

# VU Research Portal

## Incomplete Information in Social Interactions

Vuolevi, J.H.K.

2011

### **document version**

Publisher's PDF, also known as Version of record

[Link to publication in VU Research Portal](#)

### **citation for published version (APA)**

Vuolevi, J. H. K. (2011). *Incomplete Information in Social Interactions*. [PhD-Thesis – Research external, graduation internal, Vrije Universiteit Amsterdam].

### **General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

### **Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

### **E-mail address:**

[vuresearchportal.ub@vu.nl](mailto:vuresearchportal.ub@vu.nl)

# Chapter 1

## Introduction

The most elementary form of a decision is a choice between two alternatives. You walk to your living room and there is one apple and one orange in a bowl. You want to have something for a snack, and you choose either the apple or the orange. Classic economics theory, along with common sense, proposes that you make a rational choice that maximizes your utility. That is, you choose the fruit that, for whatever reason at that particular moment, you prefer the most.

This simple example becomes a lot more difficult if you enter the living room with your friend, and you are both hungry. One of you gets the apple, the other one the orange. But how would you decide which one gets which fruit? If one person prefers the apple and the other one the orange, the decision is easy. But if both prefer either the apple or the orange, the decision is more difficult. You may take the preferred fruit and leave the second best alternative to your friend, or you may let your friend have the favorite fruit.

The above example highlights three key aspects of social decision-making, and this Introduction is organized in subchapters accordingly. In the first subchapter I will describe how social decisions influence other people's outcomes, with a particular focus on situations in which self-interest and other-interest are at odds (e.g., both prefer the same fruit). I will also demonstrate how cooperation can sometimes emerge in such situation, even if people are assumed to pursue their self-interest. Second, I will review empirical evidence on people's tendencies to pursue self-interest versus other-interest (e.g., the favorite fruit to you or your friend) and demonstrate how preferences that differ from self-interest influence social interactions. Third, I will discuss how incomplete information influences cooperation. Sometimes people do not know their partner's exact preference (e.g., which fruit is the favorite) and the same behavior can be interpreted in multiple ways (e.g., did my friend choose apple because he likes it better or because he thinks that I like orange better). I will discuss in detail how incompleteness of information influences perceived cooperation as well as general evaluations of the partner—both of which may influence cooperation in return.

### Basic Principles in Social Interactions

To start with, I will introduce situations that are characterized with conflict of interest and define cooperation in this context. Second, I will introduce some basics of game theory—a framework for formal understanding of decisions in situations in which more than one person is involved. Third, I will review interpersonal strategies that people adopt in such situations and present some simulation-based as well as experimental

evidence to show that human cooperation is conditional: People cooperate with others who cooperate with them, and noncooperate with those who do not cooperate.

### **The Conflict Between Self-interest and Other-interest**

In the living room example the most elementary question is whether you and your friend prefer the same fruit or not. If you prefer different fruits, the situation is characterized by correspondence of interest. Human life is full of situations in which people's interests align and the key question is about coordination: What should we do so that we both can obtain the best outcomes (e.g., ask your friend's favorite fruit and get an honest answer). But if you prefer the same fruit, such perfect-for-everybody solution is impossible to obtain. These situations—also notably present in everyday life—are characterized by a conflict of interest and the key question is about cooperation: Should I pursue my self-interest (e.g., take the favorite fruit) or should I pursue another person's interest (e.g., offer the favorite fruit to the friend).

Behaviors that benefit another person or a collective, but are costly to the self, are defined here as cooperation. Thus, the concept of cooperation is only meaningful in the context of conflict of interest, because cooperation cannot be separated from self-interest if self-interest and other-interest align. The second important aspect of this definition is that cooperation refers to specific behaviors alone. For example, if you give the favorite fruit to your friend with an idea that you can ask a favor later, your behavior is still defined as cooperation. This is where a distinction can be made between cooperation and altruism. The former refers to specific behaviors whereas the latter, at least in its strictest sense, refers to underlying motives and ultimate goals of behaviors (see Sober & Wilson, 1998). Following terminology in Van Lange (2008), I use altruism somewhat more loosely and define it as an other-regarding motive that may underlie cooperative behaviors. The difference to the strictest definition is that altruism in this dissertation does not necessarily mean that a person would forfeit self-interest in all possible respects (e.g., donating money to a charity is still altruistic even one could argue it is self-interested if one feels good about it). The question whether all behaviors can be accounted for by self-interest is beyond the scope of this dissertation, but interested readers may refer to previous discussions (e.g., Batson, 1991; Cialdini & Fultz, 1990; Dovidio, 1984).

### **Game Theory**

Game theory is a framework for understanding social decision-making (Luce & Raiffa, 1957; Von Neumann & Morgenstern, 1944). It captures the decision options for each player and describes the way in which different combinations of behaviors influence outcomes for each player. In a dyadic case, this is often represented as a matrix where each row represents one behavioral option for Player 1 and each column represents one behavioral option for Player 2. Each cell consist of two values—outcomes for each

player that correspond with that particular behavior combination. Game theory can describe all possible ways in which two or more individuals can be interdependent (e.g., independence versus interdependence, corresponding versus conflicting interest, joint control of outcomes versus partner control of outcomes).

Table 1.1 presents an example of the best known dyadic game—the prisoner’s dilemma (Rapoport & Chammah, 1965; Tucker, 1950). In this game, both players have the same two options—either to cooperate or to defect. If both players cooperate, they receive the best possible collective outcomes (e.g., 4 points each, 8 points in total). However, if one player defects while the other one cooperates, the defector receives much higher outcomes (e.g., 6 points) than the cooperator (e.g., 0 points). If both players defect, they receive lower outcomes than they would by mutual cooperation (e.g., 2 points).

**Table 1.1: The prisoner’s dilemma.**

		Player 2	
		Cooperate	Defect
Player 1	Cooperate	4 / 4	6 / 0
	Defect	0 / 6	2 / 2

The prisoner’s dilemma is a widely used paradigm in social sciences, because it captures the conflict between self-interest and collective interest. Defection is the rational strategy, because it leads to higher outcomes than cooperation regardless what the other player does. At the same time, cooperation is the best collective strategy, because it leads to the best overall outcomes. But do people pursue self-interest or collective interest in the prisoner’s dilemma? What interpersonal strategies people apply and should apply in such mixed-motive situations?

### **Conditional Cooperation and the Emergence of Tit-for-tat**

According to game theory, players would always defect in the prisoner’s dilemma. Game theory is based on the assumptions that all players pursue their self-interest and that all players can assume that all other players pursue their self-interest, too. Defection is indeed the rational strategy that cannot be exploited by any other strategy (i.e., no strategy can gain a relative advantage over a defector). But defection is by no

means the perfect strategy in terms of utilizing the best collective outcomes provided by the situation. Two cooperators would gain twice as high outcomes as two defectors playing with each other. But at the same time, unilateral cooperation would be exploited by a defective partner.

To compare different interpersonal strategies, Robert Axelrod (1984) organized a computer tournament in which different strategies played the prisoner's dilemma game against each other. The game was iterated and the agents had complete information on their opponents past behavior, which allows more complex strategies than unconditional cooperation or unconditional defection. Leading experts in the world submitted numerous strategies, but the results emerged surprisingly simple: The best strategy, submitted by Anatol Rapoport, was tit-for-tat—strategy that makes the cooperative choice in the first trial and simply copies the partner's previous behavior for the next round (i.e., previous cooperation is responded with cooperation and previous noncooperation with noncooperation).

Population dynamics simulations provided further evidence for the effectiveness of tit-for-tat. A subpopulation of tit-for-tat agents can obtain cooperation with each other and sustain even among selfish agents (Axelrod, 1984). Tit-for-tat can also evolve from completely random strategies (Axelrod, 1997). Tit-for-tat is an evolutionally stable strategy (along with the self-interest strategy), which supports Trivers' (1971) original argument that cooperation can emerge through a reciprocal altruism (i.e., going beyond self-interest given that the partner does the same). Experimental findings also provide evidence for tit-for-tat. While some people use selfish strategies—idea that is consistent with population dynamics simulations showing that defectors also sustain—most people adopt a version of tit-for-tat in their interactions (Klapwijk & Van Lange, 2009; Komorita, Hilty, & Parks, 1991; Komorita & Parks, 1995; Parks & Rumble, 2001; Sheldon, 1999; Van Lange, 1999).

Tit-for-tat is an example of a conditional strategy and human cooperation is inherently conditional. As Axelrod has demonstrated (1984, 1997), conditional cooperation provides good outcomes with those who want to cooperate, but also a good protection against defectors. Some previous literatures refer to conditional cooperation by reciprocal altruism (Trivers, 1971) or reciprocity (e.g., Kollock, 1993; Nowak & Sigmund, 1992; Van Lange, Ouwerkerk, & Tazelaar, 2002). To avoid confusions with different definitions in social exchange theory (e.g., Homans, 1961) and self-disclosure, here I use conditional cooperation to refer to the general principle that people adjust their cooperation according to the partner.

The human brain is suited for conditional cooperation and social exchange in particular (e.g., trading money for goods, favors for other types of favors). Social exchange involves conditional logic (e.g., if P then Q) and people perform quite poorly at detecting flaws in such conditional statements in general. By contrast, when conditional statements are about social exchange (e.g., if you borrow my car, then you

must fill the tank with gas), people perform very well (Cosmides, Tooby, Fiddick, & Bryant, 2005; Ermer, Cosmides, & Tooby, 2007). This indicates that people's logical reasoning is suited for detecting social cheaters and identifying defectors from tit-for-taters.

## Social Interactions Beyond Self-interest

### Empirical Evidence and Models for Social Preferences

Game theory is a normative approach to rational social behavior and explains what people should do—given that they pursue their self-interest. However, even in a single-shot prisoner's dilemma with perfect strangers, some people cooperate (for a review, see Caporael, Dawes, Orbell, & Van de Kragt, 1989). This behavior cannot be explained by evolutionary fitness of tit-for-tat, because conditional strategies cannot be applied in single-shot games where players do not have any information on their partner's previous behaviors, nor can they distinguish defectors from tit-for-taters a priori.

Perhaps even a clearer evidence of violation of the self-interest assumption comes from experiments on *the dictator game*—game that creates a conflict between self-interest and equality. In this game, the first person (i.e., the dictator) can freely divide a fixed outcome (e.g., units of money) between the self and another person. Rational self-interest would suggest that the first person keeps all the outcomes, but most people actually allocate part of it, up to the 50-50 split, to the second person (Bolton, Katok, & Zwick, 1998; Bolton & Ockenfels, 2000; Camerer & Thaler, 1995; for a recent developmental evidence, see Fehr, Bernhard, & Rockenbach, 2008). These findings are typically interpreted as evidence for egalitarianism. Self-interest allocations would create a high degree of inequality, and people sacrifice their self-interest to obtain a smaller difference in outcomes between the self and another person (e.g., the inequality aversion model; Fehr & Schmidt, 1999; see also Deutsch, 1975; Güth & Tiez, 1990; Roth, 1995).

However, the egalitarianism explanation can not explain cooperation in a single-shot prisoner's dilemma, because, without knowing what the partner would do, the cooperative choice is not a better choice than noncooperation in terms of equality in outcomes (i.e., outcomes are equal if both players make the same choice, and equally unequal if one player cooperates while the other one does not). To understand this behavior, some researchers have argued that people tend to be somewhat altruistic, in a sense that other people's outcomes have a positive weight and that people are willing to sacrifice, to some extent, their own outcomes for another person (e.g., Batson, 1991; Van Lange, 2008). This alternative explanation would also explain behavior in the dictator game. If people care about the second player's outcomes, they indeed would allocate some outcomes to that person.

The ring measure of social values is a variant of the dictator game, in which people make 24 binary choices between two outcome distributions for self and another person (Liebrand, Jansen, Rijken, & Suhre, 1986). These outcome distributions are sampled in such a way that self-interest, egalitarianism and altruism motives can be distinguished from each other. The findings provide clear evidence that self-interest is the main motive, but also that both egalitarianism and altruism influence social behavior (Van Lange, 1999).

First and foremost, these behavioral experiments show that classic economics and game theory, relying on the self-interest assumption, make quite a reasonable assumption that self-interest is indeed the strongest single motive in interpersonal behavior. At the same time, the self-interest assumption fails to capture some important aspects of social decision-making. The egalitarianism motive has been incorporated in virtually all modern social decision-making models (e.g., Fehr & Schmidt, 1999; Loewenstein, Thompson, & Bazerman, 1989). The integrative model of social values also includes the altruism component (Van Lange, 1999).

### **The Interdependence Approach to Social Interactions**

In the previous section I discussed social preferences—outcome distributions that people prefer between self and others. I provided evidence that preferences often deviate from self-interest, but social preferences are not the only determinant of behavior in social interactions. Preferences are, literally, just preferences, and actual outcome distributions are also importantly shaped by the interaction partner's behavior as well as the structure of the social situation.

In this section I provide a framework that captures interdependence the same way as game theory does, but which is not limited to self-interest social preferences. This framework is called interdependence theory, which, similar to game theory, uses a matrix (or a similar representation) that describes outcomes as functions of interaction partners' possible behaviors (Kelley & Thibaut, 1978; for an overview, see Rusbult & Van Lange, 2003). Unlike game theory, interdependence theory does not make the assumption that people would make their decisions based on the game theoretical matrix alone.

Interdependence theory posits that people *transform* their motives from immediate self-interest (i.e., the game theoretical, given situation) to the effective situation, which incorporates motives broader than immediate self-interest. Transformations occur for many different reasons including long-term considerations, norms, and equality (for an overview, see Rusbult & Van Lange, 2003). Here, I focus on explaining how different social preferences are related to transformations, and provide a few case examples how transformations change the game theoretical situation. To illustrate, I will use the integrative model of social value orientation to explain how social decisions can be influenced by self-interest, altruism and egalitarianism motives (Van Lange, 1999).

**Table 1.2: The prisoner's dilemma before and after altruism transformations. Numbers after the arrow sign are transformed outcomes.**

		Player 2 = $0.5S + 0.5O$	
		Cooperate	Defect
Player 1 = $0.5S + 0.5O$	Cooperate	4 4	$6 \rightarrow 3$ $0 \rightarrow 3$
	Defect	$0 \rightarrow 3$ $6 \rightarrow 3$	2 2

Table 1.2 represents the prisoner's dilemma before (i.e., the given situation) and after altruism transformations (i.e., the effective situation, after the arrow signs). In this example, both players put the same weight on their own as well as the partner's outcomes (i.e.,  $0.5 \times \text{Self} + 0.5 \times \text{Other}$ ). Outcomes within each cell are now identical, because equal preferences for self and other corresponds with the idea that both players prefer the highest possible combined outcomes, regardless how much each player gets. Now, cooperation gives one point more independent of what the partner does. Thus, if both transform their motives like this, mutual cooperation occurs.

**Table 1.3: The prisoner's dilemma before and after egalitarianism transformations. Numbers after the arrow sign are transformed outcomes.**

		Player 2 = $0.5S + 0.5E$	
		Cooperate	Defect
Player 1 = $0.5S + 0.5E$	Cooperate	4 4	$6 \rightarrow 3$ 0
	Defect	0 $6 \rightarrow 3$	2 2

Table 1.3 presents the prisoner's dilemma before (i.e., the given situation) and after egalitarianism transformations (i.e., the effective situation, after the arrow signs). In this example, both players weight their own outcomes and equality in outcomes to the same extent (e.g.,  $0.5 \times \text{Self} + 0.5 \times \text{Equality}$ ). Outcomes in diagonal cells (i.e.,



mutual cooperation or defection) do not change to the given situation. Even though both weight their own outcomes less than in the given situation (i.e.,  $S=1$ ), those two cells provide the perfect equality in outcomes, which increase their values back to the original numbers. However, the combination of cooperation and defection is particularly poor in terms of equality. For the cooperator this situation is bad in two respects: The outcomes are the lowest (e.g., 0) and the inequality is the greatest (e.g., 6 vs. 0). But also for the defector the situation is not ideal: The outcomes for self are good (e.g., 6) but because that person also weighs equality, and this combination of behaviors is very bad in terms of equality, the effective outcomes are lower (e.g., 3).

Both altruism and egalitarianism transformations change the prisoner's dilemma's outcome matrix similarly. The transformed outcomes are not identical, but similar enough that both transformations can explain why some people cooperate. Thus, transformations can explain the empirical fact that some people cooperate in the prisoner's dilemma, but one game theoretical situation is not enough to obtain information on specific transformations (see Weibull & Salomonsson, 2006).

Specific transformations can be derived by measuring behavior across different social situations. Previous experiments have manipulated the outcome structure of the prisoner's dilemma and revealed that people may defect for two distinct reasons. People may defect because they are greedy and want to obtain the best possible outcomes, or because they are fearful that the other player may defect (Coombs, 1973; Rapoport, 1967; Van Lange, Liebrand, & Kuhlman, 1990; Yamagishi & Sato, 1986). Using a similar approach, Simpson (2003) demonstrated that males are more likely to defect for greed and females for fear.

To summarize, interdependence theory posits that behavior in social interactions is a product of the situation and its characteristics (e.g., conflicting vs. corresponding interest) as well as transformations, which can account for individuals' different outcome preferences between self and others. This is an important addition to game theory and social decision-making models, because the former focuses only on the situation and the latter only on general and person-specific outcome preferences. In a way, interdependence theory is a formalization of Kurt Lewin's (1936) original idea that behavior (B) is a product of the person (P) and the environment (E). Even though Lewin presented this idea as an equation (i.e.,  $B = f(P, E)$ ), interdependence theory is the only theory that actually describes the relationship between person variables, situational variables, and the interaction between the two, as illustrated in this section.

### **Incomplete Information in Social Interactions**

Conditional cooperation clearly has its benefits as partners can protect themselves against noncooperation and still benefit from mutual cooperation with those who want to cooperate. But this benefit comes with a cost: Compared to unconditional

cooperation or unconditional noncooperation, conditional strategies require information on the partner's previous behavior.

The assumption that people have complete information on their partner's past behavior has been a prevailing, and often unquestioned, assumption in both theory and practice. Experimental research has a long tradition of representing games as matrixes that reveal complete outcome information. Such experiments also provide complete information on past behaviors (i.e., one particular choice option was chosen over the others). Thus, in vast majority of experiments of cooperation—which conclude that people use tit-for-tat—partners' have complete information on each other's previous cooperation.

In this section, I will first distinguish three ways in which information in social interactions can be incomplete (Kelley et al, 2003). Second, I will review previous literature that suggests how people might perceive their partners' cooperation when information is incomplete. Finally, I will discuss how incompleteness of information may influence both perceived cooperation and perceived transformations. I will also demonstrate how misperception of the partner's transformations may have a long-lasting influence on mutual cooperation.

### **Incomplete Information about Cooperation and Transformations**

*Incomplete behavioral information* refers to uncertainty about the partner's exact behavior. People may know what the partner could do (i.e., the choice options), but they do not know for sure which one of these possible behaviors was chosen or will be chosen. The most extreme case of incomplete behavioral information is no information at all. For example in a single-shot prisoner's dilemma, players have no information on each other's behaviors. The second example of incomplete behavioral information is probabilistic information: One may know the probability of the partner's cooperation (e.g., 80%), but does not know whether a particular behavior was cooperation or noncooperation. Third, information regarding a specific behavior can be erroneous: People may receive information on cooperation whereas the actual behavior was noncooperation or vice versa (e.g., erroneous second-hand information). In this case, incompleteness of behavioral information is related to unreliability of information.

Table 1.4 illustrates a simple situation in which Person A can either cooperate or noncooperate, and Person B can interpret this behavior either as cooperation or noncooperation. The two diagonal cells represent the correct perception: Cooperation is correctly perceived as cooperation or noncooperation is correctly perceived as noncooperation. These are the two possible scenarios when information about Person A's behavior is complete—no errors in the perception of behavior is possible if information is explicit and deterministic.

**Table 1.4: Illustration of incomplete behavioral information**

		Person B's perception of Person A's behavior	
		Cooperation	Noncooperation
Person A's behavior	Cooperation	Cooperation correctly detected	Noncooperation incorrectly detected
	Noncooperation	Cooperation incorrectly detected	Noncooperation correctly detected

The two other possibilities are that Person A had not cooperated but Person B actually perceived this noncooperation as cooperation (i.e., cooperation incorrectly perceived), or that Person A had cooperated but Person B actually perceived this cooperation as noncooperation (i.e., noncooperation incorrectly perceived). These possibilities may occur only when people have incomplete information on their partner's behaviors. Perhaps more realistically, many behaviors are not binary choices between cooperation and noncooperation, but people may cooperate to a different degree (e.g., between 0 and 10). In this case, people may not know the exact level of cooperation (e.g., 5), but they may know the range (e.g., between 4 and 6). Thus, incomplete behavioral information may also refer to distributional information regarding the partner's behavior.

*Incomplete situational information* refers to missing outcome information. Often people know the value of the partner's behavior for themselves (e.g., a colleague commenting a manuscript), but they do not necessarily know the cost or benefits to the person who helped them (e.g., how much time it took). Also, people may not know all the behavioral options for the partner which makes it difficult to judge the cooperativeness of a specific behavior.

Table 1.5 illustrates this example. Person B has complete behavioral information by knowing whether the partner gave comments on the manuscript or not (i.e., misperception of behavior is not possible). Also, person B probably knows the benefits of receiving valuable feedback from a colleague (i.e., complete situational information with outcomes associated to the self). However, Person B can probably only estimate the costs and benefits of commenting or not commenting the manuscript that incurred for Person A (e.g., did the colleague intrinsically enjoyed reading the manuscript or how much time it took). Person B is probably also quite unaware of alternative behaviors and their outcomes for Person A (e.g., whether the colleague would have been busy with self-interest goals or not).

**Table 1.5: Illustration of incomplete situational information**

		Person B's perception of Person A's behavior	
		Commented the manuscript	Did not comment the manuscript
Person A's behavior	Commented the manuscript	Received feedback Took time? Enjoyed the paper?	
	Did not comment the manuscript		Did not receive feedback Busy for a good reason?

*Incompleteness of transformational information* refers to uncertainty about the partner's general cooperative versus noncooperative tendencies across different social situations. Different individuals put different weights on self-interest, egalitarianism, and altruism, and therefore exhibit a wide range of behaviors ranging from competition to cooperation (e.g., social value orientation; Balliet, Parks, & Joireman, 2009; Van Lange, 1999; Van Lange, De Cremer, Van Dijk, & Van Vugt, 2007). People may know their own transformations (e.g., egalitarianism preference), but information about others' transformations is almost always incomplete. This is because motives and intentions that underlie others' behaviors cannot be directly accessed, but they must be inferred from behavior (for an overview, see Pronin, 2008; see also Ross & Sicoly, 1979). As I have discussed, even with complete information people may make mistakes in inferring their partner's transformations. With incomplete behavioral or situational information, the underlying transformations are even more difficult to attain.

### **Perceived Cooperation Under Incompleteness of Information**

How people perceive their partners' level of cooperation under incompleteness of information? Incomplete information often sets some boundaries (e.g., behavior cannot be extremely noncooperative or cooperative), but the exact level of cooperation must be inferred. This is a special feature of incomplete information: The missing pieces of information must be filled in.

Previous research suggests that inferences might be driven by the assumption of other people's self-interest. Research on the "norm of self-interest" reveals that global judgments about unknown others are guided by a belief in self-interest (see Miller & Ratner, 1996, 1998). For instance, people overestimate the impact of financial rewards

on their peers' willingness to donate blood. People also attribute responsibility in a self-serving way. For example, people think that their spouses are more responsible for negative than for positive events in their relationships, whereas people think of themselves being responsible for both positive and negative events (Kruger & Gilovich, 1999). Further evidence shows that these cynical theories about other people are more pronounced and lead to more selfish behavior when people are encouraged to think more about others' thoughts (e.g., Epley, Caruso, & Bazerman, 2006; Vorauer & Sasaki, 2009).

Another line of research demonstrated that dispositional attributions are also guided by self-interest. Research on interpersonal biases reveals a stable trait bias in that people think of others as more selfish and less fair than they think of themselves (Alicke, Dunning, & Kruger, 2005; Allison, Messick, & Goethals, 1989; Dunning & Cohen, 1992; Messick, Bloom, Boldizar, & Samuelson, 1985; Van Lange & Sedikides, 1998). Interestingly, this research reveals that in comparison to many other attributes (e.g., those linked to competence) such better-than-average (i.e., superiority) effects tend to be most pronounced for attributes that are strongly linked to social qualities (e.g., morality, honesty).

Thus, the way in which people interpret missing information may be influenced by the belief in other people's self-interest. If this is indeed true, people would systematically underestimate their partners' cooperation. Because beliefs can influence perceived cooperation only when incomplete information is present, the mere incompleteness of information can reduce cooperation.

Assuming that perceived cooperation is influenced by the partner's actual cooperation and pre-existing beliefs (e.g., self-interest), incompleteness of information might have somewhat different effects on those who behave in a generous vs. stingy manner. Given the assumption that people tend to rely on beliefs of people's self-interest, the observation of generous behavior is more conflicting with the observer's *a priori* beliefs than the observation of stingy behavior. With increasing incompleteness of information, people should become more doubtful of another's generosity than another's selfishness. People might fill in the blanks (i.e., the lacking information) with self-interest, and people need more instances of generous behaviors to believe that the other is indeed a generous person than stingy behaviors to believe that the other is indeed a stingy person. Thus, whereas all kinds of behaviors can be communicated under complete information, generous behaviors might be more difficult to communicate under incompleteness of information.

### **Perceived Transformations Under Incompleteness of Information**

The way in which people perceive their partner's cooperation under incompleteness of information may quite directly influence perceived transformations. People tend to make dispositional attributions (e.g., a nice person) on other people's behaviors (e.g.,

cooperative behavior) in a too straightforward manner, while largely neglecting situational factors (see fundamental attribution error; Ross 1977; correspondence bias; Jones, 1990). If people perceive that their partner cooperates less under incompleteness of information, they may make inferences that the partner is generally less cooperative (i.e., perceive more self-interested transformations than is warranted). Such dispositional inferences may be particularly harmful, because they may reduce cooperation in future interactions. If the partner is perceived as less cooperative, the willingness to cooperate may drop. Equally important, the partner who is initially judged as noncooperative may be interpreted as such—especially under incompleteness of information—even if the partner started cooperating more.

**Table 1.6: The prisoner’s dilemma with asymmetric egalitarianism transformations. Numbers after the arrow sign are transformed outcomes.**

		Player 2 = $0.8S + 0.2E$	
		Cooperate	Defect
Player 1 = $0.5S + 0.5E$	Cooperate	4 / 4	6 → 5 / 0
	Defect	0 / 6 → 3	2 / 2

Misperception of transformational information can have a crucial impact on social interactions. Table 1.6 describes the prisoner’s dilemma game with egalitarianism transformations, similar to Table 1.3. Both players engage in strong egalitarianism transformations (e.g.,  $0.5S + 0.5E$ ), which would suggest that both players cooperate—given that they know that their partners have these transformations. Table 1.6 is drawn from Player 1’s perspective. Player 1 knows the transformation for the self, but underestimates the egalitarianism transformation for Player 2 (e.g.,  $0.8S + 0.2E$ ).

Mutual cooperation would be the preferred option for Player 1, but at the same time, Player 1 can reasonably—albeit erroneously—expect Player 2 to defect (i.e., defection yields better perceived outcomes for Player 2). Now, given that Player 2 is expected to defect, Player 1 should also defect, because it yields better transformed outcomes than cooperation. This initial misperception and defection may have a long-term impact on mutual cooperation. If Player 2 correctly perceived Player 1’s high level of egalitarianism, Player 2 would cooperate in the first round. But Player 2 would

probably use tit-for-tat and retaliate Player 1's defection in the second round—a cycle that may lead to mutual defection even though both players' initial preferences were cooperative.

## **The Present Dissertation**

In the next four chapters, I will present nine experiments that aim to test four key premises: First, people systematically underestimate others' unselfish motives and attribute too much self-interest to others' imagined behavior (Experiments 2.1 & 2.2) as well as others' overt behavior with incomplete information (Experiments 3.1, 3.2, & 3.3). Second, the more incompleteness of information is present the more self-interest beliefs reduce cooperation (Experiments 4.1, 4.2, 5.1, & 5.2). Third, the detrimental effects of incomplete information are more pronounced for generous rather than stingy partners (Experiment 5.1 & 5.2). And finally fourth, incompleteness of information influences perceived transformations (Experiment 5.1 & 5.2). Each chapter represents an independent research article that has been published or is under review for publication. The following overview aims to convey the key contributions of each chapter.

The first empirical chapter (Chapter 2) examines motives that people attribute to other people, and compares these motives to those that people display in their own social behavior. Using a behavioral paradigm that yields relative weights for self-interest, altruism, and egalitarianism motives, Experiment 2.1 reveals that people think that egalitarianism has a smaller impact on other people's social decisions than it has on people's own social decisions. Using an evaluation paradigm in which people rate other's outcome allocations, Experiment 2.2 reveals that people expect other people to rate equal allocations as less positive than people themselves rate equal allocations. Hence, Chapter 2 indicates that in the absence of any information, people systematically underestimate the role of egalitarianism in others' social behavior.

The second empirical chapter (Chapter 3) examines interpretations that people make on others' overt behaviors when they lack some pieces of information. Using a new method—the dice-rolling paradigm—Experiments 3.1 and 3.3 provide evidence that people “fill in the blanks” with self-interest. That is, when people are given only partial information about their partner's behavior, they tend to use their self-interest beliefs to fill in the part of information that is not given. As a result, people tend to overestimate the role of self-interest not only in the imagined behavior of others (see Chapter 2), but overt behavior of others is also filtered and attributed through self-interest beliefs. Additionally, Experiment 3.2 compares people's actual behavior and predictions regarding other people's behaviors, and shows that people expect more self-interest from other people than people exhibit themselves. Hence, Chapter 3

indicates that behavior with incomplete information is filtered through the belief in other people's self-interest.

The third empirical chapter (Chapter 4) examines the influence of incomplete information on cooperation in dyadic interactions. Building on previous findings showing that people hold self-interest beliefs about other people (Chapter 2), and that such beliefs influence behavioral attributions (Chapter 3), Experiments 4.1 and 4.2 test the idea that incompleteness of information reduce cooperation. Using a new method—the coin paradigm—the results revealed that incompleteness of information undermines both expectations about another person's cooperation as well as one's own cooperation. Moreover, the detrimental effects of incompleteness of information on cooperation were mediated by expectations of other's cooperation. Hence, Chapter 4 indicates that the belief in self-interest serves to fill in the blanks when information is incomplete, which undermines expectations of other's cooperation as well as one's own cooperative behavior.

The fourth empirical chapter (Chapter 5) examines how generous versus stingy behaviors can be communicated under incompleteness of information. Incompleteness of information indeed undermines cooperation (Chapter 4), but it is not clear whether this tendency still holds when partners differ in their level of generosity versus stinginess. Experiments 5.1 and 5.2 reveal that the detrimental effects of incomplete information are more pronounced for generous than stingy partners. Second, the chapter examines dispositional attributions (i.e., perceived transformations) that people make under incompleteness of information. Both experiments reveal that under incompleteness of information, people judge the partner as less benign—the effect that is more pronounced for generous partners. Finally, the analysis suggests that such beliefs might help account for these effects on cooperation. Hence, Chapter 5 indicates that the more generosity one seeks to communicate, the more incompleteness of information undermines cooperation and perceptions regarding benign intentions of the partner.



