed trustee, the problem is similar, but more subtle. The weak-willed trustee, though intending to act according to his values and principles, finds it difficult to do so. This may well lead to serious consequences for the truster. While a hypocrite is not trustworthy to begin with, we may have good reasons to suppose that the weak-willed trustee is trustworthy, i.e. competent and adequately motivated. But when ultimately the trustee proves to lack the courage to act as expected the truster is faced with a kind of betrayal.

Benjamin mentions yet another type of person, the 'self-deceiver', who sees himself as acting on certain values and principles while in fact his conduct that is motivated by quite different, incompatible interests and desires (1990, p. 49). To persist in this discrepancy he has to compartmentalise different aspects of the self, 'the result of which is a further corruption of the wholeness or intactness' (p. 49). Here a major problem has arisen even prior to the issue of consistency. The self-deceiver cannot really be considered as participant in a relationship of trust. The price he has to pay in terms of psychic energy in order to preserve his moral 'integrity' is so high that he is hardly able to act as an autonomous person with the capacity to choose goals and values personally. He might be adequately motivated but lacks trustworthiness. In this case the only trust relationship possible is one of 'therapeutic trust'.[80]

On basis of this presentation of the four types of persons who lack integrity, I conclude that having personal integrity, including moral integrity, is an essential element for being trustworthy. To be worthy of trust one's words and deeds need to be in line with a relatively stable set of values and principles to which one is genuinely committed (cf. Benjamin, pp. 51-52). Stressing the importance of having integrity for being trustworthy does not only have a practical reason. The importance of integrity also follows from the duty to pay due respect to autonomy. The opportunist, the hypocrite, the weak-willed and the self-deceiver all fail to fulfil the duty to take autonomy seriously. The opportunist and the hypocrite may be considered autonomous in the sense

---

[80] This kind of trust primarily has a therapeutic or educational purpose. One trusts another who lacks the competence or the motivation to be trustworthy with the aim to evoke trust, i.e. to teach him the essence of trusting and of taking care. See also the example of the abuse of the trust between husband and wife in Section 5.4 (Horsburgh, 1960; Sztompka, 1999).

that they act freely in accordance with a self-chosen plan or aim.[81] However, they fail to act autonomously in the sense that the principle upon which they act cannot be ethically justified as a principle for all other agents. We cannot will it to be a principle for all of us to act as a hypocrite or as a moral chameleon. Moreover, the behaviour of the opportunist and, to a lesser extent, that of the hypocrite and the weak-willed is morally problematic because they do not take the autonomy of the truster seriously. For the weak-willed and the self-deceiver it is already questionable whether they really govern and live their life in an active sense (Scanlon, 1998, p. 105). Consequently, they do not manage to act according to the duty to respect autonomy.

In sum: to have personal integrity (which includes moral integrity) is a practical and moral *conditio sine qua non* for trustworthiness. It is a practical necessity, because a lack of a relatively stable set of values and principles to which one is genuinely committed complicates the assessment of this person's competence and motivation. It is a moral necessity, because those who lack integrity cannot act according to the duty of respect for autonomy.

This conclusion directly confronts us with the problem of trustworthiness. Above I argued that respect for autonomy results in two moral principles that encourage the trustee to accommodate to the moral viewpoint of the truster. This section has shown that acting according to these principles does not automatically make a trustee trustworthy. The practical and moral importance of integrity raises the question whether such a person can be trustworthy at all. Thus, the main question in dealing with reflected, i.e. reasoned moral pluralism in the context of trust is whether it is possible to keep one's moral integrity and act upon the moral value of accommodation at the same time.

5.7 Integrity and moral disagreement: The compatibility of compromise and trustworthiness

Trustworthiness demands respect for the truster's autonomy, and therefore results in the requirement to act according to the moral value of accommodation. On the other hand, the importance of integrity for trustworthiness implies a clear constraint on the reach of the two principles of accommodation. This tension raises problems in dealing with moral disagreements arising in moral pluralism. This tension becomes most explicit if we focus on the compatibility of trustworthiness and the making of compromises.

On the one hand, the value of accommodation implies preparedness to compromise. Confronted with a genuine moral disagreement some flexibility is required of both

---

[81] In the case of the opportunist it may be that the external influences in terms of possible gain or perceived resistance interfere with his autonomy such that his acts are of the prudential kind rather than based on moral considerations.

parties if they mutually aim to work out the problem. This often implies moral costs for those who are involved. Respect for the other party's moral viewpoint, concerns for the other's vulnerable position or the moral value of the continuation of the relationship may lead us revise our view.On the other hand, consistently acting upon moral values and principles which one sincerely believes to be true is considered laudable. We morally value those who 'have been steadfast in resisting pressures of temptations to compromise' (Benjamin. 1990, p. 1). Making compromises can be seen as a signal of problems at the level of one's integrity. People who often and quickly make compromises are seen as moral chameleons of doubtful integrity. As Van Willigenburg says, 'by giving in on some moral principle or value one may … be in danger of losing one's moral identity' (Van Willigenburg, 2000, p. 387). For trustworthiness this would be a serious problem because a person who lacks integrity cannot be trusted. It seems that moral compromise and moral integrity are incompatible. If this is true, then the question arises whether we can ever compromise on matters of ethics without compromising our integrity' (Benjamin, 2001, p. 29). Additionally, Van Willigenburg shows that this relationship is not only problematic for reasons of consistency. Compromising on moral matters becomes even more problematic if one starts out from the objective nature of moral values. This seems to make them 'immune from negotiation and barter' (2000, p. 387). If one has reasons to believe that one's view is true it is impossible to give in. 'Compromising on that belief must be a sign of serious confusion,' because 'truth is usually regarded as something that cannot be negotiated' (2000, p. 387). Theologian O'Donovan (1994) gives a clear example of the tension that arises if one starts from moral values as part of true beliefs rather than as subjective opinions. He argues that if one takes the *imitatio Christi* as point of departure for our moral actions, striking a bargain with the world would be incompatible. There is no place for compromise because the cost would be 'the loss of clear knowledge of the good' (pp. 95-96).

Thus, a compromise is indispensible or even laudable from the perspective of accommodation, but problematic from the perspective of integrity. This tension is a serious problem from the perspective of trustworthiness, because both accommodation and integrity are necessary conditions. This results in an impasse: a compromise is inevitable and to be avoided at the same time. To determine whether this is a genuine dilemma we need to take a closer look at ' compromise'.

5.7.1 Compromise: Standard, betrayal and prudence

O'Donovan's statement that striking bargains with the world is incompatible with the religious ideal of the imitation of Christ suggests that he would reject any possible role compromise may play in addressing moral disagreements. Surprisingly, this is not the case at all. In the same section, he concedes that every moral decision entails an element

of faithfulness, but also of compromise (1994, p. 96). He presents several reasons for this inclusion of the 'plausibility' of compromising. For my argument the most important reason is that he rejects a specific type of compromise, namely conflict solving via striking bargains. This shows that the concept of compromise encloses different conceptions with different interpretations and different evaluations.

Benjamin (1990) is one of the main authors who considers different interpretations of compromise on moral matters. His analysis of compromise is helpful to understand the rather mixed attitude towards compromise in the moral realm. Benjamin considers compromise as a way of 'splitting' differences. It is a special way of dealing with (moral) conflicts. In the strict sense, a compromise makes the conflict manageable, but it does not make the disagreement fully disappear. 'It makes the best of what both parties regard as a bad situation' (p. 7). Benjamin distinguishes three types of compromise: (a) compromise in the standard sense, (b) compromise as a form of betrayal, and (c) compromise as prudence. A word about each.

The standard sense of compromise. Bejamin refers to the dictionary definition of compromise as 'a settlement of differences by mutual concession' (p. 5). With respect to this standard sense, he stresses the need of making a conceptual distinction between compromise as outcome and as process. The process of compromise does not always lead to a compromise as outcome. For instance, if two parties initially holding opposing positions (A and B) freely opt for 'position S that is a synthesis that combines the strongest features of A and B, while avoiding their agreed-upon drawbacks' we do not speak about a compromise in the strict sense (p. 7). There has been no concession by either one of the parties. This is a process of rational conflict solving, which leads both parties to the belief that S is to be preferred over their initial held positions. In the end, there is nothing to be 'split'.

Compromise as betrayal. This kind of compromise does split the difference, but in a way that entails the betrayal of fundamental views and principles. It includes both self-betrayal and the betrayal of others. In self-betrayal, making a compromise results in infringement of one's own fundamental values and principles. Above, we saw that if a person is quick to compromise one may wonder whether he has any principles at all. This self-betrayal directly affects the person's integrity and consequently his ability to be trustworthy. Those who are quick to compromise become vulnerable to the power of others and are less likely to be adequately motivated in situations of trust. Benjamin even holds that they 'become dangerous to others and unable to be trusted fully' (1990, p. 9). Moreover, self-betrayal is often perceived as a signal of a person who betrays others too. She is the moral chameleon, the opportunist who is always ready to give in on her so-called principles if this suits her better than adhering to her initially held position. Consequently, 'if we count on her too heavily, she is liable to betray us as she betrays herself' (p. 8). Thus, Benjamin accurately concludes that this kind of compromise makes a person both unreliable and untrustworthy.

Compromise as prudence. Compromise for reasons of prudence relates to external conditions that limit the freedom of a person. As with compromise as betrayal, it implies a clear impact on what one considers as valuable and important. However, in contrast to this former type, compromise as prudence does not entail an element of betrayal. Certain circumstances force us to limit our aims and make compromises on what we sincerely value. However, this does not necessarily result in a betrayal of our aims or principles. Benjamin presents the example of an injured athlete (p. 10). Her aim is to perform as good as possible on the next championship. However, because of her injury she can only participate in the event at high risk of turning her current injury into a lifelong disability. If this athlete decides not to participate in the tournament, she genuinely makes a compromise. Nonetheless, her aim is still to perform on the highest level at this tournament. Another case would be a Christian political party with a strong reformed orthodox background in the context of a democracy. In principle, they aim for a theocracy in which the Bible is the one and only guideline for politics and society. However, given the external circumstances, e.g. their minority position, they decide to accommodate to the principles of democracy. Nonetheless, within this framework they still strive for the original aims. Thus, one can argue that they do not really betray their moral integrity. They certainly make a compromise, but it is one of prudence rather than a kind of betrayal of their aims and principles. In these two cases, the element of prudence is rather clear. In practice the line between prudence enforced by external circumstance and betrayal is not always easy to draw (Benjamin, p. 11). For instance, an opportunist's choice to compromise may not be forced, but nonetheless often occasioned by external circumstances.

This analysis of the different types of compromise shows that moral compromises include more than striking bargains on moral issues, which is nothing but a betrayal of someone's moral values and principles. Compromise is not by definition detrimental to our integrity. The 'compromise as prudence' example indicated this. Moreover, Benjamin rightly argues that making compromises is not restricted to the compromises we make with others. We also make internal compromises. We are capable of making internal compromises even on moral matters 'while experiencing no deep threat to personal integrity' (p. 22). This is not a mere inconsistency that actually does affect our integrity. Even if we perceive integrity in terms of the integration of the different parts we value in our life and with the elements which constitute our identity, integrity is not a matter of all or none (p. 22). We face conflicts within ourselves and often we find ways to deal with them, to split the difference without perceiving it as a betrayal of our integrity. This internal process is relevant for the discussion whether one can be trustworthy and making compromises on moral matters at the same time.

We now can turn to the question whether making compromises on moral matters can coexist with being trustworthy. In other words, whether it is possible to accommodate in situations of moral disagreement without endangering one's integrity.

From the above analysis, it is relevant to start to discuss two problematic answers. The first is the answer that a process of bargaining is the preferable strategy. This option is problematic for several reasons. To begin with, the idea of bargaining suggests that moral values, ideals and beliefs can be the input of negotiations just like money or time. However, what holds for resources like money does not hold for moral values. In the negotiation about the price of a house, the two negotiators may agree on a price that is fifteen thousand euros beyond the buyer's first offer and nineteen thousand euros lower than the asking price. This can be considered as a good price that does not include feelings of betrayal even though both have made a compromise. This way of bargaining is incompatible when the negotiations involve moral values. Suppose that John has a disagreement with a trustee in the commercial sector about the moral acceptability of child labour. Based on the inherent worth of children, John considers it unacceptable that the children are forced to work at that age. His belief is not negotiable. A compromise such as in the case of the house sale would be fully implausible. A price between two offers can be an unproblematic outcome of a bargaining process. However, if the outcome of the bargaining process would be that he agrees with child labour in 25 percent of the situations, and at the same time – without any additional arguments – still condemns 75 percent of the similar cases, it would make many doubt the sincerity of John's moral point of view. Certainly this outcome would not contribute to his trustworthiness. Half the price is still a price, but half values are no values at all. Bargaining ignores the special character of moral values and moral ideals. Therefore, the result of bargaining on moral matters is often a compromise as betrayal. This kind of compromise on one's moral values and principles complicates trustworthiness, because it infringes on one's integrity.

A second problematic way of compromising leads to procedural solutions only. As such, procedural solutions are not problematic. We often use the logic of democracy or of the market to address moral conflicts. However, a focus on procedural solutions in terms of searching for a majority or the balance between supply and demand is only relevant from a pragmatic perspective. It does not make a trustee trustworthy. Take the example of the moral acceptability of child labour. Since John cannot check personally whether children are involved in the production of his jeans, he has to rely on others. Procedural measures are only of limited help in this case if he wants to trust. On the one hand, there is the option that if it turns out that his jeans were made by children he will be financially compensated. From a market perspective, this can be a justifiable solution. However, as a truster John wants to be assured that this cannot happen, rather than being financially compensated for the fact that his jeans turn out to be manufactured by

children. The financial compensation does not make the trustee more trustworthy. Neither does it make John more willing to trust. On the other hand, financial incentives may play a role in steering the behaviour of the company. The introduction of a system of fines or other legal measures can influence the behaviour of the company. This can be helpful from the perspective of trustworthiness. These procedural incentives can help to improve the competence and steer the motivation of the company in a way that makes them trustworthy. Nonetheless, in the context of trustworthiness the procedural measures can never serve as the only reason for trust. If one relies on the company only because of the procedural system of fines it is actually reliance on the agencies that introduce and maintain the system of control and penalties. The competence and especially the motivation of the company are only of marginal importance for an assessment whether or not to trust. Thus, a procedures-only approach is not enough in the context of trustworthiness.

5.7.3 Moral compromises, but no real impact on integrity

Having said this about the problems of bargaining and the shortcomings of a strict focus on procedural solutions, we can return to the question whether it is possible to accommodate in situations of moral disagreement without posing a danger to one's integrity to such extent that it affects one's trustworthiness. The answer is positive

- if the acting on the principles that follow from the moral value of accommodation implies a compromise, but does not negatively affect the trustee's integrity or
- if the involved parties all agree that the impact of the compromise on the trustee's integrity is fully justified by the value of the relationship or a common (moral) aim.

The first situation holds for the compromise that is implicit in accepting conflicting views as worthy of further deliberation and debate. Let me explain this with an example. Suppose two vegetarians who both have strong moral beliefs about the way animals ought to be treated and a farmer who is actively involved in intensive husbandry meet in a public meeting about the future of food. The vegetarians clearly do not accept the moral view of the farmer. Nonetheless, the two vegetarians react differently. One of them refuses to enter the debate with the farmer. He thinks it to be a compromise if he were to start a discussion with the farmer. According to him entering the debate would suggest that this moral view is at least legitimate,[82] while he rejects any legitimacy in the idea of using animals for food production. The other vegetarian, however, enters the debate with the farmer. Not because her moral beliefs are less strong or because she thinks intensive farming is morally acceptable, but her respect for the farmer leads her to

---

[82] I define 'legitimate' in line with Postema (1998), who argues that a reason is legitimate if it recognised by others in public deliberation as 'relevant and appropriate from this similar perspective' (p. 444). See also Van den Hoven, 2006, pp. 168-170.

the view that his position has to be taken seriously, although his moral view lacks a binding force on her.

The reaction of second vegetarian includes a compromise, but does not result in a betrayal of her integrity. Even though we sometimes can have strong moral beliefs and ideas, moral pluralism puts severe constraints on our ability to prioritise between conflicting views and on our knowledge about what is the right position. We often remain uncertain about how we can obtain the answers to questions concerning which moral values should take priority, and if we do have answers we are often uncertain about their status (Van Willigenburg, 2000, p. 395).[83] This general uncertainty about how to prioritise between conflicting moral views implies that the moral claims that underlie them have to be taken seriously. Thus, even if we disagree, as in the case of the vegetarian and the farmer, there is a strong reason to respect the other's moral view. It is a kind of moral precaution. Because we are confronted with the moral beliefs of autonomous agents who are our moral equals we had better err on the safe side and enter the debate. This does not mean that we have to accept the other's moral point of view. The respect that is minimally necessary to be trustworthy is, in terms of Postow, a 'stance of modest respectful disapproval' (2007, p. 191; 202). With this stance, she means that a person disapproves the behaviour of another in a *modest* way, because she is not telling him what he should do. And she disapproves in a *respectful* way, i.e. that she makes 'special efforts to "think outside the box" to find a way to accommodate [his] moral view in the particular concrete case that confronts them' (2007, p. 202). From this perspective we can have a genuine intention to enter the debate and still be convinced of our own view. Thus the compromise that is implicit in the reaction of the second vegetarian is a way of dealing with a part of our human predicament in a pluralistic society, rather than a sign of moral weakness.

The first condition is also applicable when the value of accommodation is an important element in the trustee's personal ethics. Even though the truster and the trustee can have conflicting views with respect to the object of trust, the accommodation by the trustee does not need to be detrimental to his integrity. Toleration, respect, openness and loyalty can also be part of one's personal moral outlook. Thus, it would be a simplification to see a moral compromise merely as negative for one's integrity. Someone for whom toleration and personal liberty are important values, making a compromise on one moral issue can be perfectly legitimate. It does not imply any problematic impact on his integrity. Quite the reverse, if he would deny accommodating to the moral view of the truster, one may question the sincerity of his aim to be open and tolerant. This shows that integrity understood in terms of wholeness and consistency does not only put clear constraints on the ability of being a trustworthy compromiser. It

---

[83] Van Willigenbrug speaks at this point about 'epistemic indeterminacy' and 'ontic indeterminacy'.

also can be the reason to compromise on moral matters. Benjamin rightly claims that if we acknowledge 'the complex characterization of the larger network of our values and principles' integrity in all its meanings 'will occasionally require moral compromise' (1990, p. 74).

We now can turn to the second condition in which making compromises on moral matters can be acceptable from the perspective of trustworthiness.

5.7.4 Moral compromises and being trustworthy: A justified impact on integrity

Being a trustworthy compromiser is possible if the involved parties agree that the impact of the compromise on the trustee's integrity is fully justified by a common aim or the value of the relationship. This agreement can arise in different contexts.

First, a clear overarching value may result in such an agreement. Suppose an ultra orthodox Christian physician is confronted with an emergency case on a Sunday. He is trusted to help the patient. If he does so, he has to compromise on his moral beliefs about the acceptability of working on Sundays. However, given the value of what is at stake – the life of a person – both the physician and the patient may agree on the moral acceptability of the compromise that comes with acting in the trusted way. This compromise does not make the physician a moral chameleon or a hypocrite nor does it affect his integrity in some other negative way. His behaviour neither confronts his consciousness with a problem of integrity in terms of the wholeness of the different parts of his moral outlook nor does it lead others to question his integrity in terms of consistency.

Second, the value of a relationship can be a reason for all involved parties to agree that the impact of a compromise on the trustee's integrity is fully justified. In this case accommodating to the truster's view implies an impact on the trustee's integrity, but the moral value the truster and the trustee attach to their relationship justifies this impact to such extent that it does not negatively affect the trustworthiness of the trustee. Suppose, Kate is strongly against travelling by plane, because of the adverse effects for the environment. At a certain moment, a close relative who lives in the United States suddenly becomes terminally ill. The medical prognosis is that she has but a few days to live. She trusts that all her relatives will visit her. The only option for Kate to act in the trusted way is to travel by plane. If she decides not to travel she obviously harms the trust of her relative. If she decides to visit her, she compromises on her moral belief with respect to the acceptability of travelling by plane. In the end, Kate decides to act in the trusted way. She makes a compromise, but it does not make her untrustworthy. The value of the relationship counterbalances the impact of the compromise. Both Kate and her relative consider their relationship more valuable than the costs entailed by the compromise. The fact that, in contrast to her moral beliefs, she decides to fly is not a signal of a lack of any moral principles or lack of integrity as such.

Finally, a common (moral) aim may justify the impact of a compromise on the trustee's integrity. In the case of war, you may be trusted to lie to the enemy about some important military intelligence. You are convinced, though, that it is morally wrong to lie. Nonetheless, both you and the truster may agree that your common aim of fighting the enemy justifies the lie and does not make you untrustworthy at all. Given the circumstances, you even become more trustworthy if you act contrary to your standard moral opinion with regard to lying. Your behaviour is not a sign of arbitrariness. It shows that you have discretionary power that enables you to assess the situation and to decide whether or not you act according to the principle of not lying.

At this point, it is important to stress that there are clear constraints on the extent to which a compromise can be justified in this way. There are limits on the need to make a compromise to be trustworthy and on what can reasonably count as a justification. First, it is important to stress that the duty of respect for autonomy does not only result in the need to compromise, but also limits this need. As moral equals, the truster and the trustee have to strive for some 'balance in the sort and amount of concessions by different parties' (Van Willigenburg, 2000, p. 401) that does justice to their moral equality.

Second, to remain trustworthy someone should express his respect as far as is consistent with his own moral view and his (moral) competence. Some issues are too important to compromise on. If such a core value is the object of trust and a compromise on this value is inevitable one has a valid reason not to act in the trusted way. The compromise would imply an impact on the trustee's moral integrity in a way that harms her trustworthiness in general. This would be the case if a woman who seeks an abortion trusts a physician who is an active and public spokesperson of the pro-life movement. The woman knows this, but considers this physician to be the best qualified person in the field. She is competent enough to be trusted. Moreover, the woman has a strong wish to be treated by a female physician, and this doctor is the only female with this expertise in her region. Nonetheless, making a compromise with respect to abortion would imply an infringement on the integrity of the physician. She cannot hold a position that is strongly against abortion and treat this woman at the same time without running the risk of being a hypocrite or an opportunist. There is no relationship or shared moral value that can justify the moral costs of accommodation. In this case, acting in the trusted way would have an impact on her integrity such that others may question her trustworthiness The physician may nonetheless decide to help the woman, e.g. by giving information or arranging a meeting with another female physician. This may not affect her integrity, but in doing so she does not act in the trusted way. This shows a clear constraint on the principle of accommodation.

Moreover, the compromises that result from the attitude of respect should be consistent with the trustee's (moral) competence. Because trust includes an appeal to respond, trustees are for the most part inclined to act in the trusted way. This inclination is a laudable default position given the duty of respect for autonomy. Nonetheless, if it

implies that the trustee has to make moral compromises it should be consistent with his competence. For instance, a group of consumers and a company may have a disagreement about the safety of a specific product. In essence this debate is not only about the (mis)understanding of the facts and figures, but also touches the question of what is an acceptable level of risk and uncertainty in a food chain. The consumers require risk levels that are near to zero. The company has the sincere belief that higher risk levels are justifiable without compromising on safety, health and well-being. Suppose that in spite of this conflict, the consumer is in a dependent position and has to rely on this company. He trusts this company to apply the 'near to zero risk' strategy. In contrast to the last example, what is at stake is not one of the core values of this company. Thus, from this perspective compromising to the consumer's position is feasible. However, in practice the company does not have the competence to apply the 'near to zero risk' strategy. Even if they would agree on moral superiority of this view, it would not be feasible for them. In this case it is clear that acting in the expected way only would imply an escape in a virtual world. In the end they can be accused of hypocrisy and it will turn out that they are untrustworthy.

This shows that although it is possible to be a trustworthy compromiser, the mutual respect for autonomy, the moral importance of the object of trust and the competence of the trustee put clear constraints on the need and the possibility to compromise on moral matters in a trusting relationship.

5.8 Summary & conclusion

Respect for autonomy is essential to trustworthiness. In this chapter, I have argued that trustworthiness involves the duty to show due respect for one's own autonomy and that of the truster. To ignore this duty has far-reaching consequences. It implies either a disrespect of the truster as moral agent and as moral equal, which will not easily motivate that person to come to trust, or it is a signal that the trustee does not take his own autonomy seriously, which raises questions about his competence and motivation.

This leads to two conclusions. First, any attempt to be trustworthy should start from respect for the autonomy of the truster. Trustworthiness begins in the recognition of him as a moral agent and as a moral equal. This conclusion results in some clear constraints on what counts as trustworthy behaviour. In general, it implies that a truster ought not to be considered just as a vulnerable person, but that he should be treated as person who has the capacity to choose goals and values. Furthermore, this conclusion implies the demand to take the moral dimension of what is entrusted seriously. The duty of 'due respect' for the autonomy of the truster is a necessary, though not sufficient, condition for trustworthiness. To be trustworthy the trustee must also be competent and adequately motivated.

Second, to be trustworthy the trustee cannot ignore her own autonomy. This complicates the duty of respect for autonomy. Problems especially arise if the truster's moral expectations towards the trustee conflict with her belief about what reasonably can be expected of her. This problem is only intensified by the fact that we are confronted with a plurality of moral views that is irreducible, and for which we lack an easy rule telling us which moral position should have priority.

To address the problems of trustworthiness that arise if the moral expectations of the trustee and the truster conflict, I have proposed a balance between accommodation and integrity. On the one hand, the duty to respect autonomy results in two principles of accommodation for a trustee: (a) The principle to defend and act on one's moral beliefs in a way that minimises harm to the truster in his position as moral agent with moral beliefs. The vulnerability of the truster provides the trustee with a strong moral reason to take additional care not to harm the truster's moral beliefs while he, as trustee defends and acts on his own moral beliefs. (b) The principle that the trustee should act in a way that minimises damage to the relationship between truster and trustee in a case of moral disagreement. These two principles guide a trustee in order to respect the moral equality and moral views of the truster. On the other hand, the duty to respect autonomy makes integrity essential for trustworthiness. Those who lack integrity, e.g. the opportunist, the hypocrite or the weak-willed person all remain untrustworthy even if they would act upon the principles of accommodation. This shows the necessity of integrity for being trustworthy and illustrates the need of a balance between the demand to act according to the principles of accommodation and the practical and moral necessity of having integrity. A monopoly of one over the other will not result in trustworthiness.

Thus, making compromises is essential for being trustworthy in situations of conflicting moral expectations. The above discussion has shown that making compromises is not only necessary from the perspective of accommodation, but also feasible from the perspective of integrity. It is possible to reach a balance if making a compromise does not harm the trustee's integrity, or if, according to both the truster and the trustee, the impact on the trustee's integrity is justified, e.g. by the moral importance of their relationship or a common (moral) aim. Nonetheless, there are two constraints on the respect for the conflicting view of the truster, and thus on the possibility of making compromises on moral matters and being trustworthy at the same time. Some issues are too important to compromise on and accommodation should always be consistent with the competence of the trustee.

Given this balance between accommodation and integrity and the feasibility of making compromises on moral matters it is possible to do justice to both the autonomy of the truster and that of the trustee. In other words, it is possible to act in a trustworthy manner even if the truster and the trustee have conflicting views.

164

CHAPTER 6

TRUSTWORTHINESS IN AN INSTITUTIONAL CONTEXT

Politics will, to the end of history,
be an area where conscience and power meet,
where the ethical and coercive factors of human life
will interpenetrate and work out
their tentative and uneasy compromises.
R. Niebuhr, 1934, p. 4

6.1 From trust in persons to trust on an institutional level

Trust in the agricultural and food sector is mostly trust in institutions, in governmental bodies, non-governmental organisations and market parties. Consumers still have face-to-face relationships with individuals. However, in order to buy and consume food they depend not only on the performance of individuals with whom they have personal relationships but also on the competence and motivation of numerous, mostly anonymous others in the food chain. Current developments characteristic for the agro-food sector, such as the process of globalisation, the increase of the use of technology in food production and the recent food-related scandals[84] complicate the identification of the agents we need to trust. And even if it is possible to identify them, it is often difficult to assess the trustworthiness of these agents. Even if we grant that the agro-food sector can be conceived as the result of individual human agency, it is difficult and sometimes even impossible to trace a specific state of affairs back to the individual acts of human agents.

Consequently, personal trust in specific individual agents has largely turned into impersonal trust in institutions. With this kind of trust, the trustee is no longer one person, but is a collective of agents operating operates towards a common aim according to internal, shared rules. This enables a truster to rely even if it is unclear which individual she has to rely on or, if this is clear, whether this individual is trustworthy. For example, I do not know all who are involved in assessing the safety of the food products I buy, and even if I could identify them, I would be confronted with the problem of assessing their trustworthiness. Nonetheless, I trust the government, and especially the

---

[84] An analysis of these developments was presented in Chapter 2.

national food safety authority with matters of food safety. Not because I have personal contact with the staff of this organisation, but because I consider them trustworthy given their mission, aim, strategy and track record. This shows that institutional trust can take the form of trust in the organisation of the institution. My trust is not in the authority's staff as a group, but in the relevance and efficiency of the operating procedures and norms that structure the acts of the staff. I rely on the efficiency of the organisation, rather than on the agents who operate in the institution. Consequently, I can trust specific persons as representatives of the institution. At the same time, the relation between trust in persons and in the organisation is more subtle than the picture of pure institutional trust suggests. We also trust institutions because of the behaviour of their representatives. An expert, well known from television, can directly influence my trust in the scientific governmental body he represents. As long as this expert does not betray my trust, I can confidently rely on his organisation. The difference with personal trust in this case is that I trust the institution because of the behaviour of the individual agent, yet my trust is not in the individual himself. Although I consider him an expert, I do not think he is competent enough to assess issues of food safety on his own. I trust him, because he employs his competence in the institutional setting. Additionally, his behaviour gives the impression that his colleagues in the institution are competent and adequately motivated to be trusted in matters of food safety. Although my trust is related to the acts of this individual person the individual is not the subject of my trust. That is to say, the performance of an individual representative is less strongly related to my trust in the institution than in the case of personal trust.

This change from personal to institutional trust raises questions with regard to the analysis of trust and trustworthiness different from those presented in the former chapters. First, it raises the question of whether it is possible to speak about 'trust' on an institutional level. In the earlier chapters I have defined trust as an attitude which (a) we adopt towards agents, (b) enables the truster to deal with the freedom of these agents in situations of uncertainty, and which (c) is based on an assessment of the trustee's competence and motivation. In Section 6.2, I discuss each of these conditions from the perspective of trust in institutions and conclude that trust can be applicable on an institutional level. Moreover, trust includes expectations about the trustee. This raises the question whether we can have expectations of an institution or of its representatives only. I frame this discussion in terms of the possibility to attribute collective responsibility to an institution and argue that we can have expectations of an institution that are specific for that institution and cannot be reduced to expectations that we might have of individuals (Section 6.3).

Having argued that an institution can be the subject of trust, the question of trustworthiness arises. It is evident that institutions can act reliably according to specific patterns, but this does not yet make them trustworthy. In the sections 6.4 to 6.6 I elaborate on the importance for an institution to be trustworthy and I discuss whether an

institution can be trustworthy. First, Section 6.4 discusses whether the necessary conditions for trustworthiness, i.e. competence in the trusted matter and adequate motivation, can be applied on an institutional level. I conclude that these conditions do not raise fundamental problems for speaking about the trustworthiness of an institution.

Given this conclusion, Section 6.5 discuses whether an institution can be trustworthy in a way that incorporates respect for the truster as a person capable of autonomous agency. From this respect it follows that, if an institutional agent aims to be trustworthy, he should have the competence to deal with the moral dimensions of trusting that come with the truster's capacity of moral agency. I analyse to what extent this respect for the truster is a feasible condition on an institutional level. Furthermore, the respect for the truster's autonomy implies that the trustee's response to what is entrusted should be motivated by this respect. In Chapter 5, I presented the obligation to show due respect for the autonomy of the truster. However, it is complicated to apply this duty on an institutional level. The autonomy of the individual trustee makes the duty binding on an individual level. This duty cannot be grounded in the autonomy of an institution. The duty to act in a way that respects the truster as autonomous person needs another basis. The basis for this obligation is, I argue, the *implicit promise* of the trustee. The way institutions are organised, the reason for their establishment, and the way they communicate often result in an (implicit) invitation to be trusted, i.e. an implicit promise that they will act as they say. The promise implies the obligation.

Finally, Section 6.6 elaborates on the implications of the respect for the truster as an autonomous person for the way a trustworthy institution can deal with conflicting moral expectations. Respecting the truster as a moral agent means that his moral beliefs should be taken seriously, but this cannot be equated with a requirement always to act according to the expectations of the truster. To deal with the plurality of moral views among trusters and between trusters and the institutional trustee, an institution needs to define its position on the balance between acting upon the principles of accommodation and respecting one's integrity. This balance, presented in Chapter 5, illustrates how important it is for institutions to have integrity. Integrity is crucial for being trustworthy, because it helps to clarify what the truster can expect of the institutional trustee, and shows the trustee constraints on the demand of accommodation.

6.2 Trust in institutional agents: Possibility and conditions

Trust in institutional agents such as governmental bodies, NGOs and market parties differs from interpersonal trust that has traditionally been the object of analysis. However, there is profound discussion about the link between both and about the question whether reliance on institutions can be defined as trust. Some claim that it is possible to speak about trust if the subject is an institution. They argue that trust in

groups and institutions is derived from trust in individuals (cf. Lahno, 2002a, p. 37) or a by-product of behaviour towards others (Misztal, 1996, p. 199). Others claim that 'trust in government and trust in people don't have much in common' (Uslaner, 2002, p. 158). This diversity of views shows the need to discuss the question of whether we can speak about genuine trust if someone trustfully relies on an institution or that it would be better to speak of mere reliance or confidence. In the former chapters, I have defined trust as an attitude which (a) we adopt towards agents, (b) enables the truster to deal with the freedom of these agents in situations of uncertainty, and (c) is based on an assessment of the trustee's competence and motivation. In this section, I discuss to what extent these characteristics of trust can also be recognised on an institutional level.

6.2.1 Trust and agency

When I trust my friend, it is easy to define the subject of my trust and to assess whether I am confronted with an agent. In institutional trust these things are more difficult. It is not directly clear what we mean when we say that an individual consumer trusts the government. The answer depends on who or what is seen as the subject of trust. Many positions are possible on the continuum between a specific person who is part of an institution as subject of trust, and the efficacy of the procedures that structure the behaviour of the agents in the institution as subject of trust.

If we start at the 'personal' end of the continuum, we first have to stress that if one's reliance is only based on the competence and motivation of this individual, we speak about personal trust rather than about institutional trust. Nonetheless, trust in an institution can be understood as trust in persons who represent the institution. Such persons are 'access points' (Giddens, 1990, pp. 84-85), who enable individuals to trust an institution. For instance, a medical doctor operates in personal relationships, but at the same time, she represents the hospital and medical profession at large. For that reason, her behaviour can directly influence my trust in the hospital and even in the medical profession. As long as she does not betray my trust, I can confidently rely on the institutions she represents. In this context, it obvious that the trustee is an agent. The fact that one is part of an institution does not change the ability to personally choose one's goals and values among alternatives, and to knowingly and willingly do or not do something. Being part of an institution does not make an individual similar to a bridge that, though having a purpose or goal, does not show any purposeful behaviour and so cannot act. An individual within an institution is still an agent whom one can rely on, even if he conforms his actions to his role within the institution and to the norms and principles of the institution that guide the behaviour of those who are part of it. In these cases, then, the condition that genuine trust is an attitude towards an agent is fulfilled.

On the other hand, one can stress that we often rely on an institution in spite of its representatives. Confronted with a police officer who turns out to be corrupt, I still may

believe the police as such is trustworthy. This illustrates that trust in institutions does not necessarily depend on trust in the specific persons who represent the institution. It is possible to rely on the 'institutionalized practices or procedures based on the belief that if followed they will produce best result' (Sztompka, 1999, p. 44). The presence of these institutional structures and procedures provides a framework which we lack in interpersonal trust. If this framework is the reason to rely, then it is no longer a matter of trust. The truster does not rely on the agency of the representatives of the institution, but on the efficacy of the 'practices and procedures'. He expects this framework to make those who operate within the institution to do something that is close to the example of the bridge: one almost cannot but act in the expected way because the acts of the 'trustee' are fully structured and determined. It is obvious that in such an extreme situation, 'trust' is not the most appropriate term to denote the way people rely on institutions.

In less extreme situations the condition of agency is fulfilled even if the truster relies strongly on the 'practices and procedures' of the institution. The reason is that the efficiency of the practices and procedure only becomes explicit in the behaviour of the representatives of the institution. It is possible to trust an institution despite the behaviour of some individual representatives. If we return to the example of my trust in the police, it is not unreasonable to believe that the police is a trustworthy institution although one police officer acts untrustworthily. However, if half of the police officers were like him, it would no longer help that the police, as an organisation, has very strict rules and procedures. Despite its procedures, the police as institution becomes untrustworthy.

On the basis of the discussion of the continuum between personal trust and reliance on the organisation of an institution, I claim that the subject of institutional trust remains people and their actions (cf. Sztompka, 1999, p. 46). In the case of trust in 'access points', it is evident that people are the subject of trust and that they are agents. But even if we trust individuals because of their role in an institution, the design of the role leads to incentives to get these role holders to do what they must do if the organisation is to fulfil our trust (Hardin, 1998, p. 22). The institutions are the structure in which agency takes place, but are no obstruction to the agency of the trustee (cf. Nooteboom, 2002, p. 31). It is like in drama. The role of the actor is fixed by the script of the play, but the specific performance on the stage depends on the actor's skills and inspiration.

Thus, trust in institutions meets the first of the above-mentioned characteristics of trust. This conclusion directly illustrates that institutional trust also comprises the second characteristic of trust. Trust on an institutional level is an attitude that enables the truster to deal with the freedom that comes with agency, rather than with uncertainty as such. This characteristic is often the reason to start and to establish institutions. Given the complexity of the issues we face in the agro-food sector, as well as in many other domains, dealing with the freedom inscribed in agency is highly difficult or even

impossible. Often it implies that we have to deal with numerous variables in many relationships. For instance, assessing the competence and motivation of all agents involved in the production process of a ready-made pizza is hardly possible for an individual consumer. Nevertheless, it is possible to trust in this situation, because 'certain institutional arrangements convert our particular personal judgment problems into problems of generalized assessments' (Hardin, 2002, p. 8). Thus, trust on an institutional level is still a way to deal with the uncertainty that results from the freedom that comes with the capacity of agency.

6.2.2 Reliance based on predictability or trustworthiness?

The third characteristic that makes trust distinct from reliance or confidence is the assessment of competence and motivation. Trust is an attitude that is based on the truster's assessment of the trustee's competence and motivation. Dunn, however, stresses that this assessment is difficult on an institutional level. According to him, both 'good intentions' and 'practical capacities' are relevant for trust, but in a political or other institutional context, 'it is wiser ... to opt for trust in practical capacity' (1988, pp. 89-90). This solution is as elegant as it is problematic. Leaving out the most complicated issue certainly makes an assessment of trustworthiness more feasible, but at the same time we no longer really trust the other. We merely rely on his or her professional competence. This is also possible if we are confronted with criminals or other persons who may be competent but do not show an adequate motivation towards the truster.

Another option is to stress that the motivations of agents within an institution are structured by the introduction of an organisational framework which makes it possible to rely on institutions independent of the motivation of the individual representatives. The structures and procedures of an organisation provide it a predictability on which one can rely. Hardin would define this as 'quasi trust'. We do not really trust an organisation, but we 'depend on its apparent predictability by induction from its past behaviour' (Hardin, 2002, p. 156). This leads Hardin to the conclusion that we had better speak about confidence than about trust in institutions, especially in the government (2006, pp. 65-66). On this point he follows Luhmann (1988), who argues that both trust and confidence refer to expectations, but that the default case is that of confidence. 'You are confident that your expectations will not be disappointed: that politicians will try to avoid war, that cars will not break down or suddenly leave the street and hit you on your Sunday afternoon walk' (p. 97).

The confidence in these examples is clearly based on the predictable patterns that surface from past behaviour, rather than on the assessment of competence or motivation of one or more agents. Hardin stresses that this last aspect of trust is not feasible at all on an institutional level. According to him, the problem is not only that we do not know enough of all the representatives of an organisation in order to assess their

trustworthiness. We already have difficulties with respect to the assessment of the relevance and efficacy of the organisation. He argues that few individuals have 'the capacity to recognise, understand and evaluate the role and structures that make an institution reliable' (Hardin, 1998, pp. 22-23). Consequently, an assessment of trustworthiness would be a bridge too far.

This, however, is a rather bold conclusion. I do not deny that it can be highly difficult to assess the organisational structure of an institution, but to say that it is almost impossible has far-reaching consequences. If this were true, one could seriously question the need and relevance of all kinds of information services, communications strategies and tools to increase the transparency of institutional agencies. If no one can evaluate the organisation these measures would be superfluous anyway. Nonetheless, it is often considered to be important and relevant to inform and communicate with consumers and citizens about the aims, mission, rules and principles of an organisation. This is not merely a kind of 'window-dressing'. It factually helps individuals to make an evaluation of both the rules and principles of an organisation and the competence and motivation of the agents within the organisation. It can do so because of three characteristics of the assessment of trustworthiness. First, the assessment is individual. From a third-person perspective, one may question whether the available information is enough to come to trust an organisation, but for an individual truster this information can be sufficient to assess the competence and motivation of the organisation and its representatives. The assessment of (institutional) trustee competence and motivation is based on the information that the individual truster conceives as necessary to make the assessment. Second, the assessment does not require full knowledge of all the organisational rules or of the capacity of all involved agents. We need knowledge to come to trust, but we do not need knowledge of every bit and piece about the trustee before we can trust him. It is possible to trust without having full knowledge of the institutional rules or know everything about every representative. Third, the assessment has an emotional component. The problem that Hardin sees betrays a very cognitive view on trust. Trust is informed by knowledge, but as I have argued elsewhere in this thesis,[85] trust has an emotional component as well. The presence or absence of trust already colours the information we have and the amount of information we need in order to assess the trustworthiness of a trustee. That can explain why individuals can be similar with respect to the extent that they depend on a trustee and with respect to the information they have about the trustee, but strongly differ in their attitude towards this trustee.

Lahno (2002b) presents an example of trust in science that illustrates these three characteristics. To trust science one does not need to know and understand every part of science. Quite the contrary, the reason that one does not understand everything about the institution is often the reason to trust. Thus, if one would have full knowledge the reason

---

[85] See Chapter 3.

for trust would disappear. This conclusion, however, does not imply that one can only trust science because of its track record of the last decades. In spite of the imperfect knowledge about the system of science, we often can make an assessment about the competence and motivation of those who operate in the institutional frameworks in which science takes place. This ability can explain why it is possible to keep trusting or even start trusting an institution when one is confronted with new situations in which it is not easy to predict what we can expect or with situations in which the available patterns are conflicting. Then predictability- based reliance becomes highly complicated, while trust based on an assessment of competence and motivation is possible, though not easy.

Hence, the truster's attitude towards an institution can be more than mere reliance on procedures and structures. Trust on an institutional level can include an assessment of competence and motivation of the trustee.

On the basis of the discussion of these conditions for trust, I conclude that trust, in the strict sense, can be applicable to some kinds of reliance in institutions. Trust in institutions is a way to deal with the uncertainty that results from the freedom that comes with the capacity of agency. Furthermore, the truster's assessment is not necessarily limited to the structures and procedures of an organisation. It also can include the assessment of the competence and motivation of the trustee. These characteristics justify reference to genuine trust rather than to reliance or confidence only. At this point, it is important to stress that I do not claim that every form of reliance on institutions can be referred to as trust. As citizen or consumer, one often relies on institutions based on the specific patterns and routines that structure the acts of the institutions. This does not meet the criteria of trust, but can be an effective way to deal with uncertainty and lack of control. My claim is that trust in institutions is possible. This is relevant for those situations in which one lacks predictable patterns or if routines conflict.

6.3 Collective responsibility and the capacity of an institution to function as a trustee

Trust, as an attitude which enables the truster to deal with the uncertainty resulting from the freedom that comes with the capacity of agency, includes specific expectations towards the trustee. A truster has expectations towards the trustee about the way the trustee will or even should act. For example, as truster, I have expectations towards my greengrocer about the quality of the products he sells. I trust him to offer me fresh and tasty vegetables. The fact that I can have these expectations of him makes him a trustee. This raises the question whether an institution can be a trustee in the sense that one can have trustful expectations of an institution rather than of its representatives only. The

answer to this question depends on the interpretation of the relation between the organisation and the individual members.

6.3.1 The relation between the institution and the individual representative

Whether an institution can be an independent trustee depends on the relation between the institution and the individuals who operate in the organisation. If one starts from methodological individualism and argues that collective actions and processes can only be understood from an individualistic perspective, then only individuals can be trustees (cf. Valasquez, 2003). An institution cannot be an independent trustee if collective agency is only interpreted as the sum of individual actions. However, I argue that an institution can be an independent trustee because the relation between the individual agent and the organisation is more dynamic.

An institution cannot be reduced to the sum of individual actions. Even though an institution consists of individual agents, it is more than the sum of the acts of these agents. Institutions are often established for the reason that the total sum of individual actions does not result in the desired effect. Food safety authorities have been established all over Europe not because of a lack of individual experts, but on the assumption that an institutionalised food safety assessment results in a more efficient, more professional and robust assessment, which helps to increase food safety and food quality. Similarly, multinational food companies are not just the result of the megalomania of management teams. They are organised in this way in order to operate more efficiently in term of costs and energy in a global market. Consequently, companies such as Unilever and Ahold can achieve aims that would be impossible without the organisational structure even if the same number of equally qualified individuals were working for these aims.

I agree with Wempe and Kaptein (2002) that a corporation can be distinguished from the individuals operating within the corporation, and that it can be understood as an independent[86] moral entity (pp. 145-146). I believe that this also holds for other institutions, such as governmental bodies and NGOs. These institutions have an organisation, a culture or a tradition that structure individual agency, and arrange the relation between individuals in a specific way. This results in collective agency, which leads to results that cannot be achieved without the organisational structure. This is a first sign that an institution can function as an independent trustee. Even if we concede that the actions are performed by individuals, the mission, rules and principles of an institution coordinate these individual actions in a way that allows us to speak about collective agency (cf. Pettit, 2001, Chapter 5).

---

[86] Wempe & Kaptein speak about a corporation as an 'autonomous' moral entity. Because my use of the concept of autonomy is somewhat different, I prefer 'independence' here. See Section 6.5.

6.3.2 Collective responsibility and trustful expectations

Collective agency is a first sign that an institution can function as an independent trustee. The institution's structure and culture do not only result in coordinated and controlled individual actions, they also lead to a specific responsibility. Collective agency comes with responsibilities that can be related but not reduced to the responsibilities of the individuals operating in the institution. An institution such as a national food safety agency has responsibilities in matters of food safety that no individual staff member has.

I see two reasons why we can speak of collective responsibilities of an institution. First, an institution is a more stable and a more permanent agent than its representatives. Even if we take the position that responsibility is ultimately a matter of individual agents, the responsibility that is attributed to a food safety agency does not depend on the individuals within in the organisation. Even if the entire staff of the agency were to change within five years, the responsibility of the institution would not change. Therefore, it is possible to speak about a collective responsibility that is related to, but independent of the specific responsibilities of the individual representatives.

Second, an 'exclusive emphasis on the individual ignores the importance of social structures in shaping individual consciences and behavior' (Frankel, 1989, p. 110). The individual acts are embedded in the organisational structure of the institution. Therefore, the acts may be performed by individuals, but the result of all these individual acts can be ascribed to the collective agent. The structure and culture of the organisation result in a situation in which the sum of individual responsibilities is not the same as the responsibility of the institution. In the case of the food safety agency, individual staff members perform the various aspects of risk assessment or have a role in the inspection of food safety. Given their role within the institution, they have a responsibility to perform one of these tasks. However, the individual acts are embedded in the organisational structure of the institution, and the institution is embedded in a legal and political context. This leads to a coherence of the individual actions, and this in turn allows us to attribute responsibility to the agency itself.

Attributing collective responsibilities to an institution is relevant from the perspective of trust. It shows that we can have expectations of an institution that are specific for that institution and cannot be reduced to expectations that we might have towards individuals. Especially with respect to objects of trust that are highly complex by nature this possibility is significant. Suppose, my neighbour is a leading expert in the field of food risk assessment. Despite his expertise, I cannot expect him to be in the position to take care of my concerns with respect to food safety. In spite of his expertise, the whole process of risk assessment, including the control and inspection of food products to safeguard public health, is beyond his individual responsibility. Nonetheless, I can have expectations towards the food agency where he is employed, because the agency presents itself as an organisation which takes responsibility in matters of food safety. This makes the agency an independent trustee even though the specific acts are

performed by its staff. Thus an institution can be addressed as an independent trustee. This leads to a next question: What makes an institution trustworthy?

6.4 An institution as trustworthy

The main question of the sections below is: What makes an institution trustworthy? (Sections 6.4–6.6). In this section I first discuss the importance of trustworthiness on the institutional level and analyse whether the necessary conditions for trustworthiness, i.e. competence in the trusted matter and adequate motivation, can be applicable on an institutional level.

It is evident that institutions can be trusted if they are reliable. Stable patterns and routines often provide sufficient reason to rely on institutions. When I buy food at my local supermarket, I simply rely on routines. I expect the safety and quality of the products to be guaranteed although I cannot check this personally. This reliance can be based on past experiences, rather than on an assessment of some agent's competence and motivation. As long as these patterns are available, reliability can be sufficient to deal with uncertainty.

However, this pattern-based reliability can become complicated if we lack a clear routine or if the available routines conflict. For instance, the introduction of new technologies often results in situations in which predictable patterns are not available or conflict. In such situations reliance can no longer be based on predictability. Nonetheless, in personal relations we then still can rely on another. As long as we consider the trustee as competent and adequately motivated, it is possible to rely on her even if we lack a clear pattern or routine. If I ask my friend to care for my newborn son, I trust him. Not because of his excellent track record of baby-sitting, but I trust him because I consider him to be competent enough to do so and because I consider him to have a benevolent motivation.

This shows the importance of trustworthiness next to reliability based on patterns, and it raises the question of what makes an institution trustworthy. In the former chapters, I have shown that trustworthiness comprises the elements of competence and motivation. These conditions are also applicable on an institutional level. But on an institutional level these conditions are relevant in two respects. The trustworthy institution, on the one hand, incorporates the conditions for trustworthiness (Section 6.4.1). On the other hand, such an institution stimulates the trustworthy behaviour of the individual agents working in the institution (Section 6.4.2).

6.4.1 Competence and motivation of the institution

To be trustworthy an institution has to be competent in the relevant matter and its motivation to respond has to incorporate the fact that as organisation it is trusted. The first criterion is relatively clear. If an institution is not competent regarding the object that is entrusted, it is obvious that it cannot act in the trusted way and thus cannot act trustworthily. For instance, if a governmental body is established to prevent terrorist attacks but is understaffed and short of tools, rendering it incompetent the organisation cannot be trustworthy, no matter how dedicated and properly motivated the staff may be. This shows that the element of competence is even more important here than on an interpersonal level. Objects are usually entrusted to an institution in view of the complexity and comprehensiveness involved. The reason for entrusting the government with issues of national security is that it is a collective problem that requires a competence beyond individual reach. Therefore, a governmental body that is established to take care of issues of national security can only be trustworthy if it is competent in matters of national security. With institutional trust, the aspect of competence is often the main incentive to trust an institution. With interpersonal trust, the fact that the trustee is your friend or your colleague may be an additional reason to trust him rather than someone else. In the case of institutional trust the incentive to trust mostly starts in its presumed competence in the relevant matter.

The condition of motivation is essential on an institutional level too. An institution has to be adequately motivated in order to be trustworthy. However, on an institutional level this requirement is somewhat more complex. On the institutional level, the strict interpretation of this requirement in terms of having goodwill towards you as a truster (Baier, 1994a) is hardly feasible.[87] It is often very difficult to determine whether a public or market institution has goodwill towards you personally. Mostly they do not even know you, and thus it is unlikely that an institution is exclusively motivated by their goodwill towards you. Further, an institution can have many reasons to act in the expected way. For instance, a company can have strong prudential reasons to produce safe products. If they would not do so, their competitors will get a bigger market share, which will negatively affect the turnover and interest of the company. This prudential motivation does not make the company untrustworthy, but it is not enough for trustworthiness. Nonetheless, it is possible for an institution to be trustworthy.

Institution may not personally know you who trust or at least rely on them, but they often are aware that they are being relied upon. Once the institution becomes aware that people rely on its procedures and routines, this can act as a motive to keep these routines. The institution may also feel obliged to do so because they are trusted. Consequently, the concerns of the trusters are taken into account, and the fact that these persons are

---

[87] See Chapter 3 for the discussion of the interpretation of this condition of motivation in terms of goodwill.

depending on the institution and that they can be harmed will influence the decisions of the institution (cf. Blackburn, 1998b, p. 33). At minimum, the motivation implies that the institution's actions are influenced by the fact that consumers or citizens in general rely on it with respect to matters in which it is competent. This illustrates that the motivation does not need to be directly related to the individual's trust. The fact that the organisation is motivated by the trust of consumers in general can be enough. This motivation can make the institution trustworthy rather than reliable only, because the truster has the assurance that the institution can and will change its actions as a result of the fact of being trusted.

Second, an institution can be committed to the object of trust. Institutions are often established with a certain aim or mission. The aim results in specific performances and does not only make the institution more predictable, but also enables the truster to have a personal commitment to the institution and its aim(s). For instance, if I am committed to sustainable food production and especially to the improvement of organic farming, my dependence on a commercial, organic food producing organisation can be characterised as genuine trust. It is not just confidence based on structures or on their past performance. I consider them to be motivated by their goodwill with respect to the improvement of sustainable food production. If they would not act in the trusted way, I would not consider it as just an unexpected break in their predictable behaviour. I would see it as a failure of being adequately motivated by the object, which we both value. I do not just consider them unreliable, but untrustworthy.

Thus, on an institutional level the condition of motivation comprises (a) the motivation from commitment with the object of trust or (b) the motivation from the awareness of being trusted.

6.4.2 Promoting trustworthy behaviour

A trustworthy institution is not only competent and adequately motivated, but also has the ability to stimulate the trustworthiness of the individual agents in the institution.

First, a trustworthy institution stimulates the trustworthiness of individual representatives by enhancing their competence. Institutions often have specific procedures to recruit competent agents and to test and certify their competence once they are part of the institution. In this way, the institution monitors the capacities of those who are its representatives. That can explain why I can trust a person who presents herself as a medical doctor and offers to look after my wife's knee. The institutional framework that constitutes the medical profession provides me with some clear indications about her competence as part of this profession. However, the example also illustrates that the role of the institution is at the same time restricted. The institutional framework is often about the minimal conditions that apply to the individual representatives. For instance, even if the physician whom I have to rely on is competent

from the perspective of the medical profession, I still do not know whether she is highly experienced or just started last year. Or whether she is an expert in treating knee injuries or that she has only rarely come across such problems.

Second, a trustworthy institution can stimulate the trustworthiness of individual representatives by its influence on the level of motivation. The organisational structure, which is constitutive for institutions, does not only make individual representatives reliable in the sense that they become predictable. The organisational framework can also influence the motivation of the individual representative. It has the ability to 'stabilise certain practices, and make the performance of the trustee less dependent on his benevolent motivation at the specific moment of action' (Blackburn, 1998b, p. 33). This ability does not change the position of the individual representative as a free agent. One still has the capacity to act contrary to the structures and procedures of the institution. The regular examples of fraud or other criminal activities within institutions illustrate that organisational structures do not change this fundamental freedom. Nonetheless, the organisational framework can influence the motivation of the individual. The fact that an institution can do this need not overrule the different roles that individual agents often have. For instance, someone is not only a civil servant. She is also a parent, a member of a church, a consumer and so on. A trustworthy institution does not deny these roles, but provides a structure and gives guidelines to deal with situations in which these roles and the related interests conflict. Consequently, it is possible to trust the national food safety authority, although one is aware that the organisation consists of many employees who are all individual agents with more social roles and personal interests. Nonetheless, one can expect the staff to be motivated in accordance with the aim of the institution.

These two elements illustrate that a trustworthy institution does not automatically result in trustworthy agents. The institution only provides a framework of minimal competence. The individual agent remains a free agent who can act contrary to the organisational structures. Nonetheless, the institution provides the structure which secures the level of competence that is necessary to act in the trusted way and it influences the individual trustee's motivation such that he will act as expected even if this conflicts with his personal roles or interests.

On the basis of the above discussion of competence and motivation, I conclude that these conditions do not raise fundamental problems for speaking about the trustworthiness on an institution level.

6.5 Trustworthiness and the motivation that starts in respect for autonomy

We have seen that institutions can be more than just predictable. They can be trustworthy in the sense that they are competent in the relevant matter and that they can be adequately motivated to respond to trust. In Chapter 5 however I have argued that trustworthiness should start in the respect for the truster as a person who is capable of autonomous agency. That is to say, if an institutional agent aims to be trustworthy, he should have the competence to deal with the moral dimensions of trusting that come with the truster's capacity of moral agency. Furthermore, respect for the truster's autonomy implies that the trustee's response to what is entrusted should be motivated by this respect. This leads to the question whether these are feasible conditions for institutions. I first deal with the condition of the competence to deal with the moral dimensions of what is entrusted (Section 6.5.1). In the next two sections I turn to the feasibility of the 'respect for autonomy' condition and propose another moral basis for the duty to show respect towards the truster as an autonomous person (Sections 6.5.2 and 6.5.3).

6.5.1 Institutions and dealing with moral dimensions

In the chapter above I argued that an autonomous agent can attach moral value to certain goods and states of affairs. If she has to rely on others with respect to such a good and becomes a truster, she has to entrust others with an object with a moral dimension. In this case, trustworthiness has to incorporate the competence to deal with this moral dimension. For instance, if an individual truster has moral beliefs about the way livestock should be treated, and if he has to rely on a public or market institution with respect to the treatment of animals in the food chain, then this institution has to take this moral dimension seriously in order to be trustworthy. Hence, the question is whether it is feasible for an institution to deal with the moral dimension of what is entrusted. To answer this question, we have to differentiate market institutions from public institutions.

According to some, the requirement of taking the moral dimension seriously is not realistic for a commercial organisation. Milton Friedman's well-known statement is that

> There is one and only one social responsibility of business – to use its resources and engage in activities designed to increase its profits so long as it stays within the rules of the game which is to say, engages in open and free competition without deception or fraud. (1970)

As he sees it, dealing with moral dimensions that do not coincide with the principles of the free market cannot be a responsibility of a commercial organisation. Also Wempe notices that in the 1980s and early 1990s many were of the opinion that 'entrepreneurship and ethics simply do not go together' (2005, p. 211). However, the situation has changed. For different reasons companies like Monsanto, Parmalat, and

Shell, are criticised for not allowing themselves to be guided by moral standards. From the perspective of Friedman, a company like Monsanto could not really be criticised in the case of the introduction of GM-food in Europe. They may have made some mistakes at the level of marketing, but morally culpable they were not. However, in reality they were severely attacked by different consumer and environmental organisations. They were accused of harming consumer autonomy and of using natural resources in a way that is irresponsible. These accusations did not follow from different views on the principles of the free market. They were the result of explicit moral beliefs about what can be legitimately expected of a company. As a consequence of this and many other examples, companies become more and more aware of the importance of this moral dimension and include it in their business strategy. Their reasons for this are often prudential. Companies operate nowadays in a 'global fishbowl' which makes it rather difficult to hide problems and incidents (Kaptein & Wempe, 2002, p. 37). Thus, taking responsibility and explicating basic (moral) assumptions in doing business is seen as an efficient way to prevent public scandals. On the other hand, many companies genuinely acknowledge that they must accept a responsibility 'for upholding certain principles or standards that are – directly or indirectly – relevant for bringing into existence certain social goods such as sustainability or respect for human rights' (Dubbink, 2007, p. 290). Thus, although 'there were serious doubts expressed about combining ethics and business, the link now seems to have become self-evident' (Wempe, 2005, p. 211). It is obvious that the *interpretation* of the link between ethics and business is not equally self-evident. However, it shows that the condition of having the competence to deal with moral issues is feasible for a commercial institution.

For public institutions, the formal condition that the trustworthy institution has to be able to deal with the moral dimension of the object of trust does not raise serious problems. Governmental bodies often are established to deal with issues that include moral dimensions, such as national security, social justice or health care. Nonetheless, there is an important point of discussion. Trusters do not only have moral beliefs about what is morally right. They also have ideals about what is morally good.[88] Trusters are concerned with both levels. Governmental agencies, however, often have difficulties with responding to those objects that include ideas of the moral good. This is, on the one hand, the result of the broad and vague character of these ideals. Consequently, it is difficult to grasp the practical consequences of the moral ideals and to respond to them as a government in an appropriate way. On the other hand, there is a more profound problem. This is the priority of the right over the good in political liberalism (cf. Rawls, 1972; 1988). This liberalism that, *mutatis mutandis*, can be recognised in the governmental structure of many European governments, does not deny that values and

---

[88] Following a rather old, but still concise description, '"the good", or "the best", is the ideal pattern of the life, while "the right" is … the concretion and particularisation of the good into a *hic et nunc*' (Taylor, 1939, p. 289).

ends play a role in society. It denies the priority of one particular vision of the good life that provides an authoritative interpretation in moral matters. Thus, from a liberal political view a governmental trustee should respond to a truster in ways 'that do not presuppose any particular conception of the good' (Sandel, 1984, p. 83). This, however, does not imply that a government cannot deal with broader moral concerns which are based on particular conceptions of the good. According to Rawls, the priority of the right does not imply that 'a liberal political conception …. cannot use any ideas of the good except those that are purely instrumental; or that if it uses noninstrumental ideas of the good, they must be viewed as a matter of individual choice, in which case the political conception as a whole is arbitrarily biased in favor of individualism' (1988, p. 251). Even in political liberalism, the right and the good are complementary, and 'the priority of right does not deny this' (Rawls 1988, p. 252).

This illustrates that, despite the priority of the right over the good in political liberalism, a liberal government is permitted to deal with the ideals of the moral good of those who trust them. It does not imply that, in order to be trustworthy, the government is obligated to act in line with the particular conception of the moral good of a truster. It is, however, possible for a government to care about this element so that all aspects of the moral agency of the truster are taken seriously, which is a precondition for being a trustworthy agent.

I conclude that (a) if an institution aims to be trustworthy, it should have the ability to deal with the moral dimensions of trust that follow from the acknowledgement of the truster as a moral agent and that (b) this necessary condition is feasible for both public and market institutions. In principle these institutions can be trustworthy even if trustworthiness implies that the institutions have to respond to objects with a moral dimension.

6.5.2 Trustee motivation and respect for autonomous agency

Having the competence to deal with moral matters is necessary, but insufficient for being trustworthy. Chapter 5 presented respect for the truster as autonomous agent as an essential element in the motivation that underlies the trustee's response. Whatever the specific response of the trustee may be, the truster should be treated as a person who is capable of autonomous agency and therefore to be respected. One can be trusted, but not trustworthy if the truster is merely seen as a vulnerable individual who can be treated as an end to the trustee's means. This importance of respect for the truster holds for institutions as it holds for individuals who aim to be trustworthy.

However, for individuals we can formulate a duty to show respect for autonomy, which follows from their own autonomy. As an autonomous agent, the individual trustee should act according to the imperative 'to treat humanity, whether in your own person or in the person of any other, always at the same time as an end, never merely as a means'

(Kant, *Groundwork,* 1999 [1785], 4: 429). From this imperative, it follows that if the trustee is confronted with an autonomous being, the only appropriate response is to respect this person as an end in himself.

This line of argument is complicated on an institutional level. An institution cannot be autonomous in the way an individual can be. Sometimes it is argued that an institution can be considered as an 'autonomous moral entity' (Kaptein & Wempe, 2002, pp. 6, 107). However, autonomy then is used somewhat differently. It means that an institution such as a corporation can be seen as an 'independent moral unit' (p. 107). Consequently, the institution can be object of moral evaluation (p. 146). Nonetheless, this independence does not include the capacity of self-obligation, which is presumed in the argument of autonomy. Only the capacity of self-obligation of autonomous agents results in a duty to show respect for the truster as an autonomous agent.

Consequently, the duty to show due respect cannot be directly transferred from the individual to the institutional level. This conclusion does not imply that an institution is not able to respect the truster as an autonomous person. It only says that, on an institutional level, the obligating character of the respect for autonomy and the related protection of the truster cannot be grounded in the autonomy of the trustee. We need an additional argument to safeguard respect for the truster. This confronts us with the question what moral argument can obligate an institution to respect the truster as an autonomous person.

The special character of institutional trust does not only raise this question, but it also entails an answer. Mostly, we trust a particular institution with respect to a specific matter. For instance, I trust the bank in matters of finance, the university in matters of academic education, and the national food authority in matters of food safety. This trust can be understood as the result of my assessment of the competence and motivation of the bank, the university and the food authority. It is no coincidence that these institutions are trusted. They create expectations. They distinguish themselves as experts in the fields of finance, academic education and food safety. Even if institutions do not explicitly say so, they are often established with a certain aim or mission and are structured in a way that creates expectations. In this way, these institutions invite us to trust them. They create expectations that come with an implicit promise about their trustworthiness.

This promise element involves additional responsibility for the trustee. As Scanlon (1990; 1998) argues, a promise creates an expectation and 'thereby an obligation' to keep the promise (1990, p. 213). The stronger the promise is expressed, the stronger the obligation to keep the promise. The obligation follows from the combination of two elements. First, the institution gives the other agent reason n to believe that they attempt to persuade him that they have a settled intention of doing x, i.e., that they will act as the truster expects. And that the institution knows that the truth of this belief is important for the truster. Second, the institution leads the other agent to believe that they know and take seriously the fact that it would be wrong to make this attempt unless they really had

the settled intention to do so (1990, p. 213). If both elements are applicable, the trustee becomes a promiser who has an obligation to act as he has promised to do.

From this obligation Scanlon formulates four principles that should guide the behaviour of the trustee and protects the position of the truster.

6.5.3 The promise in the invitation to trust and four principles for the trustee

From the obligation to keep a promise, Scanlon formulates four principles to show the moral wrongs in not keeping one's promise (1990, pp. 202-210). My claim is that these principles also hold for the institutions that (implicitly) invites trusters to rely on them.

First, if one makes a promise one should act according to the 'Manipulation Principle' (p. 202). This principle states that it is not permissible to manipulate the other by intentionally raising false expectations about how one will react. This principle clearly presupposes the acknowledgment of the other as someone who cannot be addressed as mere means to the institution's ends and underscores the incompatibility of trustworthiness with exploitation and manipulation of the truster.

Secondly, the promiser's behaviour should be guided according to the principle of 'Due Care'. This principle requires that one must 'exercise due care not to lead others to form reasonable but false expectations about what one will do when there is a reason to believe they would suffer significant loss as a result of relying on those expectations' (p. 204). This principle obviously requires more than just abstaining from manipulation. It acknowledges the vulnerable position of the truster and says that the trustee, in his position as promiser, is required to take care that he does not create 'reasonable but false' expectations. The addition 'reasonable' to the false expectations shows a clear limit on this principle. Consumers and citizens can have unreasonable expectations of the government or market parties. Suppose that the government's explicit commitment to animal welfare results in the expectation of an individual that the government will make the food chain free of animal products. This expectation is not only false. It is also unreasonable, i.e. it is an expectation that can be rejected by others as a basis for informed and unforced general agreement (cf. Scanlon, 1998, p. 153). Nonetheless, individuals can also formulate expectations that are false, but not unreasonable. With respect to a food company that distinguishes itself as a corporation that takes social responsibility seriously, it can be reasonable to expect that it sells products that aim to reduce the incidence of obesity. The company considers this issue as mainly the responsibility of public health services and that it should not interfere with individual food choices. The individual's expectation is false, yet not unreasonable. In this case, the principle of due care requires the company to be as clear as possible about what can be expected of it given its commitment to corporate social responsibility. A further important limit to the requirements that follow from this principle is related to the condition that the truster would suffer 'significant loss' if the trustee would not act as

promised. As a result of commercials, one may get the impression that many food products are still produced in a traditional way, according to regional recipes and with natural ingredients only. The language suggests this and the impression is not unreasonable. Nonetheless, it is false. The question is whether this tension between the practice of food production and the consumer expectations about it results in significant loss on the part of the consumers who rely on the companies. There is no general answer to this question. If a consumer believes, on the basis on the communication strategy of a company, that products that belong to the Italian cuisine still come from Italy although this is not case, he may not suffer a significant loss. On the other hand, if the communication of a company leads a vegetarian consumer to the belief that certain products are meat free, although they are not, then this consumer may suffer a significant loss. He runs the risk that he consumes a product that conflicts with his moral beliefs. In this case, I would argue, the company is required to make sure that its language in no way encourages such false expectations.

This last point directly leads to the third principle to guide the acts of the institutional trustee, the 'Loss prevention principle' (Scanlon, 1990, p. 204). This principle implies that one must take reasonable steps to prevent any loss that a person would suffer as a result of the other's reliance on the expectations that one has intentionally or negligently led them to form. Baier (1991) argues that this principle 'could demand a very great deal of us, if we really tried to live by it'. In Scanlon's version, however, the proportionality criterion makes it a rather moderate claim. The required steps should be in accordance with the proportion of the expected loss. If serious losses are at stake, more profound measures are necessary. From this perspective certain actions in the food chain can be understood as expressions of the loss prevention principle. For example, when a product proves to lead to serious health risks it is bound to be recalled or banned. When the situation poses less hazard the principle can result in the requirement of a timely warning that the company cannot fulfil the promise or that one will provide compensation. Suppose that, as a result of technical problems, eggs from free-range chickens are not accurately labelled. Consequently, it is unclear whether the eggs meet the free-range standards. In this case recall would be an exaggerated measure. Nonetheless, given the expectations of consumers, it is essential to give a timely warning that the eggs are not properly labelled, and thus to prevent that the consumer uses eggs she does not want.

Finally, Scanlon formulates a fourth principle, because the centre of promising lies in assurance rather than compensation. This results in the 'fidelity principle', which requires a person who makes a promise to do precisely what he *assured* the other he would do. In Scanlon's version the fidelity principle contains six steps that together lead to the requirement that, in the absence of some special justification, the promiser must do

x unless the other agent consents that *x* is not performed (1990, p. 208).[89] If we apply the six steps to the context of institutions, the principle implies that if

(a) a company voluntarily and intentionally leads its consumers to expect *x,*

(b) the company knows that the consumers want to be assured of this and

(c) acts with the aim to provide this assurance,

(d) the consumer knows that the company has these beliefs and intentions,

(e) the company has in mind that the consumer has this knowledge, and that the consumer knows that the company has this in mind,

(f) the consumer knows that the company has this knowledge and intent, then this company should act in the promised way.

The 'fidelity principle' stipulates that the company is morally bound to *perform* (unless the consumer explicitly waives this requirement). The 'loss prevention' principle provides the consumer with a far weaker kind of assurance, namely independent withdrawal of its promise via warning or compensating. Hence the stronger account of assurance that is entailed by the fidelity principle is needed to make keeping a promise a moral duty and to provide the other party a 'right to rely' (1990, p. 209). This principle does not imply that any kind of breaking a promise is morally wrong. 'It would not render promises pointless to recognize … a cost which is both quite unexpected and much more serious than what is at stake for the promisee' (1998, p. 200). Yet it shows that not performing in the expected way should have the consent of both parties.

These four principles are relevant for trustworthiness on an institutional level. They can protect the truster in his vulnerable status and start in respect for him as a person rather than as a means to the institution's ends. The first three principles provide an additional safeguard to prevent that the vulnerability of the truster is misused or that his vulnerability results in serious (preventable) loss. This reflects the acknowledgment that the truster is a person worthy of respect and protection. The fidelity principle is even more explicit in showing respect for the autonomy of the other, which is reflected in the requirement of mutual consent as a necessary condition for the justification of non-performance.

These principles are significant for trustworthiness on an institutional level also because their moral justification does not start in the capacity of self-obligation, which makes the duty to pay due respect for autonomy binding. Even if an institution lacks the capacity of self-obligation, these principles are valid for an institution if it invites others to trust by creating specific expectations about their future behaviour. The invitation

---

[89] The fidelity principle holds that 'if (1) A voluntarily and intentionally leads B to expect that A will do *x* (unless B consents to A's not doing *x*); (2) A knows that B wants to be assured of this; (3) A acts with the aim of providing this assurance, and has good reason to believe that he or she has done so; (4) B knows that A has the beliefs and intentions just described; (5) A intends for B to know this, and knows that B knows it, and (6) B knows that A has this knowledge and intent; then, in the absence of some special justification, A must do *x,* unless B consents to *x's* not being done' (1990, p. 208).

incorporates a promise and making a promise implies an obligation that results in these four moral guidelines. The four principles express respect towards the truster as an autonomous person and adherence to them results in the protection of the (vulnerable) truster.

In the daily practice of trusting, we have to concede that institutions only rarely formulate the promise such that it explicitly meets the terms of the fidelity principle. Quite often there is no 'right to rely' (Scanlon, 1990, p. 210), because institutions do not formulate clear promises. However, this explicitness is not necessary for the validity of the other three principles. The expectations that are created in others and underlie promising do not necessarily imply that one speaks directly to the other. It also can be done by all kinds of hints. Therefore, the other three principles are often valid and provide a sufficient protection for the truster not to be exploited or manipulated.

Thus, if an institution creates, by communication or by their organisational structure, expectations about their trustworthiness, this promise element results in an obligation to show respect for the truster and to take his vulnerability seriously. The validity of this obligation is independent of the autonomy of the institution and the obligation directly results in action- guiding principles, such as the 'Manipulation principle', the principle of 'Due care' and the 'Loss prevention' principle.

I conclude that the condition that a truster should be treated as a person who is capable of autonomous agency and therefore worthy of respect is feasible on an institutional level. Institutions can be trustworthy even though the duty to show respect to the truster as an autonomous agent cannot be grounded in the institution's capacity of self-obligation. The promise implicit in the reasons that underlie the establishment of an institution, the way it is structured and in the way it communicates, gives rise to an obligation for the institutions to respect the truster and to make sure that the truster is taken seriously despite his vulnerable position.

6.6 Trustworthiness and the importance of institutional integrity

Having argued that it is possible for institutions to deal with the moral dimension of trusting and that an institution can be motivated by the respect for the truster as an autonomous person, I can turn to the problem of dealing with conflicting moral expectations in a trusting relationship.

Respecting the truster as a moral agent means that her moral beliefs should be taken seriously. In a case of conflicting moral expectations, this acknowledgement results in moral reasons to accommodate to the view of the truster. Accommodation means that the trustee should be open to the other's moral view, should be prepared to change his own view and should be willing to actively search for new ways to deal with the conflict. However, accommodation cannot be equated with a requirement always to act according

to the expectations of the truster. To act according to the principles of accommodation[90] is an even more complex task for an institution than it is for an individual. As on an individual level, the institutional trustee is confronted with the question when one should accommodate and when one has legitimate reasons not to do so. Moreover, the institutional trustee is confronted with the question to whose view one should accommodate. Institutional trustees such as a governmental agency or a big retail company are trusted by numerous individuals who all have their own, sometimes conflicting moral opinions. Consequently, it is possible that the institutional trustee offends the moral beliefs of one truster by being prepared to accommodate the moral expectations of another truster. This shows that accommodation cannot imply that in a case of conflict the trustee should always act according to the expectation of the truster. On an individual level one may already wonder whether someone who always changes his view to that of the truster can be trustworthy. However, on an institutional level this is simply impossible.

Next to the willingness to accommodate, having integrity is crucial for being trustworthy. I demonstrated in Chapter 5 that dealing with the plurality of moral views in a trustworthy manner requires a balance between acting upon the principles of accommodation and respecting one's integrity. This also holds for trustworthiness on an institutional level. The institutional trustee, confronted with conflicting moral expectations, needs to define its position on the balance between acting upon the principles of accommodation and respecting its integrity. To be able to do so, an institution needs integrity. Having integrity is crucial, because integrity (a) helps to clarify what the truster can expect of the institutional trustee, and (b) shows the trustee constraints on the demand of accommodation.

6.6.1 Integrity on an institutional level

To analyse what integrity means on an institutional level, we can return to the 'pictures' of integrity as they have been presented by Calhoun.[91] These pictures are (a) 'the integration of "parts" of oneself ... into a whole', (b) integrity understood as 'fidelity to those projects and principles which are constitutive of one's core identity', and (c) the 'clean-hands picture of integrity', which is related to 'maintaining the purity of one's own agency' in 'dirty-hands situations' (Calhoun, 1995, p. 235).

Even though these pictures mainly refer to a person of integrity rather than to an institutional agent, they can be used on an institutional level too. As I have shown above an institution is an independent collective agent that has responsibilities that cannot be reduced to individual responsibilities. For that reason, it is possible to speak about the

---

[90] See Section 5.6.1 for an elaboration of these principles.
[91] See also Section 5.6.2

integrity of an institution. As a collective agent, the three pictures of integrity can apply to institutions.

The 'clean-hands' picture of integrity is often applied to an institution. Integrity requires not only coherence in general, but requires 'sticking to one's principles, moral or otherwise, in the face of temptation' (McFall, 1987, p. 7). The way an institution deals with the conflict of interests and expectations shows its integrity. However, what this integrity implies is not immediately clear. It only shows that an institution has integrity, but does not directly clarify what the truster can expect of the institutional trustee. Therefore, further elaboration of the other views seems more promising.

6.6.2 Integrity as consistency and its shortcomings

To have integrity as an institution the various parts of the institution should constitute a whole and should act in a coherent way. This coherence includes not only consistency between acts of the various parts of the organisation, but also consistency between the acts of the institution and its aims and scope. In the case of a national food authority, this picture of integrity includes both coherence between the departments of the organisation and the coherence between their task to protect human and animal health and their performance in monitoring food products. From the perspective of trustworthiness, this focus on coherence helps the truster to clarify what he can expect of the institutional trustee. This can help to prevent that 'reasonable but false' expectations are raised by the behaviour of the trustee.

However, the help of integrity in terms of consistency is limited, because consistency shows the reliability of an institution, but not necessarily its motivation. An institution can also act in a consistent way for reasons of strategy. Then its motivation can be questioned, and it still remains unclear what a truster can expect of the trustee. As a result the institution can hardly be trusted. De Bakker gives the example of suspicions about 'the development of ever higher food safety standards in the European Union.' People tend to distrust such 'ambitious plans or projects featured by "high" values or ideals.' As long as it is unclear whether the institution is 'primarily motivated by the wish to deal with consumer concerns in a responsible manner,' there remains a suspicion that there are 'hidden agendas of economic interests' (De Bakker, 2007, p. 120). Then too, consistency as such does not solve a moral conflict. If there is a profound disagreement, say about the acceptability of the use of biotechnology in food production, the consistency of the acts of the various parts of a biotech company does not change the disagreement. Integrity in terms of consistency alone is not enough for being trustworthy.

Moreover, the help of integrity in terms of consistency only is limited, because it does not directly show the trustee constraints on the demand of accommodation. Integrity in terms of coherence only provides a procedural constraint, which is often

helpful, but insufficient to be trustworthy if one is confronted with conflicting moral expectations. This shows the need of the final picture of integrity: integrity in terms of a sincere commitment to a certain goal or aim.

6.6.3 Integrity as commitment to the institution's core

Even though it is complicated to speak about the identity of an institution in the same way as we do with regard to individuals (cf. Ashman & Winstanley, 2007), it is possible to distinguish projects and principles of an institution that are constitutive of its core. Institutions are often established for specific reasons and with explicit tasks and aims. For that reason, some acts are fundamental to the existence of the institution, because they directly follow from its tasks or aim. For integrity, the institution's performance should be motivated by the commitment to its task or aim, rather than by any other possible motivation. In the words of McFall, 'if one values not just honesty but honesty for its own sake, then honesty motivated by self-interest is not enough for integrity' (McFall, 1987, p. 8). Similarly, a food safety agency, committed to the protection of human and animal health, should primarily be motivated by this commitment. The agency can have many reasons to perform risk assessments and to monitor food products, but to have integrity the main motivation should be the commitment to public health and animal health and welfare. Only with this motivation, the organisation has 'fidelity' to the task and aim that constitute its core.

Integrity in terms of commitment to the institution's core can help the truster to be clear on what can be expected of this institution in a way that shows more than the institution's reliability. A sincere commitment to certain tasks and aims that form the core of the organisation gives substance and content to the procedural demand of consistency. From this perspective, consistency is not merely a sign of clear routines, but it also can be an indication of the trustee's honesty and commitment to its task and aim.

Furthermore, commitment implies constraints for the trustee on the demands of accommodation. The question to what extent the institution is required to accommodate can be answered in the light of the commitment to the institution's core. This enables an institution to deal with conflicting moral expectations in a way that does not affect its trustworthiness. I conclude that integrity understood as a commitment to the institution's core is necessary to be trustworthy, because this view of integrity enables a trustee to define its position on the balance between accommodation and respect for its own integrity.

Finally, two remarks are necessary to clarify the consequences of this view on integrity. First, it does not imply that every consumer will trust an institution that is motivated by its commitment to its own core. A truster still can have reasons not to trust. For instance, a trustee can be held to be incompetent. Second, for being considered trustworthy it is not necessary that the truster shares the trustee's commitment. Suppose

that I prefer to buy organic food products, but the only way to do so is to buy them from a retail company that is committed to biodynamic agriculture. I do not share their holistic and spiritual understanding of nature and the human being, but nonetheless I admire them for their commitment to the production and retail of organic food products. This makes them trustworthy. Although I do not share their motivation, I think they are adequately motivated to act in the trusted way.

6.6.4 The problem of 'clean hands' revisited

Now it is possible to review the 'clean hands' problem from the perspective of integrity in terms of commitment to the institution's core. It is important to stress that the commitment is not a safeguard for an institution to keep clean hands. Given the variety of interests and expectations both in and outside the institution, this is hardly possible. This leads De Bakker to the conclusion that, 'the truth is that a corporation of great integrity faces unsolvable dirty-hands dilemmas exactly because of its great integrity. One can become painfully aware that in reality honesty and helplessness are closely connected. Integrity is of great value but can also lead to scars on the (personal or organizational) soul that do not feel ''round and sound'' at all' (De Bakker, 2007, p. 126). This view on integrity is problematic in two ways. It tends to be too simplistic and puts too great a burden on the institution. First, it starts in a simplistic view on the relation between principles that are central to the tasks and aim of the institution and conflicting expectations. The quotation suggests that a sincere commitment to a set of principles could only result in one outcome, and if this outcome conflicts with the interests and expectations at stake, the institution has an almost insurmountable problem. Principles, however, are not just fixed norms. They are characterised by flexibility, i.e. they can be adjusted and specified for a certain situation (Verweij, 1998). A principle does not merely indicate what is morally right and wrong in general, but has the ability to guide behaviour in practice. To specify these principles is not easy, but is does not make an institution helpless. Consequently, acting according to principles and being confronted with conflicting expectations does not necessarily lead to the problem of dirty hands.

Furthermore, the quote burdens an organisation of integrity too heavily. Integrity in terms of a commitment to the institution's core does not require that right and wrong are predetermined and are merely a matter of applying principles to a specific case. Integrity requires moral reflection on how the institution's principles have to be interpreted given the conflicting interests. Consequently, an institution of integrity does not need to act the same in each situation. Quite the contrary. From its commitment to its own core, an institution of integrity has the capacity to adjust its acts to the specific interests and expectations at stake. Without having integrity this might be a sign of arbitrariness or the institution may even be accused of opportunism. But if an institution is committed to

certain principles, it is possible to show that the institution still stands for what, in their best judgement, is worth doing for the institution (Calhoun, 1995, p. 260), and that the change of action is the result of discretionary powers that follow from reflection on their principles. This may not directly resolve all problems, but shows that the institution can have and keep integrity even if it is confronted with conflicts.

In sum, having integrity in terms of a genuine commitment to certain tasks and aims that form the core of the organisation is essential for being trustworthy. It helps the truster to understand what can be expected of this institution, while the 'consistency only' view merely considers an institution's reliability. Moreover, this view of integrity helps the trustee to deal with conflicting moral expectations because the institution's commitment results in constraints on the demands of accommodation. As a result of these constraints the room for accommodation becomes clearer. For instance, the institution need not to accommodate in all situations, but given the 'apparent centrality of honesty to the concept of integrity' (McFall, 1987, p. 6), it should be honest about what can be expected. To be trustworthy, it is better for a trustee to explain that the truster's expectation is not feasible for the institution, rather than making doomed attempts to live up to it. Such constraints and requirements that follow from the institution's integrity help the institutional trustee to define its position on the balance between acting upon the principles of accommodation and respecting its integrity. This position may remain 'tentative and uneasy' (Niebuhr, 1934, p. 4). Integrity in terms of commitment does not solve the problem of dirty hands. Nonetheless, it can show that the behaviour of the trustee is not a sign of arbitrariness, but is based on both respect for the truster and the institution's commitment to its core aims and tasks. Integrity provides the institution with reasons whether or not to accommodate in a specific case – reasons that they can present to the potential truster.

6.7 Summary and conclusion

This chapter focused on questions arising in shifting attention from personal to institutional trust. I argued that trust in institutions can be more than reliance and that institutions can be trustworthy rather than reliable only. An institution can be competent and adequately motivated to respond to what is entrusted. This includes competence to deal with moral matters and a motivation prompted by respect for the truster as agent capable of autonomous agency. This respect for the truster is as essential a condition for trust in institutions as it is for interpersonal trust. Despite the problem that the formulated duty to show due respect for autonomy cannot be grounded in the autonomy of an institution, it is possible to formulate principles that start out from respect of the truster and result in his protection. These principles follow from the implicit invitation of many institutions to be trusted. The way institutions are organised, the reason for their

establishment or the way they communicate include a promise that the institution can be trusted. This promise implies the obligation to take care for the truster for the sake of the truster, rather than for any other strategic reason.

Finally, from acknowledgement of the truster as an autonomous person it follows that each moral view within the plurality should be considered as prima facie valuable and as object of accommodation. However, this cannot be equated with a requirement always to act according to the expectations of the truster. To deal with the plurality of moral views among trusters and between trusters and the institutional trustee, an institution needs to define its position on the balance between acting upon the principles of accommodation and respecting one's integrity. To be able to define such a position, an institution needs integrity in terms of a genuine commitment to certain tasks and aims that form the core of the organisation. This helps to clarify what the truster can expect of the institutional trustee, and shows the trustee constraints on the demand of accommodation.

This conclusion can be added to the five steps that have been taken in this study, which I will summarise and apply to the case of trust in novel food in the final chapter.

CHAPTER 7

A FRAMEWORK TO ASSESS TRUSTWORTHINESS IN PRACTICE:

THE CASE OF PRODUCTS AT THE INTERFACE BETWEEN FOOD

AND HEALTH

> *trustworthiness is not just mechanical dependability,*
> *and trust is not merely confidence in a range of particular actions.*
> A.C. Baier, 1991, p. 117

7.1 Trustworthiness and the five steps that follow from this study

To conclude this study this chapter presents the five main steps taken in the previous chapters (Section 7.1). To show the relevance of these steps as a framework to assess trustworthiness in practice, I discuss the problem of trust in a specific case. As a result of different developments, the food sector and health sector become more and more intertwined. This creates many prospects, but also raises questions. One of the questions relates to the implications for public trust in food and health. For that reason, I apply the five steps to the problem of trust in the case of the introduction of novel food products at the interface between food and health (Sections 7.2 – 7.7).

The study began with posing the problem that people have to rely on others, but often do not know whom to trust. I called this the 'problem of trust'.

From trust to trustworthiness. As a first step, I proposed a shift to conceptualising the problem of trust as a problem of trustworthiness. For this step I offered both conceptual and moral reasons. On the one hand, it has a conceptual background. An individual cannot *decide* to trust. One may want to trust, but one cannot trust at will. On the other hand, the autonomy of the individual provides a strong moral reason for this shift. A lack of or hesitance to trust should be acknowledged as a legitimate point of view, rather than as a failure only. This does not imply that the truster cannot be wrong, but shows that the burden of the proof also lies at the level of trustworthiness. Even though a trusting relationship is by definition asymmetric and includes differences in knowledge and power, the truster should be treated as a person who is capable of autonomous agency.

Consequently, the main question is not how the individual can be changed so that he will trust, but what conditions the trustee has to fulfil in order to be worthy of trust. This first step has direct consequences for how the 'problem of trust' has to be addressed.

From risk to trust. Second, in the conceptual analysis of trust, I argued that although trust and risk taking can be relevant in the same situation there is a fundamental difference between them. Being trustworthy is different from enabling individuals to take risk or to reduce the risk at stake.

From reliability to trustworthiness. As a third step, I focused on the importance of trust as a way of dealing with the uncertainty that comes with the freedom of agency, rather than with uncertainty as such. Trustworthiness includes more than reliability that results from predictability. If a trustee would invariably act according to a predictable pattern determined by his nature or its organisational structure, it would be like relying on a machine. Machines perform, they do not deliberate; machines are programmed, not motivated. The trusted agent should not merely act in a predictable way, but should be motivated to respond to what is entrusted. Thus, if one says that some (collective) agent is trustworthy, it means that he is not just predictable, but worthy of trust even though he has the power to do otherwise.

From competence and motivation to the inclusion of respect for autonomy. As a fourth step, I argued that trustworthiness begins from the duty to show due respect for the truster as a person who is capable of autonomous agency. Despite the vulnerable status of the truster, trustworthiness is predicated on recognition of the truster as a moral agent and as a moral equal. This respect results in some clear constraints on what counts as trustworthy behaviour. In general, it implies that a truster ought not to be considered just as a vulnerable person, but that he should be treated as person who has the capacity to choose his own goals and values. Consequently, any form of exploitation or manipulation of the truster by the trustee rules out trustworthiness. Furthermore, this conclusion implies the demand to take the moral dimension of what is entrusted seriously. Trustworthiness has to include the ability and willingness to deal with the moral dimension of trust implied in the recognition of the truster as moral agent.

From arbitrary compromises to acting in a trustworthy manner. The final step is to cope with the problem of conflicting moral expectations that results from the respect of the truster as a moral agent and a moral equal. If one starts out from respect for the truster, her views have to be taken seriously even if they are in conflict with those of the trustee. This requires accommodation, which implies that the trustee should be open to the other's moral view, should be prepared to change his view and should be willing to actively search for new ways to deal with the conflict. To deal with such a conflict making compromises is often inevitable. The trustworthiness of an agent who makes compromises in every situation can be questioned. Compromises easily get a character of arbitrariness if the truster does not have legitimate reasons for the decision whether or not to act according to the expectation of the truster. Moreover, on an institutional level,

it is sometimes unclear to whose view one should accommodate given the many trusters and many expectations. Nonetheless, it is possible to remain trustworthy and make the compromises that are sometimes necessary from the perspective of accommodation. Having integrity is essential with respect to this. Integrity does not just complicate the demand of accommodation. It can provide reflected constraints on the demand of accommodation. Integrity understood as a sincere commitment of the trustee with 'those projects and principles which are constitutive of one's core identity' or with the tasks and aims that are constitutive for an institution, lead to constraints on the demand of accommodation that are not arbitrary. They are not beyond debate, but they are reflected and the trustee can give legitimate reasons for the decision whether or not to act in the expected way. Consequently, not everyone will trust this agent, but the agent is trustworthy despite of the confrontation with the moral conflict. In practice, it can be frustrating to lose trust although one has legitimate reasons not to act in the expected way. Despite this frustration, this is better than trying to look for ways to get people to trust an agent although this agent is not trustworthy, either because he is not competent or not adequately motivated. If one is trusted, but not trustworthy, the problem of trust will return in the end.

To show the implications of these five steps, I discuss the questions of trust and trustworthiness in the case of novel foods at the interface between food and health (Meijboom, 2007).

7.2 Trustworthiness and new food products at the interface between food and health

Milk that lowers your level of cholesterol, dietary advice based on genetic knowledge, new scientific information on the link between a food product and the occurrence of certain types of cancer – these are but some of the many examples of the trend to use health knowledge in the development of food and vice versa. They clearly signal that the food sector and health sector become more and more intertwined. Since there is enough safe food for everyone (in the Western world), the general policy of the food sector has shifted from increasing production rates to more quality-focused production (McInerney, 2002; Lang, 1999).

This creates many prospects, but it also raises questions. One of them is the question of what the implication is for public trust in food and health. In both sectors trust is widely considered to be crucial. With regard to health, we already have a very long tradition of relying on experts, like physicians and pharmacists. However, it is also the case that it is long ago that the average consumer was able to fully assess all aspects of food consumption. This situation is not exclusive to food and health, yet their importance in our daily life makes these two sectors special. We value food and health

highly. In surveys, health is often reported as one of the most important aspect in life. Moreover, food is more than nourishment. Food can be important to express people's lifestyle and even can be directly related to one's identity (Caplan, 1997; Meijboom et al. 2003). Thus, when we have to trust another with regard to our food or health we have to entrust something highly personal and valuable.

Given the importance of trust for food and health, it seems reasonable to state that when food aspects are introduced into the medical sector or vice versa, the need for trust will not change. There are no indications that the introduction of health-related food products yields a situation in which consumers no longer need to rely on other stakeholders. Quite the contrary. There are indications that it can have a serious impact on trust. On the one hand, the addition of health aspects to food products is often a high-tech process that requires trust in the expertise and benevolence of a whole range of agents. On the other hand, the combination of health and food entails new roles and responsibilities for the involved professionals, institutions, and consumers. This easily results in what I have called, the problem of trust: People have to rely on others, but often do not know whom to trust. In the sections following I make use of the above-mentioned five steps to analyse this specific problem of trust and show the relevance and implications of each step.

7.3 Why the problem of trust is not to be analysed as a failure of the truster

Conceiving the problem of trust as a failure of the individual consumer clearly ignores three aspects of the context in which novel products at the interface between food and health are introduced. First, it ignores the dependent position of the truster. Most individuals lack the expertise needed to understand, let alone monitor the process of developing, producing and retailing of novel products. Even if they have the expertise, the number of agents involved in this process makes it impossible not to rely on others. Consequently, one has no option other than to rely on experts and institutions. This reliance enables a truster to act despite the evident lack of control, but makes him vulnerable to the motivation and competence of the trusted agents. Consequently, if one is uncertain about which expert or authority can be trusted, it makes trust not just a bit more complicated, but can make the difference between an attitude of trust and a state of uncertainty or even distrust. The trusted agent should take this vulnerability seriously in his response to what is entrusted. Despite his vulnerability the truster should be treated as a person who is capable of autonomous agency. Therefore, the vulnerability of the truster results in a demand for the trustee to take additional care and to take the truster's hesitations seriously. This illustrates that the problem of trust cannot be understood as a problem of the truster only.

Second, to lay the problem of trust with the consumer ignores the impact on trustworthiness of some recent scandals and problems in both the health and the food sector. In the Netherlands, recently hundreds of patients were confronted with the uncertainty of an infection with HIV and hepatitis B and C. Due to an omission in the cleaning procedure of a medical device the patient's blood could be contaminated. Fortunately, in none of the cases patients were infected. Another example is the estimated 20,000 deaths and 800,000 avoidable infections a year that could be prevented if hospitals were more careful with hygiene, as has been claimed by the German Society for hospital hygiene.[92] In both cases, all agents probably wanted and aimed to act in a trustworthy manner, the net result is that the trust of patients has been harmed. Another example is the MPA scandal: In 2002, an illegal hormonal growth promoter, the synthetic progesterone MPA, was found on Dutch pig farms. The scandal showed that some agents within the agricultural and food sector deliberately abused the reliance and trust of others. These examples may be exceptions, but they affect trust in the sector, and they show that the problem of trust is at least partly a problem of trustworthiness.

Third, a one-sided emphasis on the truster ignores the implicit claims of market institutions and governmental agencies which comes with the introduction and admission of a new product on the market. If a product at the interface between food and health is launched, there is often a claim about the quality, safety or added value of the product. It is, for instance, introduced as healthy, as a necessary supplement to one's daily food or as an instrument of preventive medicine. In each case there is a claim about what may be expected of this product and of its producer or of the government that allows the product to enter the market. In two ways these claims show that the problem of trust is more than a failure of the truster.

On the one hand, there is a problem on the level of communication. If trusters question the trustee's claim and do not immediately trust the producer or the government, it would be too easy to define the problem of trust as a problem of the individual truster only. The trustee has a problem too. If the claim related to the introduction of the product is legitimate, the trustee obviously failed to make this point sufficiently clear for the truster. This conclusion does not change if we acknowledge that trust includes an emotional element; we cannot discard the problem of trust as an irrational attitude of the truster. As I argued in Chapter 3, emotions are to a certain extent beyond rational control, but nonetheless we can open ourselves to argument, persuasion and evidence, because emotions are judgements. Therefore, they can be rational in the same sense in which judgements can be rational (Solomon, 1980, pp. 262; 270). Thus, the problem of trust is not just an indication of the irrational truster. The problem of trust is sincere, even though it is sometimes difficult to articulate the reason for it. It is

---

[92] Deutsche Gesellschaft für Krankenhaushygiene e.V. – DGKH (www.dgkh.de), accessed at 2 April 2008.

possible to ask the individual to reflect on his emotion, because it is directed to a specific trustee and includes a judgement about that agent. Consequently, the emotional aspect of trust has to be taken seriously as an evaluation of the competence or motivation of the trustee; the emotional should not be shrugged off as irrational.

On the other hand, the claims that institutions attach to products incorporate a statement about their own trustworthiness. Unless a trustee hopes that he will be trusted blindly, formulating claims about safety, health or other qualities of novel products reflect the institution's competence and motivation. To present a product as a tool for the prevention of cardiovascular diseases implies an invitation to the consumer to rely on this claim of the producer or the governmental agency that allows this kind of claims. This invitation creates expectations about the trustworthiness of the institution. This (implicit) invitation results in an obligation for the trustee. As I have argued in Chapter 6, this invitation includes a promise and making a promise implies the obligation for the institution to show respect for the truster and to take care that the truster is taken seriously despite his vulnerable position. This results in principles that require of the trustee that he does not manipulate the truster, prevents that 'reasonable, but false expectations' are raised, and that he takes reasonable steps to prevent any loss that a person 'would suffer as a result of the other's reliance on the expectations that one has intentionally or negligently led them to form' (Scanlon, 1990, p. 204). This obligation and the principles that follow from it illustrate that the problem of trust cannot be reduced to the problem of the truster. The trustee, too, has an evident responsibility.

7.4 Why a focus on risk is not sufficient

Novel food products that include health claims often raise questions about the safety of the product and about the legitimacy of the health claim. As a consequence, enabling consumers to deal with risks is regularly seen as an effective answer to questions of trust. The idea is that if individuals are able to address a danger as a risk, one has the opportunity to decide how to deal with the situation rather than the restricted choice to take or leave the danger (cf. Luhmann, 1988). This approach seems promising. If a situation of danger turns into one of risk, it becomes an object of action, because one has the option of making an assessment of the danger and of addressing it as a risk that one can analyse and manage. Consequently, one has the opportunity to decide on the best way to deal with the uncertain situation. Therefore, providing information on risks and enhancement of transparency is often proposed as the most efficient (regulatory) approach to this problem. Despite the importance of both aspects, there are two problems.

First, transparency and risk communication already presume some levels of trust. Only if one already considers the provider of information reliable, the information

becomes useful. For example, a person is uncertain about the safety and efficacy of a newly introduced nutraceutical claimed to lower blood pressure. Given this person's state of health, the claim sounds interesting, but he does not know whether the product is safe and whether the claim is valid. Consequently, he cannot make a deliberative decision about whether or not to use this product. To make the assessment of the claimed health benefits and the possible risks he needs information. This information can change the uncertain situation into one that can be assessed in terms of risk. However, information about this new product will only help if he already considers the provider of the information reliable. A report by, for instance, the Health Council on the safety and efficacy of nutraceuticals would only be of help to assess the risks and benefits at stake if he considers the institution a reliable source of information. Thus, on the basis of the available information, the person can make his personal assessment whether or not to use this new product. However, he can make this assessment only because he already relies on many other agents. Therefore, an overriding emphasis on risk communication only begs the question.

Second, I have argued that trust is fundamentally different from taking risks. Trust is not the outcome of an assessment of the risks and benefits of trusting in the light of the aims and goals one pursues. In contrast to someone who takes risks, a truster is not counting, but coping with complexity. He is not calculating risks, but dealing with the uncertainty he is faced with. Therefore, better risk assessment and more risk information do not necessarily lead to more trust. To help this person it can be useful to translate the problem of known uncertainty into one of risk. Consequently, he can make his personal assessment and decide whether it is worth to take the risk involved in relying on another agent or not. However, this risk–benefit analysis does not provide a direct reason to trust. It can only show that, given his preferences, the risk is worth taking.

Trust has a different focus. It starts where a risk focus ends. It arises in situations that remain uncertain despite the attempts to turn the uncertain aspects into risk factors. The aim is not to try to make a risk–benefit analysis, but to personally assess the competence and motivation of the trusted agent. When I trust the company which offers the nutraceutical, I do not assess risks. I do not even have the instruments to do so. However, I make an assessment of the company's competence and motivation with respect to this new food product. In this process of assessing the competence and motivation, risk information may contribute to trust only as far as it serves as a signal or is proof of the competence and motivation of the trustee. If the risk at stake is very low or the risk information is of a high quality, but I do not consider the other agent, on whom I have to rely, competent or adequately motivated, I will not trust him. Maybe I consider it worthwhile taking the risk and act nonetheless, in that case I am a risk taker, not a truster. This explains why risk information or improved risk analysis is not enough to deal with the problem of trust. It does not directly influence the level of trust. It only has direct influence on a mechanism to cope with uncertainty that is different from trust. In

building trust, risk information only plays a secondary role to the extent that it provides clarity about the trustworthiness of the other agent.

7.5 Why reliability is not enough

Reliability often seems to be enough to be trusted. Especially on an institutional level, a company or the government should simply do what they are established for and act in a reliable, i.e. predictable way. This often enables us to deal with complex and uncertain situations. For instance, if one has bought a product for many years to one's satisfaction, one considers the company reliable and expects that the product's safety and quality remain unchanged the next time one buys it. However, reliance based on predictability can be complicated, because we lack predictable patterns or because the available patterns conflict. In these cases being reliable is not enough. Then it is important that a trustee shows himself trustworthy, i.e. competent and adequately motivated. The problem of trust with respect to food products at the interface between food and health can clarify this point.

Trust in food and health products is often based upon a long history of clear patterns and routines. Most trust in food products and health products is currently based on the predictability of the stakeholders. Routines and patterns are often clear enough to act on although one is confronted with uncertainty and lack of personal control. There are long traditions and relatively clear norms that give consumers clarity on what he can expect of others when he buys food or a pharmaceutical product. However, we lack such clarity with regard to the developments at the interface between food and health. We sometimes simply lack clear patterns that show the reliability of an agent. Especially when new technologies, e.g. biotechnology or nanotechnology, are used in order to enhance the health effects of food products we often lack the predictability and familiarity that can serve as a first basis for trust. Partly because we are confronted with new benefits, unknown carry-over and long-term effects of novel technologies, for another part because we lack a history in which trustees could have proven their reliability. There is no clear pattern or history that only has to be explicated or revealed in order to show the truster that his reliance is warranted. It takes time before trust based on routine is achievable. Until that moment, the problem of trust remains, because it is unclear what to expect of the other party. To be trusted in such a situation, it becomes important to show that one is competent to deal with the new context and adequately motivated to respond to what is entrusted. It requires a more active stance from a trustee than is required for reliability.

Furthermore, reliability is not enough in the face of conflicting patterns and conflicting expectations. First, the available patterns that make an agent reliable can conflict. Both for food products and for health products patterns and routines are

available that show the reliability of a company or a governmental agency. However, novel food products with a health claim can be categorised in both groups. There is not one unambiguous pattern available upon which one can formulate trustful expectations. Trust is either based on patterns of the food sector, although the advice has a health claim, or it is based on patterns of the health domain, although it is a dietary product. Thus, the introduction of such dietary advice complicates pattern-based trust. This can be explicated with the example of trust and the health claims on the labels of food products. The idea of labels with special information is not new. We are used to all kinds of labels and claims regarding food products, like 'high fat', 'low sugar', or 'new formula.' Regarding pharmaceutical products, we have a similar situation in which the expectations concerning the instructions for use are clear. Yet labels that inform you about the health/enhancing aspects of food are new. Consequently, it is not quite clear what to expect. We can consider them as the usual food labels and adapt our expectations to what we normally expect of them, or we apply our expectations regarding the instructions for use of pharmaceutical products. This indistinctness can have a direct influence on trust in food and pharmaceutical products. For instance, if a label on a traditional product tells me that it is especially relevant for elderly people, I can think that it is safe and unproblematic to use it even when I do not belong to the target group. In the case of a pharmaceutical product such an indication will alert me. Even when it is prescribed by my GP, I will ask her whether this product is safe and effective for me. Thus, suppose a dairy product with a special hormone is introduced to the market that has substantial health effects for elderly consumers only, the expectations I can legitimately formulate regarding this product are unclear. In this situation, being reliable is not enough to address the problem of trust, because there is no clear routine that tells me whether the producer or retailer is reliable. Moreover, it also affects my expectations with regard to other food and pharmaceutical products, because I am no longer sure whether my 'normal' expectations still apply to those products.

Second, reliability is not enough, because the developments at the interface between food and health provide us with questions about of whom we expect what. Traditionally, health improvement has been a matter of physicians, pharmacists, and other health professionals. Currently, other parties – mostly market parties – have started to play an important role in this field as well. Since our expectations are both sector dependent and agent relative, this shift has implications for trust. In many cases, we trust not merely someone to do what we trust all others to do, but we require a specific action that we can reasonably expect of that person. However, our expectations differ from stakeholder to stakeholder. What one expects of a supermarket is often completely different from what is expected from a pharmacy. For instance, attention to taste and aesthetics is not something we expect of a pharmaceutical company, yet we trust food companies to pay a lot of attention to taste and design before introducing a new product. Likewise, we do not expect a greengrocer to ask every consumer that buys strawberries whether he or she

is allergic to them. For a physician this is quite different. When he knows that some medication can have adverse effects for those who have high blood pressure, we expect him to take account of this fact in prescribing other medication. Normally, these differences are not very problematic as long as it is clear whom one can expect to take care of the issues at stake. However, it becomes more difficult to determine what to expect of whom, if a greengrocer were to sell vegetables claiming real health benefits. The answer to this question is crucial for trust in relation to novel foods.

In sum, the case of food products with health claims illustrates that proving oneself reliable is not enough to deal with the problem of trust. If there is no routine or if routines conflict, a trusted agent should be able to show competence with regard to the new situation and adequate motivation to respond to what is entrusted.

7.6 The importance of including the moral dimension: Respecting the autonomy of the truster

The conflicting patterns mentioned in the section above are not just a practical problem which should be sorted out. It clearly signals a moral dimension. The differences in expectations we have with respect to agents who operate in the health sector and with respect to those who operate in the food sector are not merely based on routines. The differences have a moral background. In both sectors matters such as safety, autonomy and justice are valued. However, the interpretation of these values differs. For instance, for both the food sector and health sector, safety is paramount in every introduction of a product. Nevertheless, the specific interpretation of what safety means in relation to pharmaceutical products is different from food products. What we consider as an acceptable safety standard regarding food is considered insufficient for medicine, and we accept side effects of medication that we would never accept in the case of food. These differences are not just the result of different customs, but are the result of reflection on what we consider as a morally acceptable risk, given the aim of the product. This reflection can explain that we accept adverse effects of medication for diabetes, but that we would not accept such a negative effect from a slice of bread. The aim of treating diabetes is considered to be important enough to outweigh these problems, while we lack such an aim in the case of bread. Such reflections – not only on safety, but also concerning other issues like justice – are necessary to address the problem of trust with respect to novel food products with a health claim.

The duty to show respect for the truster as a person who is capable of autonomous agency is crucial if one aims to be trustworthy. Acknowledgment of the truster as a moral equal, which follows from this stance of respect, prevents that the trustee takes advantage of the vulnerable position of the truster by using only a minimal interpretation of the norms and principles or by manipulating him by giving one-sided or even false

information. It can be in the interest of a company to use the less stringent rules for safety as they are applied in the food sector, or it can be in the interest of a profession to adopt a protective attitude instead of giving consumers the option make an informed choice. However, respect for the autonomy of the truster implies that the mere fact that it is in the interest of the trustee cannot be a sufficient argument for the interpretation of the relevant norms and principles in that way. Despite his vulnerable position, the interests of the truster should be taken into account. Furthermore, acknowledgment of the truster as a moral agent results in the demand to take his moral judgement seriously. Thus, given the moral dimension in the introduction of food products with a health claim, the question of trust should not merely be addressed as a matter of risk or predictability. Given the questions with respect to the interpretation of moral values and norms and the distribution of responsibilities, the moral dimension marks the starting point for trustee reflection on the own competence and motivation as necessary condition to operate as a trustworthy partner.

7.7 Making trustworthiness operational

From the acknowledgement of the truster as moral agent and the recognition of the moral questions bound up with the introduction of these new food products, there is a clear need for the trustee to reflect on the competence and motivation that is necessary to operate as a trustworthy partner. Trust in the field of food and health needs clarity as to what we can reasonably expect of each other. This is not an easy task, given the broad range of conflicting beliefs about how principles should be interpreted in this context and given the uncertainty about the responsibilities of involved agents. Nonetheless, such reflection is necessary in order to address the problem of trust.

First, the process of reflection should start with preparedness to accommodate the views and expectations of the truster. This implies that the trustee is open to the other's moral view, is prepared to change his view and is willing to actively search for new ways to deal with the conflict. This requirement of accommodation is not the same as simply acting according to the truster's wishes. Respect for autonomy does not only imply that the trustee takes additional care on behalf of the truster, but also that he treats her as an autonomous person, i.e. as someone with whom one can reason and who can reflect on her own beliefs and expectations.

Second, it is essential for a trustee to reflect on the question of when accommodation results in acting according to the expectation of the truster and when it does not. Without such reflection the response of the trustee seems arbitrary. Acting in line with expectations of the truster can be conceived as opportunistic behaviour, and not responding in the expected way can be interpreted as hypocritical given the trustee's keenness on accommodation. Consequently, each action can be conceived as a sign of

untrustworthiness. To prevent this, it is important that the response of the trustee can be explained and substantiated in the light of the trustee's competence and motivation.

Third, a trustee should have a clear view on the own competence and that of the truster. To reflect on this point is essential in three ways. First, it can serve as tool of precaution. If a company that launches a food product with a health claim wants to remain trustworthy, it should reflect on the question whether it has the competence to deal with the moral questions that will arise with the introduction. This does not mean that that the company must have the competence to answer every possible question, but that it is aware of the moral dimension and prepared for a debate on this. Second, reflection on one's competence is important, because of its consequences for the assumed competence of the truster. Suppose that a company does not have the competence to deal with the consumer concerns with respect to their own health status. Nonetheless, it aims to introduce a specific dietary advice with the quality to reduce the risk of a certain ailment. The company does so, because it believes that consumers get more information and are provided with more tools and therefore are in the position to be responsible for their own health status. Apart from the legitimacy of this view, it shows that the institution's ideas about its own competence have direct influence on what the trustee expects of the truster. Third, reflection on the institution's competence can prevent that (implicit) promises are made that cannot be fulfilled. Trust implies an appeal to respond and respect for the autonomy of the truster translates into strong reasons to respond. Nonetheless, this cannot be a reason to respond in a way that is not consistent with the company's competence. If there is a debate about the content and distribution of responsibilities with regard to the production and retail of nutraceuticals, the government is often proposed as the stakeholder that should intervene. Given the public dimension of both food and health issues this seems reasonable. Nonetheless, a government cannot take on all of the responsibilities that other agents are not prepared to accept. If they were to do so, they cannot manage to live up to the expectations and consequently cannot be trustworthy. Thus, even though it seems reasonable to expect that there is clear regulation on issues like safety, health claims, and other issues, the government should have a clear view on where and why there are constraints on its competence. Otherwise, it is burdened with a weight it cannot bear and runs the risk of being accused of untrustworthiness.

Fourth, as I have argued, essential to trustworthiness is integrity. Given the plurality of views of the many trusters, a trustee on an institutional level often cannot act according to the expectations of one truster without leaving the expectations of another truster unanswered. To justify one's response, an institution needs to have a clear view on its main aims and tasks and on the principles that follow from it. Integrity, understood as the sincere commitment to the institution's core tasks and aims, can help to get such a clear view. Consequently, it is possible to formulate reflected constraints on the demand

of accommodation, but also to clarify where there are possibilities for accommodation that do not affect the institution's trustworthiness.

For instance, a company that operates in the free market and introduces a nutraceutical is confronted with many expectations about the affordability of the product, the validity of the claim, the interest rates, etc. To satisfy all these expectations is impossible. Given the costs of development, the product will not be affordable for all consumers. On the other hand, the shareholder's expectations about the interest rates of the company conflict with a broad and sound study on the validity of the health claim. If the company aims to be trustworthy, it has to strive for a balance between accommodation to the truster's expectations and the integrity of the company in terms of consistency and commitment to its core. On the one hand, this illustrates that blindly acting according to what is expected does not make the corporation trustworthy. On the other hand, if the core of the corporation is defined as making money as fast as possible, this cannot be done at all costs, because this would not take the requirement of accommodation seriously. Striving for a balance can imply that the company decides to compromise on its core aim by investing in the validity of the health claim, because the company takes the vulnerability of the consumer seriously. Then again the company can decide that its aim to make money puts a constraint on the need to reduce the price of the product to make it affordable for all consumers. This balance may not result in the trust of all who formulated expectations towards the company. Nonetheless, it helps to clarify what can be expected of the company and how it is motivated by both the respect for the truster and its own integrity.

Fifth, on the basis of respect for the truster and the ability to give reasons for the way the trustee responds, the trustee should communicate with truster. This implies communication and transparency on:

- the actual competences of the institution, but also on the reasons that underlie the choice for developing these competences,
- the motivation that underlies the response to the truster, which should include an explication of how this motivation incorporates both respect for the truster as an autonomous agent and the trustee's commitment to the core of the institution and to the principles and tasks this entails.

This communication is *possible* as a way to address the problem of trust, because the truster is to be treated as an autonomous person with whom one can reason and who can reflect on her own beliefs and expectations. Furthermore, it is possible because the trustees have reflected ideas about why they consider themselves competent enough to be trusted and because they can provide reasons for their position on the balance between accommodation and integrity.

Moreover, this communication of and transparency about the trustee's reflected opinion on its competence and motivation is *essential* to address the problem of trust. In this study, the problem of trust has been defined as the problem that people have to rely

on others, but often do not know whom to trust. As the need to rely can hardly be reduced, given the complexity of our society, the most relevant way to address this problem is to reduce the uncertainty about who can be trusted. In the problem of trust, trust is at stake, but in its essence it is a matter of trustworthiness. Therefore, in communicating on the reflected ideas of the trustee about its competence and motivation, the trustee can show itself a trustworthy partner. The combination of the level of reflection on trustee's competence, the start in respect for autonomy of the truster and the trustee's position on the balance between accommodation and integrity enable a trustee to go beyond reliability.

This helps a truster to assess who can be trusted a) in those situations that remain uncertain even after the uncertain aspects have turned into risk factors as much as possible, b) when predictability fails, c) when conflicting moral values are at stake, and d) when the hesitance to trust starts in the belief that one is disrespected as an autonomous person. In these cases the trustee's reflection and communication on its competence and motivation can make the difference. The communication provides the truster with crucial information for the assessment of the trustee. Moreover, this communication creates expectations about the future behaviour of the trustee, which result in principles that require the trustee to protect the truster from manipulation and serious losses. This does not solve all problems of trust, but provides clarity with respect to trustworthiness, which is a necessary condition to build and maintain trust.

BIBLIOGRAPHY

Adams, R.M., 1987, *The Virtue of Faith And Other Essays in Philosophical Theology*, Oxford: Oxford University Press

Ames, W., 1628, *Medulla Theologiae* (Marrow of Theology), Amsterdam: Johannes Jansonnius

Ashman, I. Winstanley, D, 2007, 'For or Against Corporate Identity? Personification and the Problem of Moral Agency', *Journal of Business Ethics*, 76/1, pp. 83-95

Arrow, K. J., 1974, *The Limits of Organization*, New York: W.W. Norton

Audi, R., 2004, *The good in the right; A theory of intuition and intrinsic value*, Princeton: Princeton UP

Bacharach, M., 2002, 'How human trusters assess trustworthiness in quasi-virtual contexts', in: R. Falcone et al. (eds.), *Trust, Reputation, and Security: Theories and Practice*, AAMAS 2002 International Workshop, Berlin: Springer, pp. 1-17

Baier, A.C., 1994a [1986], 'Trust and antitrust', in: A.C. Baier, *Moral Prejudices; Essays on ethics*, Cambridge: Harvard University Press, pp. 95-129

Baier, A.C., 1994b, 'Trust and its vulnerabilities', in: A.C. Baier, *Moral Prejudice;, Essays on ethics*, Cambridge: Harvard University Press, pp. 129-151

Baier, A.C., 1994c, 'Sustaining trust', in: A.C. Baier, *Moral Prejudices; Essays on ethics, Cambridge*: Harvard University Press, pp. 152-182

Baier, A.C., 1991, The Tanner Lectures on Human Values 13, Princeton University, pp. 107-174

Baier, A.C., 1985, *Postures of the mind; Essays on minds and morals*, Minneapolis: University of Minnesota Press

Baker, J., 1987, 'Trust and Rationality', *Pacific Philosophical Quarterly*, 68, pp. l-13

Barrera, D., 2005, *Trust in embedded settings*, dissertation Utrecht University, Veenendaal: Universal Press

Bateson, P. (1988) 'The Biological Evolution of Cooperation and Trust', in: Gambetta, D. (ed.), *Trust: Making and breaking cooperative relations*, Oxford: Basil Blackwell, pp. 14-30

BBC, 2002, 'Hormone fear forces Dutch slaughter, Tuesday, 6 August, 2002, http://news.bbc.co.uk/1/hi/world/europe/2176481.stm (accessed 28 March 2007)

Beauchamp, T.L., 1991, *Philosophical ethics; An introduction to moral philosophy*, second edition, New York: McGraw-Hill

Beauchamp, T., Childress, J.F., 1994, *Principles of Biomedical Ethics* (Fourth Edition), New York: Oxford University Press

Becker, L.C., 1996, 'Trust as noncognitive security about motives', *Ethics*, 107, pp. 43-61

Beekman, V., 2008, 'Consumer Rights to Informed Choice on the Food Market', *Ethical Theory & Moral Practice*, 11/1, pp. 61-72

Beekman, V., F. W. A. Brom, 2007, 'Ethical tools to support systematic public deliberations about the ethical aspects of agricultural biotechnologies', *Journal of Agricultural and Environmental Ethics*, 20/1, pp. 3-12

Beekman, V., 2001, *The green third way? Philosophical reflections on government intervention in non-sustainable lifestyles*, dissertation Wageningen UR

Beekman, V., F.W. A. Brom, 2000, 'Vertrouwen is de achilleshiel van de voedselketen', *Nederlandse Staatscourant*, nr. 91

Beld, T. van den, 2001, 'The Morality System with and without God', *Ethical Theory and Moral Practice*, 4, pp. 383–399

Benjamin, M., 2001, 'Between Subway and Spaceship: Practical Ethics at the Outset of the Twenty-first Century', Hastings Center Report, 31/4, pp. 24-31

Benjamin, M., 1990, *Splitting the Difference, Compromise and Integrity in Ethics and Politics,* Kansas: University Press of Kansas

Berkhof, H., 1993, *Christelijk geloof : een inleiding tot de geloofsleer*, 6^e druk Nijkerk: Callenbach

Bernstein, P.L., 1998, *Against the gods: the remarkable story of risk*, New York: John Wiley & Sons

Birchard, K., 2002 'Europe investigates hormone-contaminated waste in food chain', *The Lancet*, 360/ 9328 p. 235

Blackburn, S., 1999, 'Am I Right?, Book Review of What We Owe to Each Other by T. M. Scanlon', *New York Times*, 21 February 1999, late edition

Blackburn, S., 1998a, *Ruling Passions; A theory of practical reasoning*, Oxford: Clarendon Press

Blackburn, S., 1998b, 'Trust, cooperation, and human psychology', in: V. Braithwaite & M. Levi (eds.), *Trust and Governance*, New York: Russell Sage Foundation, pp. 28-45

Brom, F.W.A., T. Visak, F. Meijboom, 2007, 'Food, citizens and market: the quest for responsible consuming', in: L. Frewer and H. van Trijp (eds.), *Understanding consumers of food products,* Cambridge: Woodhead Publishing, pp. 610-623

Brom, F.W.A., 2002, 'Food safety: science and ethics', Background paper for the FAO expert consultation on 'food safety: science and ethics', FAO, Rome 3-5 September 2002

Brom, F.W.A., 2000. 'Food, Consumer Concerns and Trust: Food Ethics for a Globalizing Market', *Journal of Agricultural and Environmental Ethics*, 12/2, pp. 127-139

Brom, F.W.A., 1998, 'Developing Public Morality: Between Practical Agreement and Intersubjective Reflective Equilibrium', In: W. van der Burg & T. van Willigenburg (eds.), *Reflective Equilibrium*, Dordrecht: Kluwer Academic Publishers, pp. 191-202

Brown, J.R., 2000, 'Social Factors in Science', in: W.H. Newton-Smith (ed.), *A Companion to the Philosophy of Science*, Oxford: Blackwell Publishing, pp. 442-450

Busch, L., 1997, 'Grades and Standards in the Social Construction of Safe Food', Invited paper presented at a conference on The Social Construction of Safe Food, Trondheim, Norway, April, 1997, http://www.msu.edu/user/ifas/Food%20Safety%20Word.doc (accessed at 7 October 2004)

Busch. R.J., A. Haniel, N. Knoepffler, G. Wenzel (eds.), 2002, *Grüne Gentechnik. Ein Bewertungsmodell*, München: Herbert Utz Verlag

Buskens, V., W. Raub, 2002, 'Embedded Trust: Control and Learning', Group Cohesion, *Trust and Solidarity*, 19, pp. 167–202

Buss, S., 2002, 'Valuing Autonomy and Respecting Persons: Manipulation, Seduction, and the Basis of Moral Constraints', *Ethics*, 115, pp. 195–235

Byrne, D., 2002, EFSA: Excellence, integrity and openness, Speech at the Inaugural meeting of the Management Board of the European Food Safety Authority, Brussels, 18 September 2002, www.efsa.eu.int (accessed at 8 September 2006)

Byrne, D., 2004, 'The Regulation of Food Safety and the Use of Traceability/ Tracing in the EU and USA: Convergence or Divergence?', Speech during the Food Safety Conference, Washington DC, 19 March 2004, http://europa.eu.int/comm/dgs/health_consumer/library/speeches/speech168_en.pdf (accessed 8 September 2006)

BZK, Ministerie van Binnenlandse Zaken en Koninkrijksrelaties, 1999, *Vertrouwen in Verantwoordelijkheid,* Deel 1 in de reeks 'Vertrouwen in verantwoordelijkheid', Den Haag, http://www.minbzk.nl/contents/pages/5063/vertrouwen_verantwoordelijkheid_9-99.pdf (accessed 29 January 2007)

Calhoun, C., 1995, 'Standing for Something', *The Journal of Philosophy*, 92/5, pp. 235-260.

Calligan, C., 1993, *In a Different Voice: Psychological Theory and Women's Development*, Cambridge: Harvard UP

Caplan, P. (ed.), 1997, *Food, Health and Identity,* London: Routledge

CEC Commission of the European Communities, 2000, *White Paper on food safety*, Brussels 12 January 2000, COM (2000) 719 final

Cohen, L.J., 1989, 'Belief and Acceptance', *Mind*, 98/391, pp. 367-389

Coleman, J.S., 1990, *Foundations of Social Theory*, Cambridge: Belknap Press

Coleman, J.S., 1983, 'Recontracting, trustworthiness, and the stability of vote exchanges. Why so much stability?', *Public Choice,* 40, pp. 89-94

Commissie Wijffels, 2001, *Toekomst van de veehouderij, agenda voor een herontwerp van een sector,* Den Haag: Ministerie van Landbouw, Natuurbeheer en Visserij

Cordner, C., 2007, 'Three contemporary perspectives on moral philosophy', *Philosophical investigations*, 31/1, pp. 65-84

Council for the Rural Areas, 1998*, Care and Trust', the basis for food production in the twenty-first century*, Report 98/7 a.

Covaci, A., K. Grob, 2002, 'Mineral oil and PCB/dioxin analysis in some European food contamination episodes', *European Food Research and Technology*, 215/1, pp. 51–54

Creel, R.E., 2001, *Thinking Philosophically, An Introduction to Critical Reflection and Rational Dialogue*, Oxford: Blackwell

Dancy, J., 2004, *Ethics Without Principles,* Oxford: Oxford University Press

Dasgupta, P., 1988, 'Trust as a Commodity', in: Gambetta, D. (ed.), *Trust: Making and breaking cooperative relations*, Oxford: Basil Blackwell, pp. 49-72

De Sousa, R., 2001, 'Moral emotions', *Ethical Theory and Moral Practice*, 4, pp. 109-126

Dekker, P., T. van der Meer, 2004, 'Politiek vertrouwen 1997-2004', *Tijdschrift voor de Sociale Sector*, december 2004, pp. 33-35

Deutsch, M., 1958, 'Trust and Suspicion', *The Journal of Conflict Resolution*, 2/4, 265-279

Diederen, P., 2003, 'Burger, laat die consument met rust!' in: H. Dagevos, L. Sterrenberg, *Burgers en consumenten – tussen tweedeling en twee-eenheid*, Wageningen: Wageningen AP, pp. 19-30

Dower, N., 1996, in: B. Mepham (ed.), 'Global Hunger', *Food Ethics*, London: Routledge, pp. 1-17

Dreyer, M., O. Renn, 2007, 'Developing a coherent European food safety policy: The challenge of value-based conflicts to EU food safety governance', in: L. Frewer and H. van Trijp (eds.), *Understanding consumers of food products*, Cambridge: Woodhead Publishing, pp. 534-557

Dubbink, W., 2007, 'Transparency Gained, Morality Lost. A Critique of the Administrative Conceptualization of CSR, Illustrated by Dutch Policy', *Business and Society Review*, 112/2, pp. 287-313

Dunn, J., 1988, 'Trust and Political Agency', in: Gambetta, D. (ed.), *Trust: Making and breaking cooperative relations*, Oxford: Basil Blackwell, pp. 73-93

Düwell, M., 2002, 'Handlungsreflexive Moralbegründung', in: M. Düwell, C. Hübenthal & M.H. Werner (eds.), *Handbuch Ethik,* Stuttgart: Verlag J.B. Metzler, pp. 152-162

Eckel, C. C., R. K. Wilson, 2004, 'Is trust a risky decision?', *Journal of Economic Behavior & Organization*, 55, pp. 447–465

EFSA, 2006, *Corporate Brochure*, EFSA: Parma, http://www.efsa.europa.eu/etc/medialib/efsa/publications.Par.0005.File.dat/Brochure _2006 apprALG.pdf (accessed 12 February 2007)

EFSA, 2005, *Annual Report 2004, European Food Safety Authority*, EFSA: Brussels

EFSA, 2004, *EFSA News, The Newsletter of the European Food Safety Authority*, 1/1

Egestad, P.S., 2002, *Trustful relations, a perspective on trust in actor relations in forestry*, Dissertation, Wageningen University

Elster, J., 1989, *Nuts and Bolts for the Social Sciences,* Cambridge: Cambridge University Press

Erikson, E., 1963, *Childhood and Society,* New York: Norton

Etzioni, A., 2001a, *The Monochrome society*, Oxford: Princeton University Press

Etzioni, A., 2001b, 'The third way to a good society', *Sociale wetenschappen,* 44/ 3, pp. 5-40

Etzioni, A., 1993, *The spirit of community: Rights, responsibilities, and the communitarian agenda,* New York: Crown

FAO, 2003, *Expert Consultation on Food Safety: Science and Ethics*, Rome, 3-5 September 2002, Rome: FAO

FAO/WHO, 1998, *Report of Joint FAO/WHO Expert Consultation on Application of Risk Communication to Food Standards and Safety Matters*, Rome, 2-6 February 1998, Geneva: WHO

FAO/WHO, 1997, *Risk management and food safety*. Report of a Joint FAO/WHO Consultation Rome, Italy, 27 to 31 January 1997, FAO Food and Nutrition Paper 65, Rome: FAO

FAO/WHO, 1995, *Report of an Expert Consultation on Application of Risk Analysis to Food Standards Issues*, Geneva, 13-17 March 1995, WHO/FNU/FOS/95.3. Geneva: WHO

Faulkner, P., 2003, 'The Epistemic role of trust', in: R. Falcone et al. (eds.), *Trust, Reputation,and Security: Theories and Practice,* AAMAS 2002 International Workshop, Berlin: Springer, pp. 30-38

FAVV, 2000, *FAVV-business Plan*, Brussel, 13 November 2000, http://www.favv-afsca.fgov.be/images/cereus/nl/pdf/aboutAfsca/BUSFR.pdf, (last accessed at 25 October 2005)

Firth, P., 1999, 'Consuming fears', *Scientific American*, 280/1, pp. 41-43

Fischer, A.R.H., A.E.I. de Jong, R. de Jonge, L.J. Frewer, M.J. Nauta, 2005, 'Improving Food Safety in the Domestic Environment: The Need for a Transdisciplinary Approach', *Risk Analysis,* 25/3, pp. 503-517

Flood, M. M., 1958, 'Some Experimental Games', *Management Science*, 5/1, pp. 5-26

Forsythe R. H., 1993, 'Risk: Reality Versus Perception', *Poultry Science*, 72/6, pp. 1152-1156

Frankfurt, H., 2005. 'I. Taking Ourselves Seriously & II. Getting It Right', *The Tanner Lectures on Human Values*, 25, Stanford University, pp. 169-202

Frankfurt, H., 2002 'Reply to Eleonore Stump,' in: S. Buss and L. Overton (eds.), *Contours of Agency: Essays on Themes from Harry Frankfurt*, Cambridge: MIT Press, pp. 61-63

Frankfurt, H., 1982, 'The Importance of what we care about', *Synthese*, 53, pp. 257-272.

Frankfurt, H., 1978 'The Problem of Action', *American Philosophical Quarterly* 15, pp. 157-162

Frankel, M.S., 1989, 'Professional Codes: Why, How, and with What Impact?', *Journal of Business Ethics,* 8, pp. 109-115

Frevert, U. (ed.), 2003, *Vertrauen. Historische Annäherungen,* Vandenhoeck & Ruprecht

Frewer, L., 2004, *Consumers, Food, Trust and Safety. The Need for Collaboration Between the Social and Natural Sciences*. Inaugural Lecture Wageningen University, 27 May 2004, Wageningen: WUR

Frewer, L., 2000, 'Risk perception and risk communication about food safety issues', *Nutrition Bulletin*, 25, pp. 31-33

Frewer, L., J. Lassen, B. Kettlitz, J. Scholderer, V. Beekman, K.G. Berdal, 2004, 'Societal aspects of genetically modified foods', *Food and Chemical Toxicology,* 42, pp. 1181–1193

Frewer, L. J., J. Scholderer, L. Bredahl, 2003a, 'Communicating about the Risks and Benefits of Genetically Modified Foods: The Mediating Role of Trust', *Risk Analysis*, 23/6, pp. 1117-1133

Frewer, L.J., S. Hunt, M. Brennan, S. Kuznesof, M. Ness, Ch. Ritson, 2003b, 'The views of scientific experts on how the public conceptualize uncertainty', *Journal of Risk Research*, 6/1, pp. 75-85

Frewer, L.J, B. Salter, B., 2003c, 'The changing governance of biotechnology: The politics of public trust in the agri-food sector', *Applied Biotechnology, Food Science and Policy*, 1/4, pp. 199-211

Friedman, M., 1970, 'The Social Responsibility of Business Is to Increase Its Profits', *The New York Times Magazine*, September 13, 1970

Frijda, N.H., 1986, *The Emotions,* Cambridge: Cambridge University Press

Frijda, N.H., A.S.R. Manstead, S. Bem (eds.), 2000, *Emotions and beliefs: How feelings influence thoughts,* Cambridge: Cambridge University Press. FSA, 2001, *Strategic Plan 2001 – 2006, Putting consumers first*, London: FSA

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

Gambetta, D. (ed.), 1988a, *Trust: Making and breaking cooperative relations*, Oxford: Basil Blackwell

Gambetta, D., 1988b, 'Can We Trust Trust?', in: Gambetta, D. (ed.), *Trust: Making and breaking cooperative relations*, Oxford: Basil Blackwell, pp. 213-237

Gauthier, D., 1979, *The logic of the Leviathan*, Oxford: Clarendon Press

Gaskell, G., N. Allum, W. Wagner, N. Kronberger, H. Torgersen, J. Hampel, J. Bardes, 2004, 'GM Foods and the Misperception of Risk Perception', *Risk Analysis*, 24/1, pp. 185-193

Gaskell, G., N. Allum, S. Stares, 2003, *Europeans and Biotechnology in 2002; Eurobarometer 58.0* (2nd Edition: March 21st 2003), A report to the EC Directorate General for Research from the project 'Life Sciences in European Society' QLG7-CT-1999-00286

Giddens, A., 1991, *Modernity and self-Identity: Self and society in the late modern age*, Cambridge: Polity Press

Giddens, A., 1990, *The Consequences of Modernity*, Cambridge: Polity Press

Gofton, L., 1996, 'Bread to Biotechnology: Cultural Aspects of Food Ethics,' in: B. Mepham (ed.), *Food Ethics,* London: Routledge, pp. 120–137

Govier, T., 1998, *Dilemmas of trust*, Montreal & Kingston: McGill-Queen's University Press

Govier, T., 1997, *Social Trust and Human Communities*, Montreal & Kingston: McGill-Queen's University Press.

Gregerson, N.H., 2003, 'Risk and Religion: Toward a Theology of Risk Taking', *Zygon*, 38/2, pp. 355-376

Gewirth, A., 1983, 'The rationality of reasonableness', *Synthese*, 53, pp. 225-247

Guleserian, T., 2001, 'Can God be trusted?', *Philosophical Studies,* 106, pp. 293-303

Gutmann A., D. Thompson, 1990, 'Moral Conflict and Political Consensus', *Ethics*, 101/1, pp. 64-88

Harbers, H., 2003, 'Vertrouwen, ter inleiding', *Krisis, tijdschrift voor empirische filosofie*, 4/1, 3-7

Hardin, R., 2006, *Trust*, Cambridge: Polity Press

Hardin, R. (ed.), 2004b, *Distrust*, New York: Russell Sage Foundation

Hardin, R., 2004a, 'Why Know?' Lecture at the University of Waterloo, April 5, 2004, http://www.nyu.edu/gsas/dept/politics/faculty/hardin/research/Why_know.doc (accessed 13 December 2006),

Hardin, R. 2003, 'If It Rained Knowledge', *Philosophy of the Social Sciences*, 33/1, pp. 3-24

Hardin, R., 2002a, *Trust and Trustworthiness*, New York: Russell Sage Foundation

Hardin, R., 2002b, 'Street-Level Epistemology and Democratic Participation', *Journal of Political Philosophy,* 10/2, pp. 212–229

Hardin, R., 2000, 'Trust and Society', in: Galeotti, G., et al. (eds.), *Competition and structure,* Cambridge: Cambridge University Press, pp. 17-46

Hardin, R., 1999, 'Book Review, Trudy Govier, Social Trust and Human Communities', *the Journal of Value Inquiry*, 33, pp. 429-433

Hardin, R., 1998, 'Trust in Government', in: V. Braithwaite & M. Levi (eds.), *Trust and Governance*, New York: Russell Sage Foundation, pp. 9-27

Hardin, R., 1996, 'Trustworthiness', *Ethics*, 107, 26-42

Hardin, R., 1993, 'The street-level epistemology of trust', *Politics & Society*, 21, 505-529

Hardin, R., 1991, 'Hobbesian Political Order', *Political Theory*, 19/2, p. 156-80

Hartmann, M., 2001, 'Einleiting', in: Hartmann, M., C. Offe (Hg.), *Vertrauen, die Grundlegung des sozialen Zusammenhalts,* pp. 7-34

Hartogh, G. den, 2002, *Mutual Expectations, A Conventional Theory of Law*, The Hague: Kluwer Law International

Hayward, A.C.D., 1997, 'Consumer Perception', in: Tennant, D.R. (ed.), *Food Chemical Risk Analysis*, London: Chapman & Hall, pp. 336-361

Heeger, R., F.W.A. Brom, 2003, 'Ernährung, Verbraucher und Staat', in: Düwell, M. (ed.), *Bioethik*, Suhrkamp Verlag, pp. 371-377

Hees, M. van, J. Vromen, 2002, 'De vele gedaantes van de rationele-keuzetheorie' [The many faces of rational choice theory], *Algemeen Nederlands Tijdschrift voor Wijsbegeerte,* 94/1, pp. 3-15

Hertzberg, L., 1988, 'On the attitude of trust', *Inquiry,* 31, pp. 307-322

Hill, T. E., 1994, *Respect for Humanity,* The Tanner Lectures On Human Values, Delivered at Stanford University April 26 and 28, 1994, pp. 1-76

Hill, T.E., 1992a, 'Kantian Pluralism', *Ethics*, 102/4, pp. 743-762

Hobbes, T., 1651, *Leviathan or the Matter, Forme, & Power of a Common-wealth, Ecclesiasticall and Civill*, London: Andrew Crooke, http://socserv2.socsci.mcmaster.ca/~econ/ugcm/3ll3/hobbes/Leviathan.pdf

Hobbes, T., 1640, *The Elements of Law Natural and Politic*, McMaster University, http://socserv2.socsci.mcmaster.ca/~econ/ugcm/3ll3/hobbes/elelaw

Hollis, M., 1998, *Trust within reason*, Cambridge: Cambridge University Press

Hollis, M., 1994, *The philosophy of social science. An introduction*, Cambridge: Cambridge University Press

Holt, C.A., A. E. Roth, 2004, 'The Nash equilibrium: A perspective', *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*, 101/ 2, pp. 3999–4002

Holton, R., 1994, 'Deciding to trust, coming to believe', *Australian Journal of Philosophy*, 72, pp. 63-76

Horsburgh, H.J.N., 1960, 'The Ethics of Trust', *Philosophical Quarterly*, 10 pp. 343–354

Hosmer, L. T., 1995, 'Trust: the connecting link between organizational theory and philosophical ethics', *The academy of management review*, 20/ 2, pp. 379-403

House of Lords, 2000, Select Committee appointed to consider Science and Technology, Science and Technology - Third Report, 23 February 2000, www.parliament.the-stationery-office.co.uk/pa/ld199900/ldselect/ldsctech/38/3801.htm (accessed 13 December 2006)

Hoven, M.A. van den, 2006, *A Claim for Reasonable Morality. Commonsense Morality in the Debate on the Limits of Morality*, Dissertation Utrecht University

Hume, D., 2007 [1739-40], *A Treatise of Human Nature: Being An Attempt to introduce the experimental Method of Reasoning into Moral Subjects*, Oxford Philosophical Texts, edited by David Fate Norton & Mary J. Norton, Oxford: Oxford University Press

Hume, D., 1978, *A Treatise of human understanding*, Oxford: Clarendon Press

Jasanoff, S., 2001, 'Citizens at Risk: Reflections on the US and EU' in: Pasquali. M. (ed.): *EurSafe 2001. Food Safety, Food Quality, Food Ethics. Preprints*, pp. 43-49

Jensen, K.K., P. Sandøe, 2002, 'Food Safety and ethics: The interplay between science and values', *Journal of Agricultural and Environmental Ethics*, 15/3, pp. 245-253

Johnson, P., 1993, *Frames of deceit: A study of the loss and recovery of public and private trust*. Cambridge: University of Cambridge Press

Jones, K., 1998, 'Trust', in: *the Routledge Ecyclopedia of Philosophy*, London: Routledge, pp. 466-470

Jones, K., 1996, 'Trust as an affective attitude', *Ethics*, 107, pp. 4-25

Jong, G. de, R. Klein Woolthuis, 2004, *The Effects of Trust on Performance of High-Tech Business Relationships*, SOM - reports University of Groningen, http://irs.ub.rug.nl/ppn/265446481 (accessed 5 July 2006)

Jonge J. de, E. van Kleef, L. Frewer and O. Renn, 2007, 'Perception of risk, benefit and trust', in: L. Frewer and H. van Trijp (eds.), *Understanding consumers of food products*, Cambridge: Woodhead Publishing, pp. 125-150

Jonge, J. de, H.P.J.M. Noteborn, E.G. Schouten, J. Timmers, 2005, *VWA Consumentenmonitor 2004, onderzoek naar het vertrouwen van de consument in de veiligheid van voedingsmiddelen',* 5 juli 2005, Den Haag VWA

Kant, I., 1999 [1785], *Groundwork of the Metaphysics of Morals,* translated and edited by Mary Gregor, in the Cambridge Texts in the History of Philosophy series. Cambridge: Cambridge University Press

Kant, I., 1996 [1797], *The Metaphysics of Morals*, translated and edited by Mary Gregor, in the Cambridge Texts in the History of Philosophy series. Cambridge: Cambridge University Press

Kettlitz, B., 2001, Consumer perception of food risk and consumer trust, *Preprints of the Third Congress of the European Society for Agricultural and Food Ethics, 3-5 October 2001,* pp. 59-66

Kierkegaard, S., 1992, *Concluding Unscientific Postscript to Philosophical Fragments*. Edited and Translated by Howard V. Hong and Edna H. Hong. Princeton: Princeton University Press

Kjaernes, U., A. Warde, M. Harvey, 2006, 'Politicising consumer trust in food: A socio-institutional explanation to variations in trust', in: Kaiser, M. & M.E. Lien (eds.), *Ethics and the politics of food*, Wageningen University Press, pp. 162-166

Kjaernes, U., 2004, Press Note, Final Conference, Consumer Trust in Food, A European Research Project on Social and Institutional Conditions for the Production of Trust, www.trustinfood.org (last accessed 5 July 2006)

Korthals, M., 2004, *Before Dinner, Philosophy and Ethics of Food*, Dordrecht: Springer

Korthals, M., 2001, 'Taking consumers seriously: two concepts of consumer sovereignty', *Journal of Agricultural and Environmental Ethics* 14/2, pp. 201-215

Korsgaard, C.M., 2004, 'Fellow Creatures: Kantian Ethics and Our Duties to Animals', *Tanner Lectures on Human Values*, University of Michigan, February 6, 2004, Volume 25 Korsgaard, C.M., 1996, *The sources of normativity*, Cambridge: Cambridge University Press

Kramer, R.M., T.R. Tyler (eds.), 1996, *Trust in organizations: Frontiers of theory and research*, Thousand Oaks: Sage

Krebs, J., 2001, 'Building Consumer Confidence: The UK Experience', lecture during the Third Congress of the European Society for Agricultural and Food Ethics, EurSafe 2001, 3-5 October, Florence

Künast, R., 2003, 'Rede der Bundesministerin für Verbraucherschutz, Ernährung und Landwirtschaft Renate Künast, Anlass: Amtseinführung des Leiters des Bundesinstituts für Risikobewertung, Herr Prof. Hensel',

http://www.bfr.bund.de/cm/232/rede_der_bundesministerin.pdf (last accessed 10 February 2006)

Kymlicka, W., 2002, *Contemporary Political Philosophy, an introduction, second edition*, Oxford: Oxford University Press

Lagerspetz, O., 1998, *Trust: The Tacit Demand*, Dordrecht: Kluwer Academic Publishers

Lahno, B., 2002a, *Der Begriff des Vetrauens*, Paderborn: Mentis

Lahno, B., 2002b, 'Institutional Trust: A Less Demanding Form of Trust?', http://www.uni-duisburg.de/FB1/PHILO/index/Lahno-Trust.htm (accessed 15 December 2006)

Lahno, B., 2001, 'On the emotional character of trust', *Ethical Theory and Moral Practice*, 4, pp. 171-189

Lahno, B., 1995a, 'Trust and strategic rationality', *Rationality and Society*, 7/4, pp. 442-464

Lahno, B., 1995b, 'Trust, Reputation, and Exit in Exchange Relationships', *The Journal of Conflict Resolution*, 39/3, pp. 495-510

Lang, T., 1999, 'The complexities of globalization: The UK as a case study of tensions within the food system and the challenge to food policy, *Agriculture and Human Values*, 16, pp. 169-185

Lewis, J.D., A. Weigert, 1985, 'Trust as a Social Reality', *Social Force*, 63/4, pp. 967-985

Lindenberg, S., 2000, 'It takes both trust and lack of mistrust: the workings of cooperation and relational signaling in contractual relationships', *Journal of Management and Governance*, 4, pp. 11-33

LNV/ FAO, 2004, 'European Food Safety Control Systems: New perspectives on a harmonized legal basis', paper prepared by the Ministry of Agriculture, nature and Food Quality and the Food and Consumer Product Safety Authority of the Netherlands, Second FAO/WHO Global Forum for Food Safety Regulators, 12-14 October 2004, http://www.fao.org/docrep/meeting/008/j2975e.htm (Accessed 15 December 2006)

Locke, J., 1960 [1689-90], *Two Treatises of government, A critical edition with an introduction and apparatus criticus by Peter Laslett*, Cambridge: Cambridge University Press

Løgstrup, K.E., 1997, *The Ethical Demand*. With an introduction by Hans Fink and Alasdair MacIntyre. Notre Dame and London: University of Notre Dame Press [translation of *Den etiske fordring*, Gyldendal, Copenhagen 1956.]

Luhmann, N., 2000, *Vertrauen, ein Mechanismus der Reduktion der sozialer Komplexität,* 4. Auflag, Stuttgart: Lucius & Lucius

Luhmann, N., 1988, 'Familiarity, confidence, trust: problems and alternatives', in: Gambetta, D. (ed.), *Trust: Making and breaking cooperative relations*, Oxford: Basil Blackwell, pp. 94-107

MacLagan, P., 1998, *Management and Morality. A developmental perspective,* London: Sage

Mansbridge, J.J., 1990, 'The rise and fall of self-interest in the explanation of political life', in: Mansbridge, J.J. (ed.), *Beyond self-interest*, Chicago: Chicago University Press, 3-22.

Marris, C., B. Wynne, P. Simmons and S. Weldon, 2001, *Public Perceptions of Agricultural Biotechnologies in Europe*, Final Report of the PABE research project

funded by the Commission of European Communities, Contract number: FAIR CT98-3844 (DG12 - SSMI), December 2001

McFall, L., 1987, 'Integrity', *Ethics*, 98/1, pp. 5-20

Meijboom, F.L.B., 2008, 'A proposal to broaden the analysis of problems of trust regarding food and biotechnology', in: Busch R.J., G. Prütz (eds.), *Biotechnologie in Gesellschaftlicher Deutung,* München: Herbert Utz Verlag, pp. 113-130

Meijboom, F.L.B., 2007, 'Trust, Food and Health, Questions of Trust at the Interface between Food and Health', *Journal of Agricultural and Environmental Ethics*, 20/3, pp. 231-245

Meijboom, F.L.B., T. Visak, F.W.A. Brom, 2006, 'From trust to trustworthiness: Why information is not enough in the food sector', *Journal of Agricultural and Environmental Ethics,* 19/5, pp. 427-442

Meijboom, F.L.B., M.F. Verweij, F.W.A. Brom, 2003. 'You eat what you are. Moral dimensions of diets tailored to one's genes', *Journal of Agricultural and Environmental Ethics*, 16/6, pp. 557-568

Meijboom, F., F.W.A. Brom, 2003, 'Intransigent or Reconcilable: The complex relation between public morals, the WTO and consumers', in: Vedder, A. (red.), *The WTO and Concerns Regarding Animals and Nature*, Nijmegen: Wolf Legal Productions, pp. 89-99

Mill. J.S., 1999, *On liberty*, Edited by Edward Alexander, Peterborough: Broadview Press Ltd

Miller, J., 2000, 'Trust: The Moral Importance of an Emotional Attitude', *Practical Philosophy*, 3/3 pp. 45-54

Miller, T.H., 1999, 'Thomas Hobbes and the Constraints that Enable the Imitation of God', *Inquiry*, 42/2, 149-176

Millstone, E., 2007, 'Can food safety policy-making be both scientifically and democratically legitimated? if so, how?' *Journal of Agricultural and Environmental Ethics*, 20, pp. 483-508

Millstone, E., 2006, 'Can food safety policy-making be both scientifically and democratically legitimate? If so, how?', in: Kaiser, M. & M.E. Lien (eds.), *Ethics and the politics of food*, Wageningen: Wageningen University Press, pp. 36-47

Millstone, E., Van Zwanenberg P., 2002, 'The evolution of food safety policy-making institutions in the UK, EU and Codex Alimentarius', *Social Policy and Administration*, 36/6, pp. 593-609

Millstone, E., Van Zwanenberg P., 2000, 'A crisis of trust: For science, scientists or for institutions?' *Nature Medicine*, 2000, 6/12, pp. 1307-1308

Misztal, B.A., 1996, *Trust in modern societies: the search for the base of social order*, Cambridge: Polity Press

Möllering, G., 2001, 'The nature of Trust: From Georg Simmel to a theory of expectation, interpretation and suspension', *Sociology*, 35/2, pp. 403-420

Muis, J., 2004, 'De redelijkheid van het theïsme', in: Th. Boer, *Schepper naast God?, Theologie, bio-ethiek en pluralisme essays aangeboden aan Egbert Schroten*, Kampen: Kok, pp. 23-38

Musschenga, B., 2004, *Integriteit. Over de eenheid en heelheid van de persoon*, Utrecht: Lemma

Nagel, T., 1991, *Equality and partiality*, Oxford: Oxford University Press

Newton, K., 2001, 'Trust, Social Capital, Civil Society, and Democracy', *International Political Science Review*, 22/2, 201–214

Newton-Smith, W.H., 2000a, 'Introduction', in: W.H. Newton-Smith (ed.), *A Companion to the Philosophy of Science*, Oxford: Blackwell Publishing, pp. 1-8

Newton-Smith, W.H. 2000b, 'Explanation', in: W.H. Newton-Smith (ed.), *A Companion to the Philosophy of Science*, Oxford: Blackwell Publishing, pp. 127-133

Niebuhr, R., 1934, *Moral man and immoral society: a study in ethics and politics*, New York: Scribner

Nooteboom, B., 2002, *Trust, forms, foundations, functions, failures and figures,* Cheltenham: Edward Elgar

Nowak, M.A., K. Sigmund, 1998, 'Evolution of indirect reciprocity by image scoring', *Nature*, 393, 573-577

Nussbaum (2001), *Upheavals of thought; The intelligence of emotions*, Cambridge: Cambridge University Press

OECD, 2000, 'Genetically Modified Foods, Widening the Debate on Health and Safety, The OECD Edinburgh Conference on the Scientific and Health Aspects of Genetically Modified Foods', http://www.oecd.org/dataoecd/34/30/2097312.pdf (accessed 5 July 2006)

O'Donovan, O, 1994, *Resurrection and Moral Order; An outline for Evangelical Ethics,* Second Edition, Grand Rapids: Eerdmans

Okrent, D., 'Comment on Societal Risk', *Science*, 208/ 4442, pp. 372-375

Okruhlik, K., 1998, 'Gender and the Biological Sciences, reprinted in: M. Curd & J.A. Cover (eds.), *Philosophy of Science: The Central Issues*, New York: W.W. Norton & Company, pp. 192-208

O'Neill, O., 2002a, *Autonomy and Trust in bioethics,* Cambridge: Cambridge University Press

O'Neill, O., 2002b, *A Question of Trust*, *BBC Reith lectures*, Cambridge: Cambridge University Press

Pagden, A. 1988, 'The Destruction of Trust and its Economic Consequences in the Case of Eighteenth-century Naples', in: Gambetta, D. (ed.), *Trust: Making and breaking cooperative relations*, Oxford: Basil Blackwell, 127-141

Paula, L., F. Birrer, 2006, 'Including public perspectives in industrial biotechnology and the biobased economy', *Journal of Agricultural and Environmental Ethics*, 19/3, 253–267

PDV (Product Board Animal Feed), 2002, *Evaluation of the MPA affair*, English Summary, http://databases.pdv.nl/documenten/Evaluatie_MPA_UK.pdf (accessed 28 March 2007)

Pellikaan, H., Van der Veen, R. J., 2002, *Environmental Dilemmas and Policy Design,* Cambridge: Cambridge University Press

Pence, G.E. (ed.), 2002, *The ethics of food, a reader for the 21st century*, Oxford: Rowman & Littlefield

Pettit, P., 2001, *A Theory of Freedom: From the Psychology to the Politics of Agency*, Cambridge: Polity

Pettit, P., 1995, 'The Cunning of trust', *Philosophy and Public Affairs*, 24/ 3, pp. 202-225

Phillips, Lord of Worth Matravers, J, Bridgeman, M. Ferguson-Smith, M., 2000, *The BSE Inquiry*, Volume 1, Findings and Conclusions, London: Stationery Office

Poortinga, W., N.F. Pidgeon, 2004, 'Trust, the asymmetry principle, and the role of prior beliefs, *Risk Analysis*, 24/6, pp. 1475-1486

Postow, B.C., 2007, 'Toward Honest Ethical Pluralism', *Philosophical Studies*, 132/2, pp. 191-210

Potter, N.N., 2002, *How can I be trusted? A virtue theory of trustworthiness,* Lanham etc: Rowman & Little Publishers

Pritchard, M.S., 1972, 'Human Dignity & Justice', *Ethics*, 82/4, pp. 299-313.

Putnam, R.D., 1995, 'Bowling Alone: America's Declining Social Capital', *Journal of Democracy*, 6/1, pp. 65-78

Putnam, R.D., 1992, *Making democracy work. Civic traditions in modern Italy*, Princeton: Princeton University Press

Rasmussen, B., K.K. Jensen, 2005, *The hidden values. Transparency in decision-making processes dealing with hazardous activities*, CeBRA Project Report 5, Copenhagen: Schultz

Rawls, J., 2002, *The Laws of People, with 'The idea of public reason revisited'*, Fourth Printing, Cambridge: Harvard University press

Rawls, J., 2000, *Lectures on the History of Moral Philosophy,* edited by Barbara Herman, Cambridge: Harvard University Press

Rawls, J., 1993, *Political Liberalism,* New York: Columbia University Press

Rawls, J., 1988, "The Priority of Right and Ideas of the Good", *Philosophy and Public Affairs*, 17/4, pp. 251-276

Rawls, J., 1980, 'Kantian constructivism in moral theory', *The Journal of Philosophy,* 77/9, 515-572

Rawls, J., 1972, *A Theory of Justice,* Oxford: Oxford University Press

Rawls, J., 1958, 'Justice as Fairness', *The Philosophical Review*, 67/2, pp. 164-194

RLG: Raad voor het Landelijk Gebied, 1998, *Zorg en vertrouwen: de basis voor voedselproductie in de 21e eeuw*, (*Care and Trust, the basis for food production in the twenty-first century*), Publicatie 98/7 (www.rlg.nl)

Rogers, A., 2002, 'EU promises reforms after another food scandal', *The Lancet*, 360/9336, p. 857

Rogers, W.A., 2002, 'Is there a moral duty for doctors to trust patients?', *Journal of Medical Ethics,* 28, pp. 77-80

Ross, D., 2006, *Game Theory*, In: Stanford Encyclopedia of Philosophy, www.plato.stanford.edu

Rotter, J. B., 1967, 'A new scale for the measurement of interpersonal trust', *Journal of Personality*, 35, 651-665

Ruscio, K.P., 1999, 'Jay's Pirouette, Or why political trust is not the same as personal trust', *Administration & Society*, 31/5, 639-657

Sagoff, M., 1988, *The economy of the earth*, Cambridge: Cambridge University Press

Sandel, M.J., 1984, 'The Procedural Republic and the Unencumbered Self', *Political Theory*, 12/1, pp. 81-96

Savage, L.J., 1954/1972, *The foundations of statistics*, New York: Dover Publications

Scanlon, T.M., 2002, *'*Reasons, Responsibility, and Reliance: Replies to Wallace, Dworkin, and Deigh', *Ethics,* 112 (April 2002), p. 507–528

Scanlon, T.M., 1998, *What we owe to each other*, Cambridge: Belknap Press of Harvard University Press

Scanlon, T.M., 1990, 'Promises and Practices,' *Philosophy and Public Affairs*, 19, pp. 199-226

Scheuplein, R.J, 1997, 'Regulating food-borne risks', in: Tennant, D.R. (ed.), *Food Chemical Risk Analysis*, London: Chapman & Hall, pp. 418-452

Scholderer, J., L.J. Frewer, 2003, 'The Biotechnology Communication Paradox: Experimental Evidence and the Need for a New Strategy', *Journal of Consumer Policy*, 26, pp. 125-157

Seligman, A.B., 1998, 'Between public and private', *Society,* 35/3, pp. 28-37

Seligman, A.B., 1997, *The problem of trust*, Princeton: Princeton University Press

Sevenhuijsen, S., 1999, 'Too good to be true? Feminist thoughts about trust and social cohesion', *Focaal. Tijdschrift voor antropologie* 34, 207-222

Shapiro, S.P., 1987, 'The social control of impersonal trust', *The American Journal of Sociology*, 93/3, 623-658

Siegrist, M., G. Cvetkovich, 2001, 'Better Negative than Positive? Evidence of a Bias for Negative Information about Possible Health Dangers', *Risk Analysis*, 21/ 1, 199 – 206

Simmel, G., 1950, *The Sociology of Georg Simmel, translated, edited and with an introduction by Kurt H. Wolff,* NY/London: Free press

Simon, H.A., 1959, Theories of decision-making in economics and behavioural science, *The American Economic Review*, 49/3, 253-283

Six, F., 2004, *Trust and trouble, building interpersonal trust within organizations,* dissertation, Rotterdam: ERIM

Skinner, Q., 2005, 'Hobbes on representation', *European Journal of Philosophy*, 13/2, 155–184

Slovic, P., 1999, 'Trust, emotion, sex, politics, and science: Surveying the risk-assessment battlefield', *Risk Analysis*, 19/4, 689-701

Slovic, P., 1996, 'The risk game', *Journal of Contemporary Water Research and Education*, 103, p. 2-7

Solomon, R.C., 1994, 'Sympathy and vengeance: the role of the emotions in justice', in: S.H.M van Goozen, N.E. van de Poll, J.A. Sergeant (eds.), *Emotions, Essays on emotion theory*, Hilsdale: Lawrence Erlbaum, pp. 291-311

Solomon, R.C., 1980, 'Emotion and Choice', in: A. O. Rorty (ed.): *Explaining Emotions*, Berkeley: University of California Press, pp. 251-281.

Strauss, L., 1961, *The political philosophy of Hobbes; Its basis and its Genesis,* Chicago: University of Chicago Press

Strawson, P.F., 1974, 'Freedom and Resentment', in: Strawson, P.F., *Freedom and Resentment and other essays,* London: Methuen & Co, pp. 1-25

Sztompka, P., 1999, *Trust: a sociological theory,* Cambridge: Cambridge University Press

Sztompka, P., 1998, 'Trust, distrust and two paradoxes of democracy', *European Journal of Social Theory,* 1/1, pp. 19-32

Taylor, A. E., 1939, 'The Right and the Good', *Mind*, 48/191, pp. 273-301

Tennant, D.R. (ed.), 1997, *Food Chemical Risk Analysis*, London: Chapman & Hall

Thompson, P.B., 2002, 'Food Safety, Ethics and the Universal Human Right to Food, Background paper for the FAO expert consultation on 'food safety: science and ethics', FAO, Rome 3-5 September 2002

Thompson, P.B., 1997, *Food Biotechnology in Ethical Perspective*, London: Blackie Academic & Professional

TNO, 2003, Rapport: 'Samen bouwen aan vertrouwen. Definitiestudie naar consumentenvertrouwen en voedselveiligheid, STB-03-09

Trust Barometer, 2008, *Edelman Trust Barometer Europe* – Netherlands Findings 2008, www.edelmanpr.nl/downloads/Edelman_Trust_Barometer_Europe_ Netherlands.pdf

Tucker, A. W., 1980 [1950], 'On Jargon: The Prisoner's Dilemma', *UMAP Journal,* 1, 101-103

Udehn, L., 2002, 'The changing face of methodological individualism', *Annual Review Sociology*, 28, pp. 479-507

Uslaner, E.M., 2002, *The moral foundations of trust*, Cambridge: Cambridge University Press

Vaarkamp, H., 2001, *Over de ratio van coöperatie: een speltheoretische benadering met een toepassing bij de mogelijke invloed van diergeneesmiddelen op voedselveiligheid,* Dissertation Wageningen University

Valasquez, M., 2003, 'Debunking moral responsibility', *Business Ethics Quarterly,* XIII/4

Van Larebeke N, L. Hens, P. Schepens, A. Covaci, J. Baeyens, K. Everaert, J.L. Bernheim, R. Vlietinck, G. De Poorter, 2001, 'The Belgian PCB and dioxin incident of January–June 1999: exposure data and potential impact on health.', *Environmental Health Perspective*, 109, pp. 265–273

Van Zwanenberg, P., E. Millstone, 2003, 'BSE: A Paradigm of Policy Failure', *The Political Quarterly*, 74/1, pp. 27–37.

Vedder, A. (red.), 2003, *The WTO and Concerns Regarding Animals and Nature*, Nijmegen: Wolf Legal Productions

Veldhuis, H., 2000, 'Ordained and Absolute Power in Scotus' Ordinatio I 44', *Vivarium*, 38/2, pp. 222-230

Verweij, M., 1998, 'Moral Principles: authorative Norms or Flexible Guidelines?', in: Burg, W. van der, Th. van Willigenburg (eds.), *Reflective Equilibrium, essays in honour of Robert Heeger*, Dordrecht: Kluwer Academic Press, pp. 29-40

VWA, 2003, Kernpunten *Jaarplan 2004*, Voedsel en Waren Autoriteit, Den Haag

VWA, (Siekman, F.B.W.), 2002, *Inventariserend onderzoek consumentenvertrouwen in voedselveiligheid,* Voedsel en Waren Autoriteit, Den Haag

Wallace, R.J., 2002, 'Scanlon's Contractualism'*, Ethics,* 112 (April), pp. 429–470

Waltner-Toews, D., T. Lang, 2000, 'A new conceptual base for food and agricultural policy: The emerging model of links between agriculture, food, health, environment and society', *Global Change & Human Health*, 1/2, pp. 116-130

Weil, F.D., 1987, 'The stranger, prudence, and trust in Hobbes's theory', *Theory and Society*, 15, pp. 759-788

Wempe, J., 2005, 'Ethical Entrepreneurship and Fair Trade', *Journal of Business Ethics*, 60, pp. 211-220

WHO (World Health Organization), 2002, *WHO global strategy for food safety: Safer food for better health*, Geneva: WHO

Williams, B., 1988, 'Formal Structures and Social Reality', in: Gambetta, D. (ed.), *Trust: Making and breaking cooperative relations*, Oxford: Basil Blackwell, pp. 3-13

Williams, G., 2000, 'Normatively demanding creatures: Hobbes, the fall and individual responsibility', *Res Publica*, 6/3, pp. 301-319

Williamson, O. E., 1993, 'Calculativeness, Trust, and Economic Organization', *Journal of Law and Economics*, 36, pp. 453-486

Willigenburg, T. van, 2000, 'Moral Compromises, Moral Integrity and the Indeterminacy of Value Rankings', *Ethical Theory and Moral Practice*, 3, pp. 385-404

Wong, D.B., 1992, 'Coping with Moral Conflict and Ambiguity', *Ethics*, 102/4, pp. 763-784

Wood, A.W., 1999, *Kant's Ethical Thought*, Cambridge: Cambridge University Press

WRR, Wetenschappelijke Raad voor het Regeringsbeleid, 2005, *Vertrouwen in de Buurt*, Amsterdam: Amsterdam University Press

WRR, Wetenschappelijke Raad voor het Regeringsbeleid, 2004, *Bewijzen van goede dienstverlening*, Amsterdam: Amsterdam University Press

WRR, Wetenschappelijke Raad voor het Regeringsbeleid, 2003a, *Beslissen over Biotechnologie,* Den Haag: SDU

WRR, Wetenschappelijke Raad voor het Regeringsbeleid, 2003b, *Waarden, Normen en de last van het gedrag*, Amsterdam: Amsterdam University Press

Wolf, S., 1992, 'Two levels of Pluralism'*, Ethics* 102 (July), pp. 785-798

Wubben, E., E. ten Pierick, J.W. van der Schans, K. Poppe, 2003, *De Nederlandse Voedsel en Warenautoriteit in het systeem voor borging van voedselveiligheid Een verkenning naar de toepasbaarheid van een institutioneel-economisch perspectief*, Rapport 7.03.05, Den Haag: LEI

Wynne, B., 2001, 'Public Lack of confidence in science? Have we understood its causes correctly?', in: Pasquali. M. (ed.): *EurSafe 2001*. *Food Safety, Food Quality, Food Ethics Preprints*, pp. 103-106

Wynne, B, 1992, 'Uncertainty and Environmental Learning: reconceiving science and policy in the preventive paradigm,' *Global Environmental Change* 2, pp. 1-22.

Zwart, H., 2003, 'Achterdocht en overvloed. De lotgevallen van de voedselconsument', In: Est, R. van, L. Hanssen & O. Crapels (eds.): *Genes for your food – Food for your genes: Social questions and dilemma's regarding food genomics*, Den Haag: Rathenau Instituut; Working document 90, pp. 93-124

SAMENVATTING

Inleiding

In het dagelijkse leven vertrouwen we op vele anderen: familie, vrienden en collega's, maar ook op experts die we niet persoonlijk kennen, op bedrijven en overheden. In veel gevallen zijn we ons niet bewust van dat vertrouwen. Daarom is het opmerkelijk dat vertrouwen volop in de belangstelling staat. Dat geldt vooral voor het vertrouwen van burgers in de overheid en van consumenten in bedrijven. Door verschillende ontwikkelingen in de samenleving is er een groeiende noodzaak om te vertrouwen. Tegelijk is er een toenemende onzekerheid over wie vertrouwd kan worden. Deze spanning definieer ik in dit proefschrift als het 'probleem van vertrouwen'. De stelling is dat dit probleem niet gezien moet worden als het dilemma van de individuele consument of burger, maar als een betrouwbaarheidsvraagstuk.

Door het 'probleem van vertrouwen' uit te werken aan de hand van voorbeelden uit de landbouw- en voedselsector wordt duidelijk dat het probleem niet enkel op het bord van het individu kan worden geschoven. De vraag is niet hoe individuen tot vertrouwen gebracht kunnen worden, maar wanneer iemand die vertrouwd wil worden, betrouwbaar wordt. Voor deze stap van vertrouwen naar betrouwbaarheid zijn diverse redenen. De morele redenen om het 'probleem van vertrouwen' te zien als een vraagstuk van betrouwbaarheid werk ik in dit proefschrift uit. Respect voor de autonomie van degene die vertrouwt staat hierbij centraal. Iemand die vertrouwt bevindt zich in een kwetsbare positie. Desondanks moet hij gezien blijven worden als iemand die een autonoom persoon is. Dit houdt in dat hij gerespecteerd moet worden als iemand die in staat is om moreel te handelen en als iemand die in moreel opzicht gelijk is aan degene die vertrouwd wordt. Dit respect heeft directe gevolgen voor wat betrouwbaarheid betekent. Wie betrouwbaar wil zijn dient niet alleen competent te zijn. Hij moet ook integer zijn, zodat hij met verschillen in (morele) verwachtingen kan omgaan, en hij dient een motivatie te hebben die start in het respect voor de autonomie van het individu ondanks diens kwetsbaarheid.

Deze centrale boodschap wordt in dit proefschrift uitgewerkt in vijf stappen:

1. Van vertrouwen naar betrouwbaarheid;
2. Van risico naar vertrouwen;
3. Van berekenbaarheid naar betrouwbaarheid;

4. Van competentie en motivatie naar betrouwbaarheid die start in respect voor autonomie;
5. Van arbitraire compromissen over wederzijdse verwachtingen naar betrouwbaar handelen.

1. Het 'probleem van vertrouwen' als betrouwbaarheidsvraagstuk

De spanning tussen de groeiende noodzaak om te vertrouwen en de toenemende onzekerheid over wie vertrouwd kan worden, leidt tot het 'probleem van vertrouwen'. Dit probleem wordt vaak gedefinieerd als een dilemma van het individu. De gedachte is dat de consument of burger onvoldoende inziet dat hij er goed aan zou doen om te vertrouwen. Vanuit deze gedachte lijkt het bevorderen van vertrouwen de meest zinvolle aanpak van het probleem. Mijn stelling is dat het 'probleem van vertrouwen' ten onrechte gezien wordt als het probleem van het individu en dat het een ongrijpbaar probleem blijft als het benaderd wordt als het gevolg van een foute inschatting van het individu. Voor een zinvolle benadering van het 'probleem van vertrouwen' moet de exclusieve focus op vertrouwen worden losgelaten en moet het probleem gezien worden als een vraagstuk van betrouwbaarheid. Voor deze stap van vertrouwen naar betrouwbaarheid zijn drie argumenten.

Ten eerste is er een conceptueel argument. Iemand kan de wens hebben om een ander te vertrouwen, maar het is niet mogelijk om daartoe te beslissen. Vertrouwen is geen standpunt waartoe iemand besluit om het in te nemen of af te wijzen. Dit kenmerk van vertrouwen heeft als gevolg dat vertrouwen niet af te dwingen is. Door gebruik van macht of kennis is het mogelijk dat iemand op je blijft rekenen, maar dat is nog geen vertrouwen. Het is echter wel mogelijk om het 'probleem van vertrouwen' te benaderen door je betrouwbaar te tonen. Dit leidt niet direct tot vertrouwen, maar het is wel een noodzakelijke voorwaarde.

Een tweede argument volgt uit de beoordeling van een gebrek aan vertrouwen. Als gebrek aan vertrouwen door een persoon of instantie als een probleem wordt gezien dan impliceert dat een claim over de eigen betrouwbaarheid. Tenzij die persoon of instantie ervan uit gaat dat anderen hem blind vertrouwen, zal hij er vanuit gaan dat het vertrouwen gebaseerd is op een (impliciete) beoordeling van zijn competentie en motivatie, met andere woorden van zijn betrouwbaarheid. Daarom omvat de beoordeling van het gebrek aan vertrouwen als een probleem ook een impliciete claim over iemands eigen betrouwbaarheid. De persoon of instantie beschouwt het gebrek aan vertrouwen als problematisch, omdat hij van mening is dat er goede redenen zijn om hem te vertrouwen. Met andere woorden dat hij betrouwbaar is. Vanuit dit perspectief is het te makkelijk om het 'probleem van vertrouwen' enkel te beschrijven als een fout van degene die niet vertrouwt. Degene die vertrouwd wil worden heeft ook een probleem.

224

Tot slot is er een moreel argument voor de stap van een exclusieve focus van vertrouwen naar een aanpak die start bij betrouwbaarheid. Dit argument is gebaseerd op de autonomie van degene die vertrouwt. In een vertrouwensrelatie is degene die vertrouwt (de 'truster') altijd afhankelijk van degene die vertrouwd wordt (de 'trustee'). Elke vertrouwensrelatie wordt gekenmerkt door een asymmetrie in kennis en macht tussen de beide partijen. Hierdoor bevindt de 'truster' zich in een kwetsbare positie. Die kwetsbaarheid is onvermijdelijk bij vertrouwen. Zonder die kwetsbare positie zou er voor de 'truster' geen reden zijn om te vertrouwen. Toch betekent dit niet dat deze kwetsbaarheid de morele status van die persoon aantast. Degene die vertrouwt dient altijd behandeld te worden als een persoon die in staat is om autonoom te handelen.

Dit argument toont dat de beschrijving van het 'probleem van vertrouwen' als enkel een fout van de 'truster' op twee manieren voorbij gaat aan de autonomie van de 'truster'. Allereerst neemt het de beoordeling van het individu niet voldoende serieus. Op basis van een houding van respect dient een gebrek aan vertrouwen in eerste instantie gezien te worden als een legitiem standpunt en niet als een foute beoordeling door de 'truster'. Natuurlijk kan een individu een verkeerde beoordeling maken van iemands betrouwbaarheid, maar vanuit een houding van respect voor autonomie volgt dat degene die vertrouwd wil worden het oordeel van de ander serieus dient te nemen. Ten tweede gaat het voorbij aan de verantwoordelijkheid die de 'trustee' heeft met betrekking tot de kwetsbaarheid van de 'truster'. De erkenning van de ander als een autonoom persoon betekent een verantwoordelijkheid voor de 'trustee' om zorg te dragen voor de bescherming van de 'truster' in zijn kwetsbare positie.

Op basis van deze argumenten stel ik dat de kern van het 'probleem van vertrouwen' een betrouwbaarheidsvraagstuk is.

2. De landbouw- en voedselsector als casus: Betrouwbaarheid centraal

De landbouw- en voedselsector staat in deze studie centraal als casus. Door de focus op deze sector wordt het 'probleem van vertrouwen' duidelijker en komt de noodzaak om dit probleem te beschouwen als een betrouwbaarheidsvraagstuk sterk naar voren.

Wie een supermarkt binnenloopt, merkt dat vertrouwen voor de voedselketen onmisbaar is. In principe schuilt achter elke aankoop de mogelijkheid van een risico voor de eigen gezondheid, voor aantasting van de persoonlijke waarden op het terrein van duurzaamheid of dierenwelzijn, enzovoort. Toch hindert dat mensen niet om te blijven consumeren. Mensen vertrouwen er bewust en onbewust op dat anderen binnen de voedselketen hun verantwoordelijkheid nemen en hebben genomen om eventuele risico's uit te sluiten of tot een acceptabel minimum te verlagen. Vertrouwen is dus essentieel voor de landbouw- en voedselsector. Toch speelt het 'probleem van vertrouwen' in deze sector een duidelijke rol. Drie kenmerken van de landbouw- en

voedselsector dragen bij aan de spanning tussen een toegenomen noodzaak om te vertrouwen enerzijds en een toegenomen onzekerheid over wie te vertrouwen is anderzijds. De eerste van deze drie kenmerken is de globalisering van de voedselmarkt. Nederlandse supermarkten worden gevuld met producten vanuit de hele wereld en de meeste voedingsmiddelen worden geproduceerd door bedrijven die internationaal opereren. De voedselmarkt is al lang niet meer te scheiden van de diverse overeenkomsten op het gebied van wereldhandel, zoals de WTO. Door deze ontwikkeling groeit het aantal actoren dat betrokken is bij de productie en verkoop van een product en neemt de afstand tussen consument en producent sterk toe. Bovendien heeft deze ontwikkeling tot gevolg dat mensen vaak niet of zeer gebrekkig weten hoe en waar hun voeding geproduceerd wordt. Dit onderstreept de noodzaak van vertrouwen. Voor een individuele consument is het onmogelijk om alle aspecten van een product te overzien en in te schatten. Wie voeding koopt en consumeert kan niet anders dan vertrouwen, of in ieder geval rekenen op vele anderen. Het mondiale karakter van de voedselmarkt toont niet alleen de noodzaak van vertrouwen. Het maakt vertrouwen ook moeilijker. Door de vele actoren die een rol spelen in de landbouw- en voedselsector en de grote afstand tussen degene die vertrouwt en degene die vertrouwd wordt, is een inschatting van de betrouwbaarheid van de 'trustee' vaak niet eenvoudig. Hierdoor is het weliswaar duidelijk dat men niet anders kan dan vertrouwen op anderen, maar is het niet duidelijk wie vertrouwd kan worden.

Ten tweede wordt de landbouw- en voedselsector gekenmerkt door een toenemend gebruik van technologie in de productie en verwerking van voeding. Het gebruik van technologie heeft een dubbel effect op vertrouwen. Aan de ene kant stellen technologieën ons in staat om persoonlijke controle te krijgen op processen waar we tot nu toe nog geen grip op hadden. Hierdoor verdwijnt in die gevallen de noodzaak om te vertrouwen. Aan de andere kant versterkt het gebruik van technologie de afhankelijke positie van individuele actoren. De gebieden waarop technologie wordt ingezet – denk aan voedselveiligheid, kwaliteit en gezondheid – maar ook de technologie zelf zijn vaak zo complex dat we aangewezen zijn op de kennis en kunde van een relatief kleine groep experts. Hierdoor wordt vertrouwen weer relevant. Tegelijk brengt de introductie van nieuwe technologieën een dynamiek met zich mee die op gespannen voet lijkt te staan met vertrouwen. Vertrouwen vraagt om een bepaalde stabiliteit en mate van voorspelbaarheid die bij de introductie van technologie vaak (nog) afwezig is. Hierdoor wordt de betrouwbaarheid van de trustee van essentieel belang.

Ten derde zijn er diverse voedsel(gerelateerde) incidenten geweest. Deze hebben veel aandacht gekregen in de media en hebben het denken en spreken over voeding sterk gekleurd. Vanuit het perspectief van het 'probleem van vertrouwen' zijn deze incidenten relevant, omdat ze enerzijds de afhankelijke positie van consumenten in dagelijkse handelingen zoals het kopen en consumeren van voedsel onderstrepen. Daarmee tonen ze de noodzaak van vertrouwen. Anderzijds blijkt uit een aantal van de incidenten dat

sommige partners in de voedselsector niet zo betrouwbaar blijken te zijn als dat van hen verwacht werd. Dit toont precies de spanning tussen een groeiende noodzaak om te vertrouwen en een toenemende onzekerheid over wie vertrouwd kan worden.

Een analyse van overheidsbeleid dat dit probleem primair bij de 'truster' zoekt, laat zien dat belangrijke beleidsmaatregelen zoals het verbeteren van risicoanalyses, het verhogen van transparantie in de keten en verbeteren van de informatievoorziening enkel effectief is voor het 'probleem van vertrouwen' als dit probleem wordt gezien als een betrouwbaarheidsvraagstuk.

3. Een analyse van het begrip 'vertrouwen'

Ondanks dat de centrale claim van deze studie is dat het 'probleem van vertrouwen' een betrouwbaarheidsvraagstuk is, is een analyse van het begrip vertrouwen noodzakelijk. Vertrouwen wordt in het dagelijkse leven op heel diverse manieren gebruikt. Vertrouwen in mijn moeder is anders dan vertrouwen in mijn auto. Het vertrouwen in mijn collega is weer anders dan het vertrouwen dat ik heb in de overheid. Het is allemaal vertrouwen, maar het is niet alle opzichten gelijk. Die verscheidenheid in betekenis en toepassing is kenmerkend voor vertrouwen. Toch is het van belang om het begrip vertrouwen te analyseren om zo scherp te krijgen wat betrouwbaarheid inhoudt.

Een eerste stap is om het onderscheid helder te krijgen tussen vertrouwen en gerelateerde begrippen zoals controle, geloof, hoop, en 'rekenen op iemand'. Vertrouwen bevindt zich tussen controle en hoop. Wie volledige controle over en kennis van een situatie heeft, hoeft niet te vertrouwen. Gebrek aan controle betekent echter niet dat vertrouwen noodzakelijk is. Ook al hebben we geen controle over het opgaan en ondergaan van de zon, zijn we prima in staat om te berekenen wanneer morgenochtend de zon opkomt. Dit is geen kwestie van vertrouwen, maar van een gefundeerde voorspelling. Dit komt, omdat we weliswaar geen controle hebben op de zonsopkomst, maar er ook relatief weinig onzekerheid is over het moment van zonsopkomst. Dit laat zien dat vertrouwen relevant is als we geconfronteerd worden met onzekerheid. Aan de andere kant van het spectrum staat hoop. Hoop en vertrouwen zijn vaak in dezelfde situaties van toepassing. Toch is er een fundamenteel verschil tussen beide. Hoop kan volledig los staan van de reële prestaties en competenties van degene op wie men zijn hoop vestigt. Vertrouwen is echter altijd gerelateerd aan hoe de ander heeft gehandeld in het verleden en hoe hij handelt in het heden. Er vindt bij vertrouwen altijd een (impliciete) inschatting plaats van degene op wie men vertrouwt. Dat is voor hoop niet noodzakelijk.

Binnen het spectrum van volledige controle over een situatie en een gevoel van hoop vinden we niet alleen vertrouwen, maar ook andere begrippen die vaak als synoniem van vertrouwen worden gebruikt. Geloven wordt bijvoorbeeld op die manier gebruikt.

'Geloof me maar' kan gezien worden als een synoniem voor 'vertrouw me maar.' Toch zijn vertrouwen en geloven geen synoniemen en zijn ze niet uitwisselbaar. Geloof omvat altijd een element van vertrouwen, maar vertrouwen omvat niet altijd een element van geloof. Wie vertrouwt, heeft de mogelijkheid om de betrouwbaarheid van de ander te onderzoeken en te controleren zonder dat dit het einde van de vertrouwensrelatie betekent. In het geval van geloven kan dit proces van onderzoeken en controleren een veel fundamentelere invloed hebben op de relatie. We spreken dan bijvoorbeeld van geloofstwijfel. Dit heeft te maken met de mate van overgave. Bij geloof is de mate van overgave veel groter dan bij vertrouwen. Geloof is daardoor omvattender. Je kunt geloof hebben in een persoon in alle omstandigheden of geloven in God door alle tijden heen. Bij vertrouwen kunnen we nog van oprecht vertrouwen spreken ook al is het vertrouwen in bepaalde personen, voor specifieke taken en in specifieke omstandigheden. Daarnaast worden vertrouwen en 'rekenen op' (*reliance*) als uitwisselbaar gebruikt. 'Rekenen op' is echter een bredere categorie dan vertrouwen. Het is een noodzakelijke, maar geen voldoende voorwaarde voor vertrouwen. Het is mogelijk om op iets of iemand te rekenen zonder te vertrouwen. Andersom is het niet mogelijk. Het verschil zit in de eisen die gesteld worden aan de andere partij. Bijvoorbeeld op een brug kun je rekenen, maar van echt vertrouwen is geen sprake. Zo'n brug kan niet anders dan functioneren zoals hij geconstrueerd is. Als hij niet functioneert, is dat geen beslissing van de brug. Om te kunnen spreken van vertrouwen moet de 'trustee' in staat zijn om te handelen, maar ook om af te zien van een handeling.

Hieruit volgen vier belangrijke kenmerken van vertrouwen. Ten eerste is vertrouwen relevant als een persoon geconfronteerd wordt met onzekerheid en gebrek aan controle. Vertrouwen stelt ons in staat om te handelen *alsof* de onzekerheid er niet is. Ten tweede is vertrouwen gebaseerd op een (impliciete) afweging. Het handelen *alsof* is niet een vlucht uit de werkelijkheid. Het is gebaseerd op een persoonlijke inschatting van degene die wordt vertrouwd. Ten derde is vertrouwen een houding die wij hebben ten opzichte van diegenen die in staat zijn om te handelen. Tenslotte is vertrouwen een houding die de 'truster' zelf kan kiezen en waarbij degene die vertrouwd wordt ook een zekere vrijheid heeft om wel of niet te reageren op het vertrouwen.

Op basis van deze eerste schets is het nog niet duidelijk wat iemand doet als hij een ander vertrouwt. Voordat ik dit verhelder en uitwerk tot een werkdefinitie van vertrouwen, distantieer ik mij van twee opties. De eerste stelt dat vertrouwen samenvalt met de beslissing om te handelen. Deze benadering heeft twee problemen. Enerzijds, kunnen handelingen ten onrechte worden toegeschreven aan vertrouwen. Naast vertrouwen zijn er ook andere mechanismen, zoals macht, die mensen in staat stellen om te handelen in situaties die gekenmerkt worden door onzekerheid. Anderzijds resulteert vertrouwen niet altijd in handelingen. Ik kan bijvoorbeeld vertrouwen hebben in mijn huisarts, maar zolang er voor mij geen noodzaak is om haar te bezoeken, resulteert mijn vertrouwen niet in een specifieke handeling.

Een tweede optie waar ik mij van distantieer, is de definitie van vertrouwen als een handeling en in het bijzonder als het nemen van een risico. Vertrouwen en het nemen van risico's zijn allebei relevant in situaties van onzekerheid, maar ze zijn niet identiek. De overwegingen die iemand heeft om een risico te nemen, zijn nog geen motivatie om een ander te vertrouwen. Maar ook andersom, de overwegingen om te vertrouwen zijn andere dan die om een risico te nemen. Vertrouwen en het nemen van risico's zijn elkaar aanvullende, maar fundamenteel verschillende mechanismen om met onzekerheid om te gaan. Een risicobenadering probeert de onzekere situatie zoveel mogelijk te vertalen in calculeerbare risico's. De vraag is dan of het het waard is om een bepaald risico te nemen, gegeven de preferenties en belangen die iemand heeft. In het geval van vertrouwen is het doel niet om de onzekerheid te vertalen in termen van calculeerbare risico's, maar om een manier te vinden om te handelen ondanks de onzekerheid. Wie vertrouwt, loopt wel een risico, maar neemt geen risico. Het gaat bij vertrouwen om een beoordeling van de situatie en niet om een calculatie van risico's. Hierdoor is het mogelijk dat, ook al is het risico relatief laag, iemand de ander niet vertrouwt, omdat hij de 'trustee' niet als betrouwbaar beschouwt. Voor vertrouwen is de (impliciete) beoordeling van de competentie en motivatie van de 'trustee' essentieel.

Door vertrouwen af te zetten tegen het nemen van risico's wordt duidelijk waarom informatie, bijvoorbeeld over de 'trustee', essentieel is voor vertrouwen, maar ook waarom we nooit tot een objectieve maatstaf kunnen komen voor de hoeveelheid informatie die nodig is om vertrouwen op te bouwen of te behouden. Vertrouwen is gebaseerd op een oprechte overtuiging met betrekking tot de betrouwbaarheid van degene die vertrouwd wordt, maar tegelijk omvat vertrouwen een duidelijke emotionele component. Door dit emotionele aspect van vertrouwen serieus te nemen, blijkt dat niet alleen onwetendheid of verkeerde calculaties leiden tot verschillen in vertrouwen. Emotionele oordelen sturen onze perceptie van de werkelijkheid. Wie een bedrijf vertrouwt, zal een terugroepactie zien als een teken dat de organisatie de zaken voor elkaar heeft en betrouwbaar opereert. Wie het bedrijf niet vertrouwt, zal in deze terugroepactie enkel een bevestiging zien dat de organisatie niet voldoende competent is. Dit maakt duidelijk dat betrouwbaarheid meer moet zijn dan het geven van informatie.

Op basis van deze analyse definieer ik vertrouwen als: "een attitude die een actor heeft ten opzichte van een andere (collectieve) actor, die de 'truster' in staat stelt om te gaan met situaties van onzekerheid en gebrek aan controle en die gebaseerd is op een beoordeling van de betrouwbaarheid van de ander welke resulteert in een positieve verwachting ten opzichte van de ander."

4. Vertrouwen en het belang van vrijheid en actorschap

Vertrouwen is een manier van omgaan met onzekerheid. Dit is niet uniek voor vertrouwen. Ook op andere manieren is dit mogelijk. Vertrouwen biedt echter de mogelijkheid om met een specifiek type van onzekerheid om te gaan, namelijk de onzekerheid die het gevolg is van de vrijheid die handelende personen hebben. Vertrouwen wordt essentieel op het moment dat het handelen van individuen niet volledig meer te voorspellen is met, of te herleiden is tot, de sociale en morele verbanden waarin zij leven en werken.

Om te kunnen omgaan met de vrijheid van individuen in situaties waarin samenwerking noodzakelijk is, zijn veel voorstellen gedaan. Een van die voorstellen is die van Hobbes. Hij laat zien dat het door middel van dwang en sancties mogelijk is om te komen tot samenwerking ondanks de vrijheid van alle betrokkenen. Voor vertrouwen is deze benadering problematisch. Het leidt weliswaar tot orde, waarin samenwerking mogelijk wordt, maar het leidt niet tot vertrouwen. Dwang en sancties staan vertrouwen in de weg. Het erkent de ander wel als individu, maar het neemt zijn vrijheid als actor niet voldoende serieus.

Rationele keuze theorie, die een belangrijke rol speelt in de literatuur over vertrouwen, neemt dit actorschap en de vrijheid van het individu wel serieus. Deze theorie gaat uit van individuen die doelgericht handelen en primair gestuurd worden door eigenbelang. Vanuit dit perspectief is vertrouwen rationeel op het moment dat vertrouwen het individu in staat stelt om te handelen volgens zijn eigen preferenties. Het kan in iemands belang zijn om zich kwetsbaar te maken en te vertrouwen om zo een bepaald doel te bereiken. Dit vraagt om een belangenafweging, waarin kansen en risico's centraal staan. We hebben echter hierboven gezien dat een risicoafweging nog niet tot vertrouwen leidt. Daarnaast wordt de afweging van de 'truster' bemoeilijkt door het feit dat de 'trustee' ook zijn eigen belangen heeft en daardoor ook zijn eigen redenen heeft om wel of niet te handelen conform het vertrouwen. Dit resulteert in de praktijk in complexe modellen, waarmee berekend kan worden of en wanneer vertrouwen rationeel is. Deze laten zien dat vertrouwen tussen individuen die doelgericht handelen en primair gestuurd worden door eigenbelang mogelijk is en rationeel kan zijn.

Het probleem van deze benadering is echter dat veel van wat wij vertrouwen noemen en als redelijk beschouwen niet goed binnen deze theorie past. Dit heeft zijn oorsprong in de exclusieve nadruk op eigenbelang als drijvende kracht achter het handelen van het individu. Het valt niet te ontkennen dat eigenbelang een belangrijke motivatie is; toch kan menselijk handelen niet gereduceerd worden tot dat wat gedreven wordt door eigenbelang. Bij vertrouwen verwachten we in sommige gevallen niet alleen *dat* iemand iets doet. We verwachten het ook *van* hem. In die gevallen gaat het om een verwachting die ook geldt als het niet in het eigenbelang is van de 'trustee'. Daarnaast gaat het bij vertrouwen niet enkel om rationaliteit in de zin van doelgericht handelen. Het gaat ook om redelijkheid. Bij redelijkheid gaat het om de wederzijdse bereidheid om elkaars

230

belangen en overtuigingen als startpunt te nemen voor een eerlijke samenwerking. Voor betrouwbaarheid betekent dit dat de 'truster' niet enkel gezien kan worden als een instrument om de doelen van de 'trustee' te bereiken, maar als een volwaardig persoon. Concreet betekent dit dat degene die vertrouwt en degene die vertrouwd wordt elkaar erkennen als autonome personen. Autonomie is hier iets anders dan onafhankelijkheid. Als autonomie en onafhankelijkheid zouden samenvallen dan zouden vertrouwen en autonomie elkaar uitsluiten. In een vertrouwensrelatie is immers altijd één partij in een afhankelijke positie en daarmee is hij niet meer volledig onafhankelijk van anderen. Autonomie kan echter ook in een bredere zin worden gebruikt. Autonomie is dan gebaseerd op de mogelijkheid van een persoon om in vrijheid te handelen volgens een zelfgekozen plan. Deze visie op autonomie staat vertrouwen niet in de weg. Wie vertrouwt, maakt zichzelf afhankelijk van anderen, maar wordt daardoor niet minder autonoom. Deze claim heeft directe gevolgen voor het operationaliseren van betrouwbaarheid.

5. Respect voor autonomie als noodzakelijke voorwaarde voor betrouwbaarheid

Autonomie speelt bij vertrouwen en betrouwbaarheid een belangrijke rol. Iemand is autonoom als hij de capaciteit heeft om vrij te handelen volgens een zelfgekozen plan. Deze capaciteit om in vrijheid te kunnen handelen, is echter nog geen voldoende reden om een persoon te beschermen. De kern van autonomie ligt niet alleen in iemands vrijheid, maar ook in de capaciteit om zijn eigen wil te bepalen. Dat laatste stelt een persoon in staat om letterlijk 'auto-nomos' te zijn, dat is, zichzelf de wet te stellen. Wie autonoom handelt, handelt niet alleen in vrijheid, maar ook volgens morele verplichtingen die de persoon zichzelf oplegt en die los staan van mogelijke sancties of verplichtingen van buitenaf. Dit maakt een individu tot iemand die moreel kan handelen. Dit betekent niet dat autonome personen enkel doen wat goed is in hun eigen ogen. Autonomie houdt in dat bij de bepaling van iemands morele verplichtingen altijd in ogenschouw genomen wordt of deze verplichting ook door elke andere autonome persoon als morele verplichting geaccepteerd zou kunnen worden. Die ander is immers ook een autonoom persoon en dat geeft hem een gelijkwaardige morele status. Daarom stelt Kant als één van zijn drie centrale claims dat de ander altijd als doel in zichzelf dient te worden behandeld en nooit enkel als een instrument. Op basis van deze imperatief volgt dat, als we geconfronteerd worden met een autonoom persoon, respect voor deze persoon als een doel in zichzelf de enige juiste reactie is. Hij is immers moreel gezien onze gelijke.

Deze twee aspecten van autonomie, (a) het in staat zijn tot moreel handelen en (b) de morele gelijkwaardigheid van autonome personen, zijn van groot belang bij het

'probleem van vertrouwen'. Ondanks de kwetsbare positie die iemand die vertrouwt heeft, blijft hij een persoon die in staat is om moreel te handelen en iemand die moreel gezien gelijk is aan de 'trustee'. Vanuit dit perspectief kan het 'probleem van vertrouwen' gezien worden als een signaal dat de autonomie van één of beide partijen in de vertrouwensrelatie niet serieus genomen wordt.

Deze centrale rol van autonomie resulteert in een morele verplichting voor wie betrouwbaar wil zijn: Wie betrouwbaar wil zijn heeft een morele verplichting om respect te tonen voor zijn eigen autonomie en voor die van de 'truster'. Het verzaken van die verplichting heeft verregaande gevolgen. Ofwel het betekent dat degene die vertrouwt niet als morele actor en als een moreel gelijke wordt beschouwd, hetgeen die persoon niet snel zal motiveren om te ander te vertrouwen. Ofwel het is een signaal dat de 'trustee' zijn eigen autonomie niet serieus neemt en dat roept al snel vragen op over zijn competentie en motivatie om betrouwbaar te handelen. Daarom dient iedere poging om betrouwbaar te zijn te starten in het respect voor de autonomie van degene die vertrouwt. Hieruit volgen twee grenzen voor betrouwbaarheid. Het houdt in dat elke vorm van misbruik onverenigbaar is met betrouwbaarheid. Daarnaast is ook manipulatie onverenigbaar met betrouwbaarheid. Bij manipulatie is het verschil tussen betrouwbaar zijn en vertrouwd worden essentieel. Voor wie enkel vertrouwd wil worden, kan manipulatie effectief zijn. Voor betrouwbaarheid is het uiterst problematisch, omdat het een signaal is dat de 'trustee' de autonomie van degene die vertrouwt niet zo serieus neemt dat het een voldoende argument is om hem niet te manipuleren.

De verplichting om respect te tonen voor de autonomie van de 'truster' roept echter ook een probleem op. Wie betrouwbaar wil zijn kan zijn eigen autonomie niet ontkennen. Dat zorgt ervoor dat er een probleem ontstaat als de redelijke verwachtingen tussen 'truster' en 'trustee' conflicteren. De morele plicht om de 'truster' als morele actor serieus te nemen, betekent dat de 'trustee' de morele opvatting van de ander niet zomaar terzijde kan schuiven. Tegelijk moet hij ook zijn eigen autonomie serieus nemen en kan hij niet eenvoudigweg met elke morele opvatting meegaan zonder zijn betrouwbaarheid te verliezen. Dit probleem wordt versterkt doordat het aantal en de indringendheid van deze conflicten toeneemt omdat de samenleving wordt gekenmerkt door een moreel pluralisme. Er is een diversiteit aan morele opvattingen, die ook met voldoende informatie en een goede dialoog niet volledig verdwijnt.

Om in die situatie van moreel pluralisme betrouwbaar te zijn, dient een 'trustee' een balans te vinden tussen accommodatie en integriteit. Accommodatie of aanpassing houdt in dat men open is naar morele opvattingen, die conflicteren met de eigen opvattingen en actief opzoek gaat naar oplossingen voor het conflict die de overtuiging van de ander respecteert. Dit streven naar accommodatie is een direct gevolg van het respect voor de ander als autonoom persoon. De morele opvattingen van de 'truster' tellen immers net zo goed als die van de 'trustee'. Concreet resulteert dit in twee principes van accommodatie. Ten eerste, wie betrouwbaar wil zijn, dient zo te handelen dat hij

andermans morele opvattingen en positie als morele actor zo min mogelijk schaadt. De kwetsbare positie, waarin een 'truster' zich bevindt, legt een extra verantwoordelijkheid bij degene die vertrouwd wordt om zorg te dragen dat morele overtuigingen van de 'truster' niet geschaad worden. Ten tweede, wie betrouwbaar wil zijn, dient zo te handelen dat de relatie tussen beiden zo min mogelijk wordt geschaad. De relatie tussen beide actoren kan immers van essentieel belang zijn in het bereiken van moreel gedeelde doelen. Deze twee principes sturen het handelen van de 'trustee' op een manier die de ander respecteert als morele actor en als moreel gelijke.

Tegelijk is integriteit van groot belang voor wie betrouwbaar wil zijn. Aanpassen aan de opvattingen van een ander kan een teken zijn van respect, maar ook van een gebrek aan principes en morele ruggengraat. Wie integriteit mist, blijft onbetrouwbaar. Iemand kan voorspelbaar zijn en daardoor is het mogelijk om op hem te rekenen, maar dat is nog geen vertrouwen. Om betrouwbaar te zijn is het van belang dat er bij de 'trustee' een zekere samenhang is in zijn handelen en dat een identiteit zijn gedrag stuurt. Dan wordt het mogelijk om een goede inschatting te maken van iemands competentie en motivatie, met andere woorden, van iemands betrouwbaarheid.

In conflictsituaties betekent betrouwbaar zijn altijd het zoeken naar een balans tussen aanpassing als teken van respect voor de ander en integriteit als teken van een eigen morele positie. Het vinden van zo'n balans is vaak niet eenvoudig en vraagt om compromissen. Een analyse van het sluiten van compromissen laat zien dat zo'n balans mogelijk is en dat het sluiten van compromissen bij morele conflicten betrouwbaarheid niet uitsluit.

6. Betrouwbaarheid in een institutionele context

Vertrouwen in de landbouw- en voedselsector is voornamelijk vertrouwen in bedrijven, overheden en andere instituties. In de moderne samenleving is ons vertrouwen gericht, niet op individuen, maar op een collectief van actoren die functioneren in organisaties.

Deze verschuiving van vertrouwen in personen naar vertrouwen in instituties en de personen die daarin functioneren, heeft gevolgen voor wat betrouwbaarheid inhoudt. Een analyse van dit institutioneel vertrouwen laat zien dat het nog steeds gaat om echt vertrouwen dat gericht is op het omgaan met de onzekerheid die het volg is van de vrijheid van mensen, die in staat zijn te handelen. Dit roept de vraag op of een institutie ook betrouwbaar kan zijn of dat ze enkel berekenbaar kan zijn. Mijn stelling is dat een institutie betrouwbaar kan zijn. Dit vraagt van een instelling dat ze enerzijds zelf competent en adequaat gemotiveerd is om te handelen op de toevertrouwde wijze. Anderzijds dat de organisatie betrouwbaar gedrag van de individuen, die functioneren binnen de organisatie, bevordert.

Hierboven heb ik voor betrouwbaarheid een extra voorwaarde geformuleerd: de morele plicht van respect voor autonomie. Deze plicht roept op institutioneel niveau een vraag op. Bij personen heeft deze morele verplichting zijn oorsprong in het feit dat men als autonoom persoon de ander ook erkent als een autonoom persoon. Een instantie is echter geen autonoom persoon in de zin dat het in staat is zelfstandig en in vrijheid te handelen volgens een zelfgekozen plan. Daarom is een ander argument nodig als basis en verdediging voor het respect voor en de bescherming van de autonomie van de 'truster'. Hierbij speelt de impliciete uitnodiging van instanties om op hen te vertrouwen een belangrijke rol. Het feit dat we banken vertrouwen in financiële kwesties of de Voedsel en Waren Autoriteit bij vragen van voedselveiligheid is geen toeval. Deze instanties profileren zich op een bepaalde wijze en creëren daarmee specifieke verwachtingen. Zelfs als ze zich niet direct profileren als experts op een bepaald terrein dan creëren de missie, de interne regels en principes van een organisatie vaak al specifieke verwachtingen. Op basis hiervan is het mogelijk om te spreken van een impliciete uitnodiging tot vertrouwen. Er is sprake van een impliciete belofte over hun betrouwbaarheid. Dit element van belofte vormt de basis voor een nieuwe onderbouwing van de verplichting tot respect voor en de bescherming van de autonomie van de 'truster'. Een belofte komt immers met verplichtingen. In het sterkste geval gaat het om een verplichting om die belofte te houden. Maar ook al voordat er een expliciete belofte gedaan wordt, kunnen er principes worden geformuleerd die de positie van de ander beschermen en respecteren. Deze principes voorkomen dat iemand wordt gemanipuleerd of onevenredig veel schade oploopt als de ander zich niet aan zijn belofte houdt. Op basis van deze principes kan een instantie zo handelen dat het respect toont voor en bescherming biedt aan de autonomie van de 'truster'. Daarmee kan een organisatie betrouwbaar zijn.

Op institutioneel niveau komt de noodzaak om te zoeken naar een balans tussen accommodatie en integriteit terug. De diversiteit aan morele opvattingen is bij institutioneel vertrouwen nog groter. Een instantie wordt vaak door velen vertrouwd. Dit betekent dat als een instantie betrouwbaar wil zijn, het moet kunnen omgaan met alle opvattingen die leven onder degenen die hen vertrouwen. In dit geval wordt nog duidelijker dat accommodatie niet kan betekenen dat de 'trustee' altijd volledig meebuigt met de opvattingen van de ander. Zich aanpassen aan één persoon kan immers een direct disrespect betekenen voor de opvattingen van een andere persoon, die ook op die organisatie vertrouwt. Daarom is integriteit essentieel voor betrouwbaarheid op institutioneel niveau. Het gaat hierbij niet alleen dat een bedrijf of instelling consistent is, maar ook dat het zich gecommitteerd heeft aan bepaalde doelen en taken die essentieel zijn voor die organisatie. Deze vorm van integriteit is van belang voor betrouwbaarheid. Ten eerste stelt het degene die vertrouwt in staat om te verhelderen wat verwacht mag worden van de 'trustee'. Daarnaast toont het de 'trustee' grenzen aan de vereiste om te accommoderen. Deze grenzen maken tegelijk duidelijk welke ruimte

overblijft voor accommodatie. Voor wie integriteit heeft, zijn deze grenzen niet arbitrair. Ze kunnen beargumenteerd worden vanuit het respect voor de 'truster' en op basis van de betrokkenheid van de 'trustee' bij zijn eigen kerndoelen en taken.

7. Een kader om betrouwbaarheid in praktijk te beoordelen

Op basis van de bovenstaande discussie is het mogelijk om terug te keren naar de vijf stappen van deze studie. Met behulp van een casus is het mogelijk om deze stappen verder te concretiseren. Ik gebruik hiervoor de introductie van nieuwe voedingsmiddelen met gezondheidsclaims. Op dit moment worden er veel producten ontwikkeld op het grensvlak tussen voeding en gezondheid. Hierbij gaat het bijvoorbeeld om melk die een bloeddrukverlagende werking heeft of margarine waarmee het mogelijk is om het cholesterolgehalte actief te verlagen. Deze producten bieden veel nieuwe mogelijkheden, maar roepen ook vragen op, waaronder vragen op het gebied van vertrouwen.

Deze casus laat allereerst duidelijk zien dat eventuele probleem op het gebied van vertrouwen niet enkel kunnen worden bestempeld als een fout van de eindgebruiker. Door dit te doen worden drie aspecten onvoldoende onderkend. Ten eerste neemt het de kwetsbare positie van de consument onvoldoende serieus. De meeste individuen zijn niet in staat om het proces achter de ontwikkeling, introductie en verkoop van deze nieuwe producten volledig te begrijpen, laat staan zelf te controleren. Dit heeft als gevolg dat men vaak niets anders kan dan vertrouwen op anderen. Ten tweede ontkent de nadruk op vertrouwen de impact van diverse problemen en zelfs schandalen in de voedingssector en gezondheidssector. Ten derde ontkent een eenzijdige nadruk op de 'truster' de impliciete claim van betrouwbaarheid, die volgt uit de beslissing om een nieuw product te introduceren of toe te laten op de markt.

Ten tweede toont deze casus het verschil tussen het nemen van risico's en vertrouwen. Nieuwe producten roepen altijd vragen op met betrekking tot veiligheid. Dit geldt zeker als er aan producten gezondheidsclaims worden gekoppeld. Daarom is het van groot belang om eindgebruikers in staat te stellen om zelf in te schatten of ze het product wel of niet willen gebruiken en of ze bepaalde risico's acceptabel vinden. Toch is hiermee het vraagstuk van vertrouwen niet verdwenen. Allereerst vragen transparantie en communicatie over risico's al een bepaalde mate van vertrouwen. Alleen als iemand de bron van informatie betrouwbaar acht, wordt de informatie bruikbaar en relevant. Daarnaast heb ik laten zien dat wie in staat is om risico's in te schatten en te beoordelen niet noodzakelijkerwijs in staat is om te vertrouwen. Vertrouwen verschilt fundamenteel van het nemen van risico's.

Ten derde volgt uit deze casus dat betrouwbaarheid meer is dan berekenbaarheid. Hoewel berekenbaarheid of voorspelbaarheid voldoende kan zijn om te vertrouwen op anderen laat deze casus zien dat betrouwbaarheid meer vraagt dan voorspelbaar zijn.

Zowel in de voedselsector als in de gezondheidssector zijn duidelijke patronen op basis waarvan wij verwachtingen formuleren. Het probleem is echter dat deze patronen niet gelijk zijn voor beide sectoren. Dat roept vragen op wanneer een voedingsmiddel een claim heeft die tot nu toe alleen bekend is uit de gezondheidssector. Hierdoor wordt onduidelijk wat wij van wie mogen verwachten. Onze verwachtingen ten opzichte van de eigenaar van een supermarkt zijn anders dan die ten opzichte van een apotheker, ook al verkopen zij beiden dezelfde producten aan eindgebruikers. Dat betekent dat berekenbaarheid in dit geval tekort schiet om betrouwbaar te zijn.

Een volgende stap is de nadruk op respect voor autonomie. De erkenning van de 'truster' als iemand die in staat is om autonoom te handelen is in deze casus om twee redenen van essentieel belang. Allereerst is het respect voor de ander als morele actor belangrijk, omdat de introductie van deze nieuwe voedingsproducten om een morele discussie vraagt. Waarden zoals veiligheid, keuzevrijheid en rechtvaardigheid worden zowel in de voedingssector als in de gezondheidssector gerespecteerd. De concrete invulling van die waarden verschilt echter per sector. De eisen die aan de veiligheid van een voedselproduct worden gesteld zijn volledig anders dan bij een geneesmiddel. Deze verschillen hebben niet alleen te maken met gewoonte, maar ook met een verschil in opvatting over wat een acceptabel risico is bij een product. Dit betekent dat de introductie van een product op het grensvlak van voeding en gezondheid vraagt om een morele discussie over de invulling van wat we onder veiligheid en andere relevante waarden verstaan. Daarnaast is respect voor autonomie belangrijk, omdat de consument behandeld moet worden als een moreel gelijke. Concreet houdt dit in dat degene die vertrouwd wil worden geen gebruik kan maken van de kwetsbaarheid van de 'truster' door de bovenstaande waarden in hun meest minimale vorm in te vullen of zelfs door eenzijdige informatie te geven.

Tot slot toont de casus het belang van de laatste stap: van arbitraire compromissen over wederzijdse verwachtingen naar betrouwbaar handelen. De introductie van de genoemde nieuwe producten vraagt om een morele discussie over de invulling van waarden zoals veiligheid en rechtvaardigheid. In die discussie is integriteit cruciaal voor wie betrouwbaar wil zijn. In praktijk zal het vaak onmogelijk zijn om tegemoet te komen aan alle verwachtingen. Om toch in staat te zijn om te bepalen wanneer accommodatie mogelijk is, is integriteit van belang. Als een organisatie staat voor een bepaald doel en voor specifieke principes is het mogelijk om op basis daarvan te bepalen en te beargumenteren welke reactie wenselijk is. Deze balans tussen accommodatie en integriteit betekent niet dat iedereen de organisatie zal vertrouwen, maar stelt de organisatie wel in staat om betrouwbaar te handelen.

Voor betrouwbaarheid in de praktijk onderstreept dit de noodzaak van een goede communicatie door de 'trustee' met de 'truster'. Betrouwbaarheid vraagt om communicatie en transparantie over:a) de competentie van de organisatie, maar ook over de redenen die ten grondslag liggen aan de keuze om bepaalde competenties te

ontwikkelen;b) de motivatie die ten grondslag ligt aan de reactie op het vertrouwen van de 'truster'. De transparantie dient duidelijk te maken hoe de motivatie verbonden is met het respect voor de 'truster' als autonoom persoon en met de eigen betrokkenheid bij de kern van de organisatie en de taken en principes die uit die kern volgen. Dit stelt de 'truster' in staat om in te schatten wie te vertrouwen is a) in die situaties die onzeker blijven ook nadat de onzekere factoren zoveel mogelijk vertaald zijn in termen van risico's, b) wanneer voorspelbaarheid tekortschiet, c) wanneer er sprake is van conflicterende morele waarden en d) wanneer vertrouwensvragen starten in de overtuiging dat men niet serieus genomen wordt als autonoom persoon. In die gevallen kan communicatie over de competentie en motivatie van de 'trustee' het verschil maken. Deze communicatie kan cruciale informatie geven voor de beoordeling van de betrouwbaarheid van de ander. Bovendien creëert deze communicatie verwachtingen over toekomstig gedrag van degene die vertrouwd wordt, hetgeen verplichtingen schept voor de 'trustee' om de 'truster' te beschermen.

In praktijk zal dit niet alle vertrouwensproblemen oplossen. Het geeft echter helderheid over de betrouwbaarheid van de ander. Dit is een noodzakelijke voorwaarde om vertrouwen op te bouwen en in stand te houden.

CURRICULUM VITAE

Franck L.B. Meijboom is geboren op 5 november 1975 te Rotterdam. In 1994 behaalde hij zijn Gymnasium diploma aan het Marnix College te Ede. Van 1994 tot 1999 studeerde hij theologie aan de Universiteiten van Utrecht en Aberdeen (UK). Eind 1999 studeerde hij af aan de Faculteit Godgeleerdheid van de Universiteit Utrecht, met als specialisatie wijsgerige ethiek en bio-ethiek. Daarnaast heeft hij de kerkelijke opleiding van de Nederlandse Hervormde Kerk, thans de Protestantse Kerk Nederland, in 2001 succesvol afgerond.

Sinds 1999 is hij verbonden aan het Centrum voor Bio-ethiek en Gezondheidsrecht, dat in 2003 is opgegaan in het Ethiek Instituut van de Universiteit Utrecht. In eerste instantie als *executive officer*. In deze functie is hij betrokken geweest bij de opbouw van de European Society for Agricultural and Food Ethics (EurSafe). Vanaf 2001 is hij als junior docent-onderzoeker werkzaam bij het Ethiek Instituut. In diverse onderzoeksprojecten heeft hij zich gericht op het terrein van de landbouw-, voedsel- en dierethiek. Deze projecten hebben als inspiratie en achtergrond gediend voor het proefschrift over vertrouwen en betrouwbaarheid in de landbouw en voedselsector, waar hij in 2004 aan begon. Naast zijn positie als onderzoeker en docent is hij voorzitter van de dierexperimentencommissie van Wageningen Universiteit en lid van de commissie van TNO. Bij de European Society for Agricultural and Food Ethics heeft hij de functie van bestuurssecretaris.

Franck L.B. Meijboom (1975, Rotterdam) studied theology and ethics at the universities of Utrecht and Aberdeen. Since 1999, he has been affiliated to the Ethics Institute of Utrecht University as researcher and lecturer. His fields of interest are agricultural ethics, food ethics, and animal ethics. Additional to his activities as researcher and teacher, he is chair and member of two committees for the ethical review of animal experimentation and secretary of the European Society for Agricultural and Food Ethics (EurSafe).