

*Detlef Fetchenhauer, Wolfgang Jagodzinski,  
Kazufumi Manabe, Axel Ockenfels, Akira Okada,  
Gisela Trommsdorff, and Toshio Yamagishi*

## **Different Perspectives on Trust**

Results from a German Japanese Symposium at the  
University of Cologne<sup>1</sup>

### **1. Introduction**

Trust is a key word in current public debates. Bankers have become one of the most distrusted professional groups in the West. Politicians in most democracies have experienced for decades the decline of political trust. Japanese no longer trust the Tokyo Electric Power Company Tepco and their government after the nuclear catastrophe in Fukushima. Germans have lost confidence in the European currency. All of these developments sound dramatic because people seem to believe that human interaction in networks and systems cannot function without a minimum level of trust. When citizens do not trust their government, they will deliberately not pay taxes or render voluntary services to the community. When people do not trust each other, they will spend enormous resources to protecting themselves against fraud, theft, and violence. As a further consequence, the state's monopoly on the use of force is endangered. Trust is a social capital and it reduces transaction costs in social and economic exchange.

As trust seems to be important in many life domains, it is not surprising that it has become a focus of interdisciplinary research. In political science, David Easton (1965) stimulated a broad research on trust, confidence, and political support nearly 50 years ago. Psychologists have investigated the personal and situational factors that influence the trustworthiness of persons. Social scientists have studied trust as a multi-level phenomenon. On the micro level, trust is a ferment of social exchange; on the macro level, it is often seen

---

<sup>1</sup> The Symposium was jointly organized by Wolfgang Jagodzinski, University of Cologne, and Kazufumi Manabe, Aoyama Gakuin University, and supported by the Japan Foundation.

as a social capital of the respective society. Cultures seem to produce different forms and levels of trust. Generalized interpersonal trust, for instance, seems to be more widespread in democracies with market economies. Economists have investigated the role of trust in economic exchange; the application of game theory has allowed them to specify the causes and consequences of trust in mathematical models and to subject these models to experimental tests.

It seems pretty obvious that different studies are based on slightly different concepts of trust. Experiments with anonymous players focus on trust in cooperation. Non-experimental research, by contrast, mostly investigates various forms of generalized trust in persons or institutions. In the next section we will distinguish trust concepts according three criteria: the objects of trust, the levels of generalization, and nature of the trust relation. Do these different research traditions have more in common than the ambiguous name *trust*? Can they ultimately be integrated into a common theoretical framework like the one suggested by Coleman (1990)? Because it is too early to answer the latter question definitively, first we will compare the different research traditions to help detect differences and communalities among the existing frameworks, and to identify fields for further research. In this paper we will present different perspectives on trust. Section 2 briefly introduces the different concepts of trust. Section 3 presents game-theoretical experiments that seem to best fit Coleman's general concept.

## **2. Types of Trust Relationships<sup>2</sup>**

Although all of these disciplines speak of trust and confidence, it remains an open question whether they all study the same object. We suggest a distinction among the different forms or types of trust according to three criteria: object of trust, specificity, and trust component.

### **2.1 Objects of trust**

In the statement, "P has trust in O," P<sup>3</sup> is the truster and O is the "trusted object or subject"<sup>4</sup>. P can have trust in (1) occurrences of events, (2) persons, (3) actions and outcomes, (4) institutions, or (5) systems or subsystems.

---

<sup>2</sup> The following section is a summary of Jagodzinski (2012).

<sup>3</sup> Throughout this paper P symbolizes the focal person, the truster, or the subject in an experiment. The only exception is subsection 3.3 where experimental subject are asked about the expected behavior of P and O.

1. *Trust in occurrences of events*: We exclude the first kind of trust from further consideration. In colloquial language one might say, "I trust that we'll have jolly weather tomorrow," but this use of the word *trust* is not of interest here.
2. *Trust in other persons*: In the social sciences trust is mostly seen as an ingredient of an interpersonal relationship. Even if we say, "I trust that O will repay the loan," we usually consider it as a trust in the debtor O and not as trust in the action (repayment).
3. *Trust in actions and outcomes*: People often know little about their exchange partner. The less we know, the more the trust shifts from the person to the action or outcome. In modern markets the counterpart often remains completely anonymous. We do not even know whether we interact with a person or a computer. In such a case, we trust that the anonymous counterpart will behave or act in a certain way but we do not trust a person<sup>5</sup>.
4. *Trust in institutions*: Trust in institutions is derived from the trust in the persons who represent or work for the institutions and from the performance of the institution. We trust in the police as long as they effectively prevent crime and protect us, and we trust in the courts as long as the judges' sentences are in accordance with the law and meet our own standard of fairness and justice. Whether institutional trust is more persistent than interpersonal trust is an open question, but the former is at least to some extent buffered against latter. A single corrupt police officer does not need to destroy the overall trust in the police, but permanent scandals often have this effect.
5. *Trust in subsystems and systems*: What has been said about institutions in principle also applies to systems. Take a democratic system as an example, which includes several institutions like the government, the press, or the unions. People may be permanently dissatisfied with some of these institutions but nevertheless keep their trust in the democratic system.

In this paper we will primarily focus on trust in the sense of (2), (3), and (4).

## 2.2 Specificity of trust

In each of the three cases we can distinguish between more or less general forms of trust. As far as interpersonal trust is concerned, P can trust a particu-

---

<sup>4</sup> We will usually speak of the object of trust for the sake of brevity even if the object is a person. What kind of object is meant can be inferred from the respective context.

<sup>5</sup> At least in Western languages trust can refer to actions and outcomes. One can even contrast interpersonal trust and trust in general by saying: "I do not trust him in general but I have high trust that he will do what is expected from him in the given situation."

lar person (e.g., a spouse), groups of persons of varying size (family, neighbors, citizens, Christians, etc.), or human beings in general.

Interpersonal trust can also be generalized along other dimensions. P can trust her best friend O in general, or in a given context or situation. The former is usually based on the inference of O's general motives, values, and personal traits, whereas the latter takes also the specific motives and the incentive structure surrounding the relationship into account. According to Yamagishi and Yamagishi (1994), the incentive structure has no effect on trust: it generates assurance but not interpersonal trust. This distinction will be discussed in Section 4 but will not play a role in other parts of the paper. In any case, trust can be based on general or specific orientations of O.

A specification of trust relationships is often reached by fixing the location of an object in time and space. This is not the only means, however. We can also define smaller and broader sets of actions: P can trust that his best friend O will (a) help him to carry the heavy dishwasher he has just bought, (b) always help him carry heavy goods, or (c) in general be helpful. Let the first action (carrying the dishwasher) be the only element of set A, and let B and C denote the sets of actions mentioned in (b) and (c). Consequently,  $C \supset B \supset A$  more general than B, and B is more general than A. As long as we generalize along a single dimension, we can easily rank the objects of trust and accordingly distinguish among more or less specific forms of interpersonal trust. As soon as generalization is performed along several dimensions, however, the differentiation among levels of generalizations often becomes ambiguous.

The last example shows how people can have trust in more or less specific actions. Institutional trust can also exist in more or less specific forms. We can have trust in the police station in our neighborhood, the police in our community, or in the police of our state or our country.

### **2.3 Affective, evaluative, cognitive, and behavioral components of trust**

So far trust has been diffusely characterized as a relationship. But what kind of relationship is it? Is it exclusively or at least predominantly cognitive, evaluative, affective, or behavioral? Different disciplines and different schools offer different answers.

#### **1. Trust as behavior**

Many actions in everyday life are immediately understood as acts or expressions of trust. Person P expresses trust, for instance, if she pays in advance, or if she lends another person O a certain amount of money, not knowing whether O will ever return it or not. In Friedrich Schiller's poem "The Hos-

tage," Damon's friend displays the utmost level of trust by accepting the agreement that he will be killed instead of Damon if the latter should not return within three days. If trust is equated to this kind of behavior, we may speak of a behavioral understanding on trust. Accordingly, trust becomes manifest in action and in nothing else. Whether trust is accompanied by feelings, cognitions, or other internal states does not matter. What matters are behavioral manifestations or expressions of trust. The conditions under which a behavior is qualified as an expression of trust cannot be precisely defined but some necessary conditions can be given. The behavior must be displayed in the interaction with other persons and it must be evident from the action that the actor places or increases the risk of his or her own material or immaterial losses. Verbal statements alone, e.g., the responses in an interview, rarely involve such risks and are therefore at best indirect indicators of trust behavior.

Behavioral trust in this sense can be more or less broadly defined. Economists, for instance, may consider moneylending always as a manifestation of trust. The more guarantees  $P$  has that  $O$  will repay, the higher  $P$ 's trust in the repayment or, more general, *in the cooperation of  $O$* . This is, in fact, the implicit or explicit understanding of trust in most trust games (see 3.2). Yamagishi and his colleagues, by contrast, hold a narrower concept of behavioral trust. According to their view, personal characteristics of the other person  $O$  are the only basis or source of trust. Trust is interpersonal. External rewards and sanctions give assurance but do not affect interpersonal trust. Accordingly, no trust is involved in moneylending if  $P$  is effectively protected against the risk of a loss by a binding contract, bailment, insurances, etc. The more security that is reached for  $P$  by these arrangements, the less moneylending is an expression of trust. While the increase of security implies an increase of trust in the former concept (trust in cooperation), it reduces the influence of trust on behavior in the latter concept. This will be discussed further in Section 4.

## 2. Trust as a belief or the result of cognitive operations

From a rational choice perspective, people are governed by expected gains and losses. A rational actor  $P$  will only hand over the money to the borrower  $O$ , for example, if  $P$  is sufficiently sure that  $O$  will return the money and pay the interest on the loan. The expected value<sup>6</sup> of loaning has to be larger than the expected value of using the money for other purposes or keeping it. In this framework, trust could be understood as  $P$ 's belief in the trustworthi-

---

<sup>6</sup> It would be more appropriate to speak of the expected utility. Throughout this paper, however, we will assume that utilities and monetary outcomes are equivalent.

ness of O (interpersonal trust), or as P's belief that the loan agreement will be kept (trust in action).

There are other conceptualizations, however, that underlie not only many micro-economic experiments but also the theory of Coleman. In these approaches, trust is decoupled from the cognitive process either by arguing that the cognitive operations *precede* the decision of loan granting or by assuming that people actually do not make these calculations but act *as if* they were calculating expected gains and losses. Accordingly, trust is the decision to place trust in O's cooperation. If this decision becomes only manifest in P's cooperative behavior, i.e., loan granting, we remain within the behavioristic perspective.

### 3. Trust as affect or emotion

Most micro-economic experiments do not attempt to measure emotions or evaluations. They usually apply to situations in which the affective component of trust may be weak or does not vary much from person to person. This seems plausible as long as P's counterpart O is anonymous or a computer. It is less plausible in long-term personal relationships or short-term face-to-face contacts. From a theoretical point of view, emotions could be easily incorporated into the rational choice framework. One can admit immaterial in addition to material rewards. It is more difficult to find scale-equivalent measures for emotions. In many social situations emotions are more important than material rewards. It often hurts more for O to lose the trust of a close friend than even a large amount of money. P will take this into account when placing trust. Fairness and helpfulness in general is intrinsically rewarding independent of their effects on reputation.

### 4. Trust as evaluation

If P tells us that he or she has full trust in O, we often consider this statement as a summative evaluation like "O is a reliable person," "O is fair," "O is not egoistic," etc. The question of whether O shares our norms and values is therefore often the decision-maker for the emergence of trust. Value standards may change with objects and situations. Trust in the police, for instance, is presumably much more dependent on security values and security needs than trust in one's spouse.

## 2.4 Manifold forms of trust in a multidimensional space

Trust probably includes all four components, though in varying degrees. While rewards, sanctions, and restrictions may have stronger effects on trust in specific actions of anonymous partners, emotions and values may be more important in face-to-face contacts and generalized forms of trust. Even if all

four components of trust and their determinants could be integrated in a single theoretical model, the strength of the relationships and the effects would differ from type to type of trust. In this paper we do not try to create a unified, theoretical model but present empirical findings for different forms of trust.

Section 3 focuses on trust in specific actions, that is, the typical form of game-theoretical experiments. While the influence of motives and personal traits, emotions, and evaluations is reduced as much as possible in these experimental settings, they play an important role in the two forms of generalized trust. Generalized interpersonal trust (Section 4) is strongly based on attachment and feelings, and generalized institutional trust (Section 5) is governed by norms and values.

### **3. Trust in Game-Theoretical Experiments**

In this section, trust is understood as the result of predominantly cognitive operations. The object of these cognitive operations is the action  $A(O)$  of an anonymous partner  $O$ , usually called cooperation. The object is very specific, like the payment of a certain amount of money. The alternative to cooperation is defection. Trust becomes manifest in the cooperation of the focal person  $P$ , the subject of the experiments. Accordingly, both the expression of trust and the object of trust are actions.

Some of the experiments described here have the form of the well-known prisoner's dilemma. For  $P$  and  $O$ , defection is the dominant strategy but both would be better off if they cooperated. The provision with collective goods, as we know from the seminal book of Mancur Olson (1965), creates such a prisoner's dilemma. As nobody can be excluded from the consumption, rational egoists will not participate in the production of a collective good but instead defect or free-ride. This hampers the formation of institutions for the production of collective goods. Even if a group cooperates and builds up the institution, free-riding may result in a suboptimal outcome. Olson has already emphasized that the collective good problem can be more easily overcome in small groups because coordination costs are low and the benefits for a single person often exceed the total production costs. Olson calls this phenomenon the exploitation of the great by the small.

Section 3.1 describes experiments on the institution formation of small groups. The process is modeled as an  $n$ -person 3-stage game that is mathematically examined and empirically tested. Like in many other experiments and applications of the prisoner's dilemma, trust is not explicitly entered as a variable. Trust presupposes risk and there is no risk if all players behave as rational egoists in a simple prisoner's dilemma game. They all will defect. The question,

therefore, is whether the rules of the game and the incentives can be changed in such a way that some or all rational egoists will cooperate. In this way, P's trust in the cooperation of others will be increased. The multi-step, multi-player game in Section 3.1 is designed exactly for that purpose: It increases the probability that rational egoists will participate in the institution formation.

The role of trust is more explicitly investigated in trust or assurance games, some of which resemble Internet shopping. The focal person P, the buyer, initially obtains a certain amount of money that he or she can either keep or invest. If P keeps the money the game is over. If P transfers the money to the anonymous seller O, the latter can either cooperate or defect. Defection means that O keeps P's money and P loses the investment. Cooperation means that P receives more money in return than he or she has invested. The crucial question is the trustworthiness of O. If O should cooperate, P and O would benefit. If not, P will have a loss. There are many variations of the game: It can be played once or several times, P can receive smaller or larger amounts of money, and the modalities of investment and return and the number of players can be modified.

Coleman's (1990) conceptualization of trust applies to these experiments quite well. A rational actor P will calculate the potential gains G and losses L of cooperation, and defection estimates the probability  $p^*$  of O's cooperation. P will place trust in the cooperation if the expected gain is larger than the expected loss:  $p^*G > (1-p^*)L$ <sup>7</sup>. Assume, for example, that buyer P and seller O both receive 20€ at the start of the game. P can either buy by paying O the 20€ in advance or keep the money. If O cooperates, they both will benefit and receive 30€ each; if O defects and keeps the money, O will have 40€ and P nothing. In this case, P will consider two alternatives: cooperation (buying) and defection (keeping the money). P will buy if the expected value of cooperation is larger than 20€—the money P can keep when not buying. Accordingly,  $p^*30$  €, the expected value of cooperation, must surmount 20€. This is the case if  $p^* > \frac{2}{3}$ .

As L and G are fixed in advance, the crucial problem for P is the determination of  $p^*$  or the ratio  $p^*/(1-p^*)$ . Often called the trustworthiness of O, the notion is somewhat misleading in situations in which  $p^*$  is largely determined by situational factors and less by the traits and motives of O (or, more precisely, their perception by P). Thus, if the term trustworthiness is used at all, we should speak of the trustworthiness of O in a given situation with regard

---

<sup>7</sup> Coleman uses an equivalent expression:  $\frac{p^*}{1-p^*} > \frac{L}{G}$ .

to this specific cooperation. Actually,  $p^*$  is a subjective probability<sup>8</sup>—subjective because it is the probability that P believes in.

Many micro-economic experiments do not attempt to measure  $p^*$ . They examine factors that theoretically should influence  $p^*$  and test empirically whether they have the predicted effect on P's cooperation and trust. A change in the rewards and sanctions for O, for instance, should indirectly influence P's behavior. O's willingness to cooperate might also be increased by playing the game several times so that O's defection can be punished by P's<sup>9</sup> defection in the following games. Another determinant of  $p^*$  is O's reliability, which people derive from their own observations or from the information of others. A theoretically and empirically important source of trust for assessing the reliability of O is his or her former behavior. Internet platforms have set up information systems that provide exactly this kind of information about buyers and sellers. In Subsection 3.2 we offer an overview of experiments that investigate the influence of this information on trust. We also report the difficulties of implementing feedback information systems in practice.

If trust games like the previously discussed ones are played only once and participants remain anonymous, rational players should defect. However, quite a few people do not behave in this way. The percentage of cooperation is far above zero. Why do people cooperate in such situations? Are they driven by altruism or moral norms? Are they anxious to lose their reputation when behaving egoistically? How do subjects as external observers assess the game? What are their estimates of the probabilities of cooperation  $p^*$  for O and P? And does P anticipate the behavior of O correctly? Subsection 3.3 addresses some of these questions.

### 3.1 Cooperation and institution formation<sup>10</sup>

Kosfeld, Okada, and Riedl (2009) consider a game theoretic model to investigate interrelationships among trust, cooperation, and institution formation in a social dilemma situation, and report experimental observations to test theoretical predictions. The aim of this research is to scrutinize whether institutional arrangements are successful in promoting cooperation, and how trust affects the sustainability of cooperation.

---

<sup>8</sup> The probability  $p^*$  must also not be confused with Coleman's concept of trust. Even if the trustworthiness in a given situation is very high, P will not place trust in O's cooperative behavior if the losses relative to the gains are very large or, more precisely, if the inequality condition (see note 7) is not met. Vice versa, a low  $p^*$  does not exclude trust if the potential gains relative to the losses are sufficiently high.

<sup>9</sup> The most famous strategy of this type is Tit-for-Tat (see Axelrod 1984, 1997).

<sup>10</sup> This subsection is based on the presentation by Akira Okada.

In a global society, an important question to ask is how to cooperate and resolve various conflicts. This question becomes more and more important theoretically and practically as we observe many political and economic problems such as global warming and monetary crisis. It is certainly true that generalized interpersonal trust is a crucial element of cooperation. Generalized interpersonal trust, however, is not enough for cooperation; it may be fragile in an anarchic state of nature. Cooperation should be supported by strategic incentives. In other words, cooperation should be strategically stable. Some appropriate institutional arrangements are needed to change our economic incentives. Institutional arrangements should be voluntary, without a central government to enforce cooperation in our global society.

There is a well-known puzzle in the institutional approach. It is often said that, because rational individuals with self-interest have an incentive to free-ride on an institution-enhancing cooperation, they are likely to fail in forming institutions. Individuals need to cooperate at a higher level of institutional arrangements. Thus, the institutional approach is faced with the same question that it is expected to solve. How can individuals cooperate in institutional arrangements?

To examine the possibility of institutional arrangements by rational individuals, Kosfeld, Okada, and Riedl (2009) present a multi-stage game model of institution formation in a public goods economy. At the beginning of the game, each player receives an endowment that can either be contributed to the production of public goods or kept. The player's benefit from public goods is a fixed share of the total contribution, i.e., the sum of contributions of all players. Thus, the more people invest their endowments into public goods, the higher the benefits for each player. Free-riders keep their endowment. Because the game has the same payoff structure as the  $n$ -person prisoner's dilemma, the individual rational action is to contribute nothing. This non-cooperative outcome is not socially optimal. In order to attain the socially optimal outcome, players may need to have a certain institution that forces participants to contribute. A challenging question is whether such an institution can be created voluntarily by rational players with self-interests.

Institution formation is a complex process, and parties are typically involved in multi-stage bargaining. To capture a sequential structure of institution formation, our game has the following three stages. In the first stage, players decide anonymously and without knowing the decision of others whether they want to participate in an institution or not. They do not make any contribution at this moment.

The first decision divides the players into participants and non-participants (acting as free-riders in the last stage). Only the former decide in the second stage whether an institution is formed. Institution formation requires

a unanimous vote of all participants. If this condition is met, the participants are forced to contribute their full endowment to public goods. These rules are public information to all individuals. They give participants a perfect assurance that other participants will also contribute. Participants vote for or against institution formation, considering whether their benefit from the institution will outweigh the non-cooperative outcome without it. In case the number of participants is too small, they will vote against the institution formation.

In the third stage, all players, participants, and non-participants actually make their contribution decisions. When an institution is created, non-participants free-ride and benefit the most. Participants, however, will also receive more in return from the institution than in the non-cooperative outcome. In everyday life existing institutions often ask for donations and some of them are quite successful in mobilizing financial support. Such contributions are also possible in the institution formation game. Rational egoists, however, will not donate<sup>11</sup>.

Kosfeld, Okada, and Riedl (2009) analyze a non-cooperative equilibrium of the institution formation game. A key factor in the result is the size (number of participants) of an institution that is agreeable to participants. Every participant's payoff should be greater than in the case of no contribution. The minimum integer satisfying this condition is called the minimum size of an institution.

Kosfeld, Okada, and Riedl (2009) show that, for every integer between the minimum size of an institution and population  $n$ , there exists a sub-game perfect equilibrium of the game in which an institution is created by  $s$  participants. On general, there exist free-riders who do not contribute. This result has two important implications. First, institutional arrangements are possible, but may not be socially optimal. The largest, socially optimal institution is not necessarily formed. Secondly, there exist a large number of equilibriums.

Kosfeld, Okada, and Riedl (2009) have tested the theoretical predictions of the institution formation game. The experiments were conducted at the University of Amsterdam in 2003, and 164 undergraduate students participated. Subjects played a four-person game of institution formation in 20 rounds with the same partners. Two different values of the marginal per capita return from contributions to public goods (MPCR) were selected, 0.4 and 0.65. These values correspond to the minimum size of an institution, 2 and 3, respectively.

Experimental observations are summarized as follows. First, there are almost always some participants in an institution. When the minimum size of

---

<sup>11</sup> They may exceptionally donate under the condition that their gain in reputation surmounts their donation. Reputation, however, is difficult to measure and has not been taken into account in the game.

an institution is 2 (MPCR=0.4), the percent of rounds with a successful institution is 43%. When the minimum size of institution is 3 (MPCR=0.65), the same percent is 60%. These data show that subjects actually form costly institutions about half the number of rounds.

Secondly, the socially optimal four-person institution forms in most cases. Participants are likely to reject institutions with free-riders. The probabilistic belief of subjects elicited during the experiments show that this result of institution formation is not due to coordination failure among subjects, and that subjects actually anticipate the largest institution.

Thirdly, institutions with fewer than four participants are rejected with a high likelihood. There is at least one free-rider in such an institution. Participants regard such an institution as unfair and reject it. Subjects' behavior is affected by the fairness consideration.

Fourthly, the frequency of the largest institution increases as rounds proceed. Subjects learn to form an institution. Finally, it was examined how institutional arrangements affect the level of contributions. Without institutional arrangements, the average contribution decreases as rounds proceed. In contrast to this observation, institution arrangements enhance cooperation, particularly when the marginal per capita return from public goods is low.

The analysis of Kosfeld, Okada, and Riedl (2009) shows theoretically and experimentally how people voluntarily form institutions. Institutions actually enhance cooperation. In the experiments, subjects have a tendency to form the socially optimal institution. Trust plays an important role in the success of institutional arrangements, and fairness is a critical factor to stabilize trust.

### **3.2 Engineering (dis-)trust<sup>12</sup>**

Trust is a critical ingredient in virtually all social and economic relationships. From an economic perspective, the need for trust arises from the fact that contracting on every move we make is often prohibitively costly or prohibitively costly to enforce. The anonymity of geographically dispersed online traders, for instance, makes contracting difficult on the Internet: You may not be able to identify, say, your eBay seller or verify the quality of the object being sold, let alone get your money back at reasonable costs (Bolton and Ockenfels 2009).

Economic theory provides useful intuition for the role of transaction costs, asymmetric information, incomplete markets, insufficient coordination, and other frictions in explaining trust gaps, and sometimes provides guidance on how to address trust problems. For instance, one economic foundation of trust relationships is the reciprocity principle of tit-for-tat combined with reputa-

---

<sup>12</sup> This subsection is based on the presentation by Axel Ockenfels.

tion systems that store information on past performance (Greif 1993). Direct reciprocity applies to repeated relationships: "I will trust you tomorrow if you are trustworthy with me today." Indirect reciprocal systems enforce trust when the relationship is one-shot by a circuitous tit-for-tat: "I will trust you tomorrow if you are trustworthy with a third party today." Both designs work equally well, in theory, if traders are fully rational and the reputation information is complete and reliable (Wilson 1985).

However, in practice, sometimes important behavioral phenomena are not captured by theory. A system that works well in theory may not work under more realistic assumptions. In fact, one stream of our earlier research on trust investigates the performance and limits of "perfect" feedback systems in the laboratory, providing an upper bound on what can be expected from field systems in which conditions are less than perfect. We observed that feedback systems are less effective in promoting trust in indirect trading networks compared with direct trading networks, even under conditions in which, in theory, they should perform the same (Bolton, Katok, and Ockenfels 2004, 2005; Bolton and Ockenfels 2009; Bolton, Ockenfels, and Ebeling 2011, see also the references therein). In addition, traders whose trust is betrayed tend to lose trust, both in the market in general as well as in the specific person who was not trustworthy (see Buskens and Raub 2010 for a related research program in sociology). These kinds of findings suggest that behavioral complexities often cannot be ignored when we investigate the determinants of trust. Luckily, while trust is not always rational, it is also not chaotic or irrational, but rather follows systematic patterns of bounded rationality (see also other subsections in this article).

Trust interacts with institutions. The previous subsection, for example, illustrated how, in principle, institutions can be formed to promote trust and cooperation. However, real-world institutions are typically far more complex than what can be captured by simple, elegant theory. What is largely missing is a study that is devoted to the economic engineering aspects of economic design. How can we deal with practical challenges when we move from theory to workable institutions for real markets? Take eBay's feedback forum, which can be interpreted as a kind of "institutionalized gossip." The feedback provided by eBay traders is publicly available and easy to access so that each buyer can look at a seller's feedback history before he or she engages in bidding. In this way, the incentives for moral hazard are weakened by the feedback system: If eBay traders punish sellers with negative feedback information by refusing to buy from them or reducing the price they are willing to pay, then the threat of leaving negative feedback should discipline the seller. However, there are various challenges in the details of the system (Bolton and Ockenfels 2012; Ockenfels and Resnick, 2012, and the references therein). One challenge is that feed-

back information must come from voluntary self-reporting of one's own experiences with trading partners. Yet, feedback is a public good; the cost for providing feedback is paid by the provider but the benefit only goes to *other* traders, and no trader can be excluded from using the information. In addition, a variety of incentives exist to manipulate feedback, like giving good feedback to friends and bad feedback to competitors. Another challenge is that negative feedback is often retaliated by negative feedback, creating incentives not to give negative feedback. As a consequence, the value of information feedback, if given at all, is likely to be less than perfect.

Indeed, a study by Bolton, Greiner and Ockenfels (2012) finds that frequency and content of feedback giving are highly reciprocal in nature, and that this has both potentially positive effects on the system (more feedback giving) as well as negative effects (discouraging negative reports). The gaming in the production of reputation information significantly reduces the informativeness of feedback information, hampering the ability of a reputation system to facilitate trust and trade efficiency. It is demonstrated, however, that reciprocity can be guided by changing the way feedback information flows through the market system, leading to more accurate reputation information, more trust, and more efficient trade. The simple two-sided feedback system, traditionally implemented by eBay, can be improved, for example, by introducing "some blindness" in feedback giving, which increases the informativeness of the feedback presented to buyers because it reduces the scope for reciprocity in the production of feedbacks. This result is behaviorally robust in the sense that it is based on complementary evidence from very different laboratory and naturally occurring field sources—although the interaction with the specific conflict resolution mechanisms at work on eBay and other platforms seem less well-understood.

The engineering study by Bolton, Greiner and Ockenfels (2012) agrees with the theoretical literature that reciprocity is a key motivation to promote trust—but it emphasizes that reciprocity in the production of reputation information seriously hampers trust, and that a partly "blind" flow of information is recommended. Related findings on the effect of information flows on trust also originate in applied and engineering work. For example, the empirical antitrust literature as well as engineering work in spectrum auction have shown that (dis-)trust and reciprocity (when contracting is not possible) can be strongly affected by the flow of communication and other institutional and behavioral complexities (e.g., Klemperer 2004). Combined with the pioneering work by Ostrom (1990) and Roth (2002), this is a convincing argument that research in trust and social behavior can substantially benefit from the accumulation of an engineering-oriented design literature, and its feedback on fundamental science.

### 3.3 Do people trust too much or too little?<sup>13</sup>

In the binary trust game that has been used in many studies (for an overview see Dunning, Fetchenhauer, and Schlösser 2012), Persons A are endowed with a certain amount of money (e.g., €5) that they can either keep for themselves or hand over to an unknown and anonymous Persons B. If they decide for the latter, Persons B get the fourfold of that amount (e.g., €20). Persons B then have two alternatives on what they can do with this money. They can either keep the whole amount for themselves or they can share the money evenly between themselves and Persons A.

In such a paradigm, B will get with certainty 10€ when sharing the money and 20€ when defecting. The rational actor, therefore, has no choice and will keep the money. Nevertheless, about 70 to 90% of Persons B decide to reciprocate the trust that is given to them (i.e., they share the money evenly). This result was stable across many experimental conditions and in many different samples both in the United States and in Europe. Consequently, the money-maximizing strategy for Persons A was to take the risky option and send their money to Persons B.

However, it turned out that in all of these different samples, participants heavily underestimated the trustworthiness of their fellow participants. On average, they thought that only about half of their potential interaction partners would honor their trust. Hidden behind this average estimation of others' trustworthiness was a huge variance. While some participants thought that literally nobody could be trusted, others could not imagine a person being so selfish as to not replicate their trust.

Given the participants' estimations of others' trustworthiness and taking into account their general level of risk aversion (as measured by their behavior in a lottery), most participants should not hand over their money to Persons B. Interestingly, this was not observed. Irrespective of their skepticism, a majority of participants choose the risky option in a trust game even if they are pretty sure they will not see the money again. When correlating participants' estimations of others' trustworthiness with their actual trust decisions, we usually find a positive and significant correlation, but this relationship tends to be rather modest (about 0.30) (Fetchenhauer and Dunning 2009).

Apparently, people place higher trust in cooperation than they expect from others. It can be argued that these results might be the result of the fact that participants do not reveal what they truly think. An economist might argue that participants did not have any material incentive to come up with a valid estimate. A psychologist might argue that people just boost their moral self-

---

<sup>13</sup> This subsection is based on the presentation by Detlef Fetchenhauer.

esteem by pretending that they themselves are trustworthy, but most others are not.

To deal with these arguments, in one study participants had to estimate the trustworthiness of 56 persons whom they saw in short video-clips. Half of the participants were told that they could earn €0.50 for each correct estimate; the other half were not given any incentive. As it turned out, both groups did not differ in their cynicism about the stimulus-persons' trustworthiness (Fetchenhauer & Dunning, 2010).

In another study, participants had to take the role of Person A in a trust game and were informed about their chances of running into a trustworthy Person B (Fetchenhauer and Dunning 2012). This probability was either 80 or 46%. In addition, participants had to indicate their decisions in a lottery that had an identical payoff structure (i.e., either an 80 or a 46% chance of doubling one's inset). The results of the lottery were very much in keeping with what we know from research about judgment and decision making. In the 80% condition a vast majority (about 75%) of all participants decided on the riskier option, but in the 46% only a small minority (about 30%) did so. This makes sense, as even for a risk-neutral participant, the expected value in the latter condition suggested it would be the safer option.

Yet, in the trust game a very different pattern of results could be observed. In both conditions a majority of the participants decided to hand over their money to Person B. Although this was done slightly more often in the 80% condition than in the 46% condition, this difference was not significant.

In summary, trust decisions do not logically follow people's estimations of others' trustworthiness. To put it differently, by measuring trust as behavior, one does not also measure trust as cognition. Moreover, our research has shown that trust as behavior is heavily influenced by emotions that people attach to their trust decisions. Interestingly, it is not the emotions that people anticipate to experience when contemplating a certain outcome of their decision (e.g., "How would I feel if I hand over my money and Person B runs away with it?"). Rather, participants unconsciously base their decisions on the way they feel in the very situation they are in to make that decision (e.g., "How do I feel when giving my money to Person B?").

Therefore, it would be beneficial for future trust researchers to put more emphasis on distinguishing among trust as behavior, trust as cognition, and trust as emotion. Each of these three dimensions is necessary to understand the complex phenomenon that is called trust (Dunning, Fetchenhauer, and Schlösserl 2012).

#### 4. Generalized Interpersonal Trust<sup>14</sup>

Some may object to how Coleman's (1990) conceptualization defines trust too broadly by equating it to behavior under risk. Many forms of risk behavior have little to do with trust if we follow common language. Can we really speak of trust if we decide for a risky action with small probability  $p^*$  because the ratio of loss  $L$  to gain  $G$  is even smaller than  $p^*/(1-p^*)$ ? Does the roulette gambler trust the croupier to give the ball exactly the spin that is needed for the chosen number or combination? Maybe the player trusts that the casino uses unbiased roulette wheels, but that is not the focus when placing the bet.

Intuitively, trust is mostly understood as a relationship between persons or groups. The bailment of a third person  $Z$  for the debts of  $O$ , for instance, is not seen as a trust-increasing measure. Instead, it indicates distrust rather than trust when  $P$  asks for such securities before lending money to  $O$ . The bailment of  $Z$  does not strengthen  $P$ 's trust in  $O$  but gives  $P$  security of getting the money back. In the rational choice terminology, however, the bailment increases trust, more specifically the trust in the redemption of the loan. This makes sense only if the action of  $O$  or the outcome is accepted as an object of trust, as observed in the micro-economic experimental research.

Yamagishi and his colleagues follow a different track and disregard this version of trust. They define the concept more narrowly as interpersonal trust (Yamagishi and Yamagishi 1994). Secondly, they argue that  $P$ 's cooperation in a given situation is a manifestation of trust only to the extent that it is not determined by positive or negative sanctions or other conditions of the external situation. Accordingly, a distinction is made between the effects of an external situation and the effects of interpersonal trust. Whether we trust a particular person  $O$  depends on  $O$ 's traits, values, motives, etc., or more precisely on what we perceive of them. The factors that are typically manipulated in economic experiments like rewards, sanctions, or restrictions, by contrast, alter the incentive structure but not the interpersonal trust in  $O$ .

The incentive structure gives  $P$  more or less *assurance* that  $O$  will cooperate. If defection is heavily punished and  $O$  has little chances of escape, he or she will most likely cooperate. Little (interpersonal) trust is required in such a situation. Accordingly, assurance has to be distinguished from trust. The less incentives  $O$  has for cooperation, the lower in other words the assurance, the more interpersonal trust in  $O$  is required for cooperation on the side of  $P$ . If the necessary amount of trust is lacking,  $P$  will not cooperate.

Assurance is, to a large degree, calculable. If the positive and negative sanctions and the restrictions of the external situation are known, gains and

<sup>14</sup> This Subsection is based on the presentation of Toshio Yamagishi

losses of alternative actions can be determined, and in many instances, what a rational actor would do. Trust seems to be less calculable. Is it therefore nothing but gullibility? One of the strongest expressions of generalized distrust—i.e., distrust of human nature in general—can be found in a Japanese proverb: “It’s best to regard everyone as a thief” (*hito wo mitara dorobo to omoie*). An expression on the other extreme, generalized trust, can be found in another Japanese proverb: “You will never meet a devil as you walk through the social world” (*wataru seken ni oni ha nai*). In responses to questions concerning these two proverbs, 53% of Sapporo residents considered (N = 101) that those who believe the former proverb are wiser than those who believe the latter proverb, whereas only 25% considered the latter wiser than the former (the remaining 22% answered both are equally wise). Similarly, 56% considered the former more likely to be successful in life than the latter, as compared with 26% who considered it the other way around. A similar pattern was observed in a larger student sample at Hokkaido University.

These findings indicate that the majority of those who responded to these questions believed that distrust means social shrewdness and trust means gullibility. The results of experimental and survey research, however, provide evidence contrary to this popular belief. For example, Sturgis and Allum (2010) analyzed data from U.K. birth cohort studies and found that intelligence test scores at age 10 or 11 had a positive effect on the responder’s level of general trust at age 34 or 46, even after controlling their social success such as SES. Yamagishi, Kikuchi, and Kosugi (1999) analyzed data from 212 residents of Sapporo, Japan, and found that the respondents’ level of interpersonal emotional intelligence (EQ) and their self-report measure of cognitive role taking both positively correlated with their level of general trust. Furthermore, Yamagishi, Kikuchi, and Kosugi (1999) and Yamagishi (2001) examined how accurately people judge when the target persons behaved in a pro-social manner or not from their pictures and found that high-trusters were more accurate in their judgment of the target persons’ pro-sociality than low-trusters. Finally, in a more recent study, Shinada, Yamagishi, and Krasnow (2011) showed Japanese and American judges a set of five-second video clips of the target persons (Japanese nationals) and asked them to determine which target persons behaved in a pro-social manner in a previous experimental game. Both Japanese and American judges in this study were able to tell pro-social targets from pro-self targets, at least better than chance, just by looking at five-second video clips of these people reading a children’s story. Furthermore, regardless of the nationality of the judges, high trusters were able to identify the level of pro-sociality of the video clipped targets better than low trusters.

These findings suggest that trust and social intelligence have “co-evolved,” and distrust and lack of social intelligence constitute a vicious cycle.

On the one hand, generalized distrust prevents people from engaging in further social interactions. Low-trusters are unwilling to enter into potentially beneficial but risky social interactions because they focus on the risky side of such interactions. This unwillingness of distrusters to engage in potentially beneficial but risky social interactions deters them from correcting their depressed level of general trust. At the same time, their unwillingness to engage in risky but potentially fruitful interactions prevents them from improving the level of their social intelligence. The lack of social intelligence makes them vulnerable in such risky but potentially fruitful interactions. This vulnerability will then have two consequences. First, the lack of social intelligence makes them more gullible when they do in fact engage in such interactions. They will more often have experiences of failure than success in such interactions, and they will further learn to distrust others. Second, realizing this vulnerability, they will avoid engaging in such interactions. By engaging in such social interactions, they learn to distrust. By not engaging in such social interactions, they lose opportunities to improve the level of their social intelligence, that is, the ability to understand one's own and other people's internal states and use that understanding in social relationships.

The experimental and survey research (Sturdis, Read, and Allum 2010) support the claim that high-trusters are more socially intelligent than distrusters. The level of general trust reflects, both at the individual level and societal level, the overall level of opportunity costs for staying in the relatively stable and secure social relationships in which untrustworthy behaviors are well controlled. We can make sense of the findings concerning the positive relationship between social intelligence and general trust in terms of two general strategies to deal with social uncertainty and opportunity costs: opportunity seeking versus security seeking strategies. Opportunity seekers look outside the stable and secure relationships for better opportunities. In order to exploit these opportunities in socially risky situations, they invest cognitive resources in developing the ability to predict other people's behavior in an open environment. Those who have accrued social skills to successfully deal with social risks can afford to enter into risky but potentially profitable relationships. Having a high level of general trust encourages people to seek such opportunities. Security seekers, on the other hand, pay opportunity costs in exchange for the security that stable relationships provide and invest cognitive resources in assessing the nature of interpersonal relationships. They are good at detecting who would be an ally, and everyone else is regarded as a "thief." The characteristics of high-trusters (i.e., the correlation between general trust and the perceived need to cooperate with others, the sense of self-determination, and the low level of social risk avoidance) are indicative of opportunity seekers who leave the security of committed relationships to pursue better

opportunities. Conversely, the characteristics of believers in the thief proverb (i.e., the lack of a perceived need for cooperation or a sense of self-determination, social risk avoidance, and the lack of social skills) are likely characteristics of individuals who prefer not to deal with people outside of their secure relationships. The strategy that people adopt depends on the opportunities open to them. An opportunity-seeking strategy is more adaptive in a social environment in which staying in the stable and secure relationship entails large opportunity costs, so the social-explorer type of social intelligence is more likely to prosper there. In contrast, a commitment-formation strategy is more adaptive in a social environment in which stable and secure relationships do not entail large opportunity costs, so the commitment-former or security-seeker type of social intelligence will prosper there. The high correlation (at the national level) between the average level of individualism as compared with collectivism and the average level of general trust ( $r = .72$ ) provides support for this argument.

## **5. General Interpersonal Trust and Trust in Institutions<sup>15</sup>**

Trust in institutions is connected to interpersonal trust, first of all, in so far as institutions are always represented by human beings. If the representatives of an institution have proven themselves trustworthy for a long period of time, this will have a positive effect on the reputation of the institution. Vice versa, people working for the institution can benefit from its positive reputation. There are not only the institution and the incumbents of institutional roles, however, but also customers and consumers, or, in the case of public institutions, the citizens. Findings from the international survey programs suggest that the generalized interpersonal trust of the citizens increases with economic and human development. There is, on the other hand, the persistent lament that trust in institutions in Western democracies is declining. Are these apparently contradictory tendencies accidental, or are they generated by the same process of modernization and globalization? Does the latter lead to higher levels of generalized interpersonal trust, and at the same time destroy the trust in institutions? Or does each culture bear its own pattern of interpersonal and

---

<sup>15</sup> This subsection is based on the presentation by Gisela Trommsdorff. She reports about parts of the project "Developmental Conditions of Intentionality" (Principal Investigator: Prof. Dr. Gisela Trommsdorff; DFG GZ, TR 169/14-1, 2) in cooperation with the project "The Lawmaker's Intention and Its Limits" (Principal Investigator: Prof. Dr. Hans-Christian Röhl) within the interdisciplinary research group "Limits of Intentionality" (DFG Research Unit 582) at the University of Konstanz, Germany. We thank Tobias Heikamp for his contributions to Section 5.

institutional trust and its own path-dependent pattern of change? Questions like these show that the phenomenon of trust is a most valuable topic for interdisciplinary collaborative research because it bridges different social science disciplines and several levels of analyses: the individual, the group, and the institutional level. This approach is based on the idea that a person and society are interrelated by social interaction and exchange, and by sharing beliefs in the meaning of self-environment relationships (see Markus and Kitayama 1994, 2003). However, the kind and quality of such interrelations and beliefs differ among individuals and across groups and societies. This can be exemplified by the phenomenon of trust.

Functional properties of trust. Trust (measured at the national level) has been shown to be associated with subjective well-being, volunteerism, democratic attitudes, and greater preference for cooperation in government and civic life (e.g., Tov and Diener 2009). However, the direction of possible causal effects for these relationships is unknown. Therefore, we start with a discussion on possible functions of trust. Engaging in social interactions implies the exchange of trust. This applies to informal and intimate interactions, and to formal interactions on the level of political negotiation, e.g., between political parties, nations, or international organizations. In its most general form, interpersonal trust may be conceived of as individuals' judgments and evaluations of other people's reliability and willingness to cooperate. Accordingly, we distinguish here between generalized interpersonal trust and trust in institutions and discuss how these forms of trust are related to individuals' self and world views.

Generalized interpersonal trust promotes the stability of interpersonal, intergroup, and international relationships even under changing situational conditions. It is an open-ended question as to whether socio-political change will induce a fading or loss of trust (into distrust) and thus change the quality of the respective relationships. Trust allows for a wide range of positive interactions, whereas distrust limits the actor's range of interactions to defensive- and avoidance-motivated behavior. Trust has several functions on the individual, social, and institutional level. Trust may function to increase individual, social, and political security and to reduce insecurity, respectively. Accordingly, we should expect positive associations between generalized interpersonal trust and trust in institutions. International comparisons on relationships of trust at the individual level have indeed shown that generalized interpersonal trust, trust in institutions, and satisfaction with democracy are inter-related (ESS data from 24 countries) (Zmerli and Newton 2008). Does this result indicate that interpersonal trust and trust in institutions are related because of the same underlying factors? Next we will discuss the possible origins and effects of trust from a psychological and cultural perspective.

Psychological factors in the origins of trust. So far, little is known about whether the psychological conditions of generalized social trust are the same as for trust in institutions. A recent study by Schoon and Cheng (2011), however, suggested that individual differences in social attitudes and trust in institutions are linked to developmental experiences and activities (e.g., family background, cognitive ability, educational achievement). Development of generalized interpersonal trust reflecting the belief that other persons in general tend to be reliable, supportive, and cooperative can be described by the *working model of secure attachment* (Bowlby 1969; Cassidy, Kirsh, Scolton, and Parke 1996). Attachment literature has demonstrated various developmental outcomes of secure (as compared with insecure, avoidant, or disorganized) attachment. For example, secure attachment is related to the development of social competence (including positive social interactions), exploration, and openness for new experiences (which in turn activate cognitive development). A basic precondition for the development of secure attachment is the availability of sensitive caretakers (safe haven) in early childhood who are able to reduce a child's distress or negative experiences. Attachment security allows the child to develop a positive view of the self and the world; these are basic preconditions for the development of generalized interpersonal trust. Research on *interpersonal trust* at the individual level is in line with attachment theoretical notions. A positive worldview and high optimism facilitate trust in others (low fear of being taken advantage of) and promotes the motivation to cooperate with others (e.g., civic engagement in Sullivan and Transue 1999).

The question arises whether *trust in institutions* is also based on the individual's development of attachment security. Different from the study by Zmerli and Newton (2008), basic differences in generalized social trust and (political) trust in social institutions can be assumed. According to Listhaug (as cited in Levi and Stoker 2000: 498), trust in institutions "reflects evaluations of whether or not political authorities and institutions are performing in accordance with normative expectations held by the public." Instead, trust in institutions should be related to past and ongoing experience with institutions. Accordingly, trust in institutions may primarily depend on the individual's knowledge about normative expectations of the public, the motivation to share these expectations, and the evaluation of the past, present, and expected future performance. This implies that general trust in political institutions is related to the degree to which representatives of institutions (e.g., members of parliament) are perceived as trustworthy and capable of acting in the public interest (Levi and Stoker 2000).

Therefore, it should be asked whether those psychological conditions that are crucial in interpersonal trust, such as attachment, may also be conducive for trust in institutions. Values in favor of stability and security were positively

associated with trust in institutions. In contrast, values emphasizing openness to change were negatively related with trust in institutions (Devos, Spini, and Schwartz 2002). According to Schwartz' theory (Schwartz, in press), openness (freedom/self-direction) is on the opposite end of conservation (conformity, security, and social order). Individuals who are more open to change, in contrast to individuals who give priority to conformity and security over freedom, perceive institutions as less trustworthy as they contribute to the preservation of social and political arrangements in society (Devos, Spini, and Schwartz 2002).

Thus, the question becomes, what influences individuals' value priorities and national differences in values and beliefs about institutions? Next, this paper will shed light on the functions and origins of trust from a culture-informed perspective.

Cultural perspective on trust. From a culture-informed approach focusing on Japanese thinking, *trust* has been discussed in contrast to *assurance* (Yamagishi and Yamagishi 1994). Experimental studies revealed that, independent of culture, individuals tend to establish committed relationships when the perceived risk of being taken advantage of is high in order to reduce social uncertainty and to assure mutual cooperation (i.e., assurance). Moreover, individuals from cultures emphasizing independence (e.g., the United States) are more likely to trust other people in general than individuals from cultures that emphasize interdependence (e.g., Japan). In cultures in which cooperation is more likely to be based on assurance than on judgments of an individual's reputation, committed relationships with significant others are typically given priority over relationships with strangers (i.e., in-group versus out-group). This describes the basis of interpersonal relationships in Japan.

This conceptualization is in line with theorizing on cultural models of agency and related types of autonomy and relatedness (Markus and Kitayama 1994, 2003; Rothbaum and Trommsdorff 2007; Trommsdorff and Rothbaum 2008; Trommsdorff 2009, 2012). Accordingly, trust can be distinguished from assurance when referring to the concept of independence vs. interdependence (in self-construal), disjoint vs. conjoint mode of agency, and promotion vs. prevention focus.

Independence is associated with an emphasis on autonomy and need for self-reliance such as relying on one's own capabilities to evaluate other people's trustworthiness and the freedom to choose with whom to cooperate. In contrast, in cultures with a strong emphasis on interdependence, responsibilities toward others and social obligations are given priority over internal attributes (e.g., personal interests, preferences). As behavior is motivated by a need for relatedness, it aims to avoid disappointing social expectations. Therefore, commitment in interactions among members of the same group is fairly

high as cooperative behavior fosters interpersonal relationships and is perceived as socially rewarding. Basic aspects of trust reside in the self-construal as autonomous and separate and the general attribution to a person's traits, the emphasis of free choice and responsibility for one's actions. Therefore, negotiations are needed to ensure a commitment of the other person to engage in cooperative behavior. In contrast, assurance originates in the conjoint mode of agency, which regards the self as interdependent and related. Here, the expectation of cooperative behavior is based on personal agency, mutual obligations, and on prevention-focused goals that are motivated by a need for belongingness (Rothbaum and Trommsdorff 2007). Both trust and assurance are motivated by a need for security. According to Yamagishi (2011), the need for security and the avoidance of insecurity can be achieved through the two different pathways of trust and assurance.

So far it is unclear whether these concepts can be applied to the study of trust in institutions (and its possible relationship to general social trust). In a recent interdisciplinary cross-cultural project, a study explored the associations among value orientations, trust, and acceptance of political (national and supra-national) institutions (Tyler 2006). Preliminary results of self-reports of Japanese and German students in law show more similarities than differences with respect to group level means, and also with respect to associations between value orientations and "trust" (acceptance of political regulations). Further data on assurance (and general social trust) will be analyzed next in order to draw conclusions about whether and how far the culture-specific concept of assurance allows for better understanding of origins and functions of interpersonal trust and trust in institutions in Japan and Germany.

These cross-cultural analyses are expected to clarify the question of whether generalized interpersonal trust precedes institution-building and therefore is a prerequisite for the function of international organizations and cooperation (Rathbun 2011, 2012) (in political science, the reverse relationship has usually been assumed), and whether trust is related to satisfaction with political institutions, the belief in political self-efficacy, and actual civic engagement. This study attempts to contribute to a theoretical framework of aspects, origins, and functions of social trust from an interdisciplinary perspective. This is an increasingly important issue because globalization reshapes relationships among citizens, national governments, and international institutions. National interests, for instance, can be in conflict with engagement in joint collaborative activities and federal interests. Therefore, we try to understand social and cultural foundations and functions of trust in interpersonal relationships and in relationships between citizens and institutions.

Fostering trust in national institutions and promoting satisfaction with democracy may encourage civic engagement that in turn may increase trust in

international institutions through perceived opportunities to influence political decision making.

## **6. Discussion and Conclusion**

Several points have been addressed in the discussion:

- a) Different research traditions and schools use different concepts of trust, differentiated according to the criteria in Section 2 and embedded in different theoretical frameworks. Whether the latter can be integrated into a single, unifying theory remains an open question. To some extent, the existing frameworks overlap. If the behavior of the trustee  $O$  is classified as the outcome of a human system, one can say, for instance, that the outcome plays an essential role in almost all forms of trust. The past outcome tells us whether the system performance has met our expectations and partly determines the reputation of the system. The expected future outcome is either the object of trust or a central aspect of trust formation. Whether gains  $G$  and losses  $L$  can also be fruitfully implemented into the existing frameworks is much less clear. They certainly are highly relevant to economic exchange. If they were generalized to all forms of social interactions, immaterial benefits and losses had to be included in  $L$  and  $G$  too. As a further consequence, we had to find a measurable utility function that is equally sensitive to material and immaterial gains and losses. Furthermore, the usefulness of the two concepts becomes less clear, the more the object of trust is generalized. Do people really think in terms of gains and losses if they are asked about their general trust? More likely, they compare the behavior or outcome with their standards and expectations in such a case. The better the outcomes have met the standards and expectations in the past, the greater the trust. Take, for instance, the example of trust into the judicial system. People may place high trust in courts because they believe that judges are independent and by and large fair. Perhaps citizens also believe that their judicial system contributes to the welfare of the country and ultimately to their own welfare. They would not be able, however, to quantify the contribution—that is, to explain how much the judicial system contributes to the country's welfare or to their own. Adding gains  $G$  and losses  $L$  as two additional variables to the framework would only increase the measurement problems, but not improve the explanatory power.
- b) The ultimate goal of all social sciences is the understanding and explanation of human behavior. Surveys sometimes include items on past behav-

ior (cf. the vote in the last election), but typically focus on opinions, attitudes, values, or other orientations. Several studies have found a positive relationship between interpersonal trust and (reported) volunteering, but it remains unclear in these studies whether trust is the cause of volunteering or, vice versa, volunteering results in an increase of trust. Experiments are advantageous in this respect because they can control the temporal sequence of cause and effect. They can examine, for instance, whether people with high trust at the beginning of the experiment also cooperate more often at a later stage.

- c) Micro-economic and psychological experiments, on the other hand, face the problem of external validity. Is the money that the student subjects obtain at the beginning of the experiment really comparable with the money we spend in everyday life? If trust is not reciprocated in daily life, people experience a loss; if O defects in the experiment, P may perceive it as a lost profit. Micro-economic experiments on trust deal with these problems of "external validity" in an exemplary manner. Researchers do not confine themselves to experiments with students but try to utilize their empirical findings for improving economic exchange in daily life. It is not surprising that the experimental results have been, in the first place, applied to Internet platforms like eBay or Amazon because Internet shopping so far resembles games between anonymous players to a large degree. Reputation systems based on information about former transactions of buyers and sellers are therefore an important source of trust. Nevertheless, it usually turns out that reality is more complex than the experiments, and that the latter have to be redesigned. The close feedback between experiment and practice, however, offers hope that we gradually gain a deeper understanding of human behavior.
- d) Yamagishi's concept of trust (Section 4) probably comes the closest to the common understanding of trust. Trust is different from the mere rational calculation of expected gains and losses. It seems to be predominantly based on evaluations and emotions. Two problems, however, have been pointed out in the discussion. First, it seems to be nearly impossible to separate the effect of the incentive structure from the effect of personal traits. What people perceive as a personal property may, in fact, be a hidden effect of the opportunity structure. National stereotypes may serve as an example. If Japanese sellers are perceived as highly reliable, this may not be a feature of the Japanese national character but of the incentive or opportunity structure of Japanese markets. Is it a matter of trust if reliability is derived from O's character, and a matter of assurance if the reliability is seen as a feature of Japanese markets? And what about those many situa-

tions in which personal traits and motives interact with the context?

In his reply to this problem, Yamagishi pointed out that his distinction between assurance and trust is not based on the perceptions of the subjects but on objective criteria that can be controlled in the experiment. Experimentally, the two can be separated. For example, the difference in the level of trusting behavior by high trusters and low trusters will disappear when the level of assurance is raised or lowered substantially. Trust and assurance operate independently, though these two are mixed in the minds of most people. It is the role of the researchers to separate what looks the same in the mind of non-scientists. Instead of being a problem, the confusion between the two—to most people—is the reason why it is so important to separate the two.

The second problem is related to the first. The distinction between the *external* incentive structure and *internal* traits and orientations of O as sources of trust, draws a false border. P's verbal and non-verbal expressions of trust are almost certainly affected by expectations about O's behavior, which in turn is also influenced by P's knowledge of O's personality. P's social intelligence does not only determine how accurately the incentive structure is assessed but also how much knowledge about O is acquired and processed. Social intelligence is therefore also a determinant of trust or, at least, of the reactions that indicate trust. As the experiments in Section 3.3 have illustrated, people tend to underestimate the trustworthiness of others. Socially intelligent people are more competent in getting relevant information about O's intentions and traits and therefore express more interpersonal trust. They can also better protect themselves against fraud and are less often deceived. Cosmopolitans and highly educated people have a broader radius of contacts and can estimate the general trustworthiness of people more accurately<sup>16</sup>. It is possible that the pure emotional component of trust that so far cannot be appropriately measured is unaffected by information. Expressions of trust, however, are certainly heavily dependent on information. The more information we have about O's reliability, the better our judgments and predictions. It is therefore somewhat misleading to describe a trusting person as "...the one who overestimates the benignity of the partner's intentions beyond the level warranted by the prudent assessment of the available information" (Ya-

---

<sup>16</sup> Many other factors come into play: Citizens in open societies and consumers in open markets will, on average, interact more often with anonymous people and therefore also have better knowledge about other people. Vice versa, the citizens of societies in which committed relationships prevail should also express lower generalized interpersonal trust.

magishi and Yamagishi 1994, p. 136). Trust should also not be equated with “default” expectancies for the trustworthiness of other people (Yamagishi et al. 1999). Rather, high interpersonal trust is the result of a prudent assessment of other people—either of a particular person or, as generalized trust, of the traits and motives of human beings in general. Expressions of trust are as much dependent on information about other persons as assurance is dependent on information on the incentive structure.

- e) As far as the measurement of generalized trust is concerned, two problems were addressed. First, as questions usually do not specify the frame of reference, they are susceptible to framing effects. The statement, “Most people can be trusted,” for example, can refer among others to “most people I know,” “most people in my country,” or “most people in the world.” The choice of the frame of reference may be partly affected by the preceding questions (the so called halo-effects), or by the radius of contacts of the respondents (Delhey, Newton, Welzel 2011). The frame of reference of a cosmopolitan, for instance, almost certainly differs from that of a parochial. Secondly, it is unclear to what extent translations convey the meaning of the master questionnaire. Take the response alternative “complete trust,” for instance: Do Germans or Japanese express the same evaluation, emotions, and behavioral intentions as Americans when choosing this answer in their mother tongue? Have they performed the same cognitive operations? Languages are profoundly shaped by cultural traditions and values and it remains unclear whether translations are equivalent.

There are, undoubtedly, various unsolved issues that open up a vast field for future interdisciplinary research. On the other hand, the symposium has shown how research on trust and confidence has made considerable progress during the last decades. Game theory has helped us to model human interaction processes more accurately. Micro-economic experiments have proven particularly useful for engineering trust in Internet markets. Psychological and sociological research has given us a better understanding of cultural differences in trust formation.

## Literature

- Axelrod, R. 1984. *The Evolution of Cooperation*. New York: Basic Books.
- Axelrod, R. 1997. *The Complexity of Cooperation*. Princeton: University Press.
- Bolton, G. E., and A. Ockenfels. 2012. “Behavioral economic engineering.” *Journal of Economic Psychology* 33(3): 665–676.

- Bolton, G. E., and A. Ockenfels. 2009. "The limits of trust." In *eTrust*. K. Cook, C. Sniijders, and V. Buskens, eds. New York: Russell Sage. 15–36.
- Bolton, G. E., B. Greiner, and A. Ockenfels. 2012. "Engineering Trust – Reciprocity in the Production of Reputation Information." *Management Science*, 58(12), 2225–2233.
- Bolton, G. E., C. Loebbecke, and A. Ockenfels. 2008. "Does competition promote trust and trustworthiness in online trading? An experimental study." *Journal of Management Information Systems* 25(2): 145–169.
- Bolton, G. E., E. Katok, and A. Ockenfels. 2004. "How effective are online reputation mechanisms? An experimental investigation." *Management Science* 50(11): 1587–1602.
- Bolton, G. E., E. Katok, and A. Ockenfels. 2005. "Cooperation among strangers with limited information about reputation." *Journal of Public Economics* 89: 1457–1468.
- Bolton, G. E., A. Ockenfels, and F. Ebeling. 2011. "Information value and externalities in reputation building." *International Journal of Industrial Organization* 29(1): 23–33.
- Bowlby, J. 1969. *Attachment and Loss: Vol. 1., Attachment*. London: Hogarth Press.
- Buskens, V. and W. Raub. Forthcoming. "Rational choice research on social dilemmas." In *Handbook of Rational Choice Social Research*. R. Wittek, T. A. B. Sniijders, and V. Nee, eds. New York: Russell Sage Foundation.
- Cassidy, J., S. J. Kirsh, K. L. Scolton, and R. D. Parke. 1996. "Attachment and representations of peer relationships." *Developmental Psychology* 32: 892–904.
- Coleman, J. S. 1990. *Foundations of Social Theory*. Cambridge: Harvard University Press.
- Delhey, J., K. Newton, and C. Welzel. 2011. "How general is trust in 'most people'? Solving the radius of trust problem." *American Sociological Review* 76(5): 786–807.
- Devos, T., D. Spini, and S. H. Schwartz. 2002. "Conflicts among human values and trust in institutions." *British Journal of Social Psychology* 41: 481–494.
- Dunning, D., D. Fetchenhauer, and T. Schlösser. 2012. "Trust as a social and emotional act: Noneconomic considerations in trust behavior." *Journal of Economic Psychology* 33, 686–694.
- Easton, D. 1965. *A Systems Analysis of Political Life*. New York: Wiley.
- Fetchenhauer, D. and D. Dunning. 2009. "Do people trust too much or too little?" *Journal of Economic Psychology* 30: 263–276.
- Fetchenhauer, D. and D. Dunning. 2010. "Why so cynical? Asymmetric feedback underlies misguided skepticism regarding the trustworthiness of others." *Psychological Science* 21(2): 189–193.

- Fetchenhauer, D., and D. Dunning. 2012. "Betrayal aversion versus principled trustfulness. How to explain risk avoidance and risky choices in trust games." *Journal of Economic Behavior and Organization* 81(2): 534–541.
- Greif, A. 1993. "Contract enforceability and economic institutions in early trade: The Maghribi Traders' Coalition." *American Economic Review* 83(3): 525–548.
- Jagodzinski, W. 2012. "Trust: Concepts and theoretical frameworks." Cologne: Unpublished paper.
- Klemperer, P. 2004. *Auctions: Theory and Practice*. Princeton, NJ: Princeton University Press.
- Kosfeld, M., A. Okada, and A. Riedl. 2009. "Institution formation in public goods games." *American Economic Review* 99: 1335–1355.
- Levi, M., and L. Stoker. 2000. "Political trust and trustworthiness." *Annual Review of Political Science* 3: 475–507.
- Markus, H. R., and S. Kitayama. 1994. "A collective fear of the collective: Implications for selves and theories of selves." *Personality and Social Psychology Bulletin* 20: 568–579.
- Markus, H. R., and S. Kitayama. 2003. "Culture, self, and the reality of the social." *Psychological Inquiry* 14: 277–283.
- Ockenfels, A., and W. Raub. 2010. "Rational und Fair." In *Kölner Zeitschrift für Soziologie und Sozialpsychologie. Sonderheft 50: Soziologische Theorie kontrovers*. G. Albert, and S. Sigmund, eds. 119–136.
- Ockenfels, A., and P. Resnick. 2012. "Negotiating reputations." In *The Oxford Handbook of Conflict Resolution*. G. Bolton, and R. Croson, eds. Oxford: Oxford University Press. 223–237.
- Olson, M. 1965. *The Logic of Collective Action: Public Goods and the Theory of Groups* (2nd Ed., 1971). Cambridge: Harvard University Press.
- Ostrom, E. 1990. *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge: Cambridge University Press.
- Rathbun, B. C. 2011. "Before hegemony: Generalized trust and the creation and design of international security organizations." *International Organization* 65(2): 243–273.
- Rathbun, B. C. 2012. *Trust in International Cooperation: International Security Institutions, Domestic Politics and American Multilateralism*. New York: Cambridge University Press.
- Roth, A. E. 2002. "The economist as engineer: Game theory, experimental economics and computation as tools of design economics." *Econometrica* 70(4): 1341–1378.
- Rothbaum, F., and G. Trommsdorff. 2007. "Do roots and wings oppose or complement one another? The socialization of autonomy and relatedness in

- cultural context." In *The Handbook of Socialization*. J. E. Grusec, and P. Hastings, eds. New York: Guilford Press. 461–489.
- Schoon, I., and H. Cheng. 2011. "Determinants of political trust: A lifetime learning model." *Developmental Psychology* 47(3): 619–631.
- Schwartz, S. H. In press. "Values and religion in adolescent development: Cross-national and comparative evidence." In *Values, Religion, and Culture in Adolescent Development*. G. Trommsdorff, and X. Chen, eds. Cambridge: Cambridge University Press.
- Shinada, M., T. Yamagishi, and M. Krasnow. 2011. "General trust and accuracy of trustworthiness judgment." Hokkaido University Center for Experimental Research Working Paper, No. 121, May 9.
- Sturgis, P., S. Read, and N. Allum. 2010. "Does intelligence foster generalized trust? An empirical test using the UK birth cohort studies." *Intelligence* 38: 45–54.
- Sullivan, J. L., and J. E. Transue. 1999. "The psychological underpinnings of democracy: A selective review of research on political tolerance, interpersonal trust, and social capital." *Annual Review of Psychology* 50: 625–650.
- Tov, W., and E. Diener. 2009. "The well-being of nations: Linking together trust, cooperation, and democracy." In *The Science of Well-Being: The collected works of Ed Diener*. E. Diener, ed. New York: Springer. 155–173.
- Trommsdorff, G. 2009. "Culture and development of self-regulation." *Social and Personality Psychology Compass* 3: 687–701.
- Trommsdorff, G. 2012. "Development of 'agentic' regulation in cultural context: The role of self and world views." *Child Development Perspectives* 6: 19–26.
- Trommsdorff, G., and F. Rothbaum. 2008. "Development of emotion regulation in cultural context." In *Regulating Emotions: Social Necessity and Biological Inheritance*. S. Ismer, S. Jung, S. Kronast, C. v. Scheve, and M. Vandekerckhove, eds. New York: Blackwell. 85–120.
- Tyler, T. R. 2006. "Psychological perspectives on legitimacy and legitimation." *Annual Review of Psychology* 57: 375–400.
- Yamagishi, T., and M. Yamagishi. 1994. "Trust and commitment in the United States and Japan." *Motivation and Emotion* 18: 127–166.
- Yamagishi, T., M. Kikuchi, and M. Kosugi. 1999. "Trust, gullibility and social intelligence." *Asian Journal of Social Psychology* 2: 145–161.
- Yamagishi, T. 2001. "Trust as a form of social intelligence." In *Trust in Society*. K. S. Cook, ed. New York: Russell Sage Foundation. 121–147.
- Yamagishi, T. 2011. *Trust: The Evolutionary Game of Mind and Society*. Tokyo: Springer.

- Wilson, R. 1985. "Reputations in games and markets." In *Game Theoretic Models of Bargaining*. A. Roth, ed. New York: Cambridge University Press. 27–62.
- Zmerli, S., and K. Newton. 2008. "Social trust and attitudes toward democracy." *Public Opinion Quarterly* 72: 706–724.