Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

The Sources of the Communication Gap

Simin He, Theo Offerman, Jeroen van de Ven

informs
http://pubsonline.informs.org/journal/mnsc/
MANAGEMENT SCIENCE
Vol. 63, No. 9, September 2017, pp. 2832–2846
ISSN 0025-1909 (print), ISSN 1526-5501 (online)

The Sources of the Communication Gap

Simin He,[a] Theo Offerman,[b,c] Jeroen van de Ven[d,c]

[a] School of Economics, Shanghai University of Finance and Economics, 200433 Shanghai, China; [b] Center for Research in Experimental Economics and Political Decision Making, University of Amsterdam, 1001 NJ Amsterdam, Netherlands; [c] Tinbergen Institute, 1082 MS Amsterdam, Netherlands; [d] Amsterdam Center for Law and Economics, University of Amsterdam, 1001 NJ Amsterdam, Netherlands

Contact: he.simin@mail.shufe.edu.cn (SH); t.j.s.offerman@uva.nl (TO); j.vandeven@uva.nl (JvdV)

Received: May 28, 2015
Revised: October 27, 2015; February 29, 2016
Accepted: March 4, 2016
Published Online in Articles in Advance: August 26, 2016

https://doi.org/10.1287/mnsc.2016.2518

Abstract. Face-to-face communication drastically increases cooperation rates in social dilemmas. We test which factors are the most important drivers of this communication gap. We distinguish three main categories. First, communication may decrease social distance. Second, communication may enable subjects to assess their opponent's cooperativeness ("type detection") and condition their own action on that information. Third, communication allows subjects to make promises, which create commitment for subjects who do not want to break a promise. We find that communication increases cooperation very substantially. In our experiment, we find that commitment value is an important factor, but the largest part of the increase can be attributed to type detection. We do not find evidence that social distance plays a role.

History: Accepted by Uri Gneezy, operations management.

Funding: Financial support from the Research Priority Area Behavioral Economics [Grant 201212161012] of the University of Amsterdam is gratefully acknowledged.

Supplemental Material: Data and the online appendix are available at https://doi.org/10.1287/mnsc.2016.2518.

Keywords: communication • cooperation • prisoner's dilemma • social distance • type detection • commitment

1. Introduction

Many people highly value face-to-face interactions. A survey among Forbes subscribers showed that the vast majority considers face-to-face meetings to be essential for negotiating contracts, closing deals, and building long-term relationships (Forbes Insight 2009). Many businessmen are willing to travel large distances to meet clients, thereby incurring considerable costs. They may be right to do so. Evidence from controlled laboratory studies shows that, compared to anonymous interactions, face-to-face communication has a major influence on people's behavior. Most notably, face-to-face communication drastically increases the cooperation rate in social dilemma games.[1] On average, the option to meet and communicate with one another increases the cooperation rate by 40 percentage points (Sally 1995, Balliet 2009). We will refer to this as the "communication gap."

There are several reasons for why communication may have such a large impact. One possibility is that face-to-face interaction reduces the social distance between people, by making the other person identifiable (e.g., Bohnet and Frey 1999, Hoffman et al. 1996). It may be harder to free ride if it has been clarified who will be harmed. Another reason is that the content of communication can induce a commitment to cooperate. In particular, making a promise can create a commitment value to cooperate for people who are averse to deceiving others (e.g., Ellingsen and Johannesson 2004b, Charness and Dufwenberg 2006, Vanberg 2008).

The existing literature has mostly paid attention to social identification and commitment as explanations for the communication gap (e.g., Kerr and Kaufman-Gilliland 1994). There is, however, an alternative explanation, which may be equally, if not more, important: type detection. Face-to-face interactions can increase cooperation because it enables people to distinguish cooperative people from noncooperative people. People often pay attention to the appearance and body language of someone else, looking for verbal and nonverbal cues that may reveal the intentions of the other person (Eckel and Petrie 2011).[2] Detecting others' types is especially important for conditional cooperators. They are willing to cooperate provided that they perceive the other to be cooperative as well. Interacting with an anonymous person may make them reluctant to cooperate. Since a very large fraction of people are conditionally cooperative (Fischbacher et al. 2001), it is plausible that type detection is responsible for a substantial part of the communication gap.

The contribution of this paper is to decompose the communication gap, estimating the importance of each of the above three factors in a single design: social identification, type detection, and commitment value.[3] We set up a laboratory experiment in which we let subjects play a social dilemma game. In the baseline treatment, subjects play this game without meeting each other before they make their decisions. In a second treatment, "Silent," subjects can identify the other before making their decision, but they are not allowed to communicate. A novelty of our experiment is to implement the "Restricted Communication" treatment, in which subjects are given time to interact face-to-face before they make their decisions, but without being allowed to make promises. Finally, in the "Unrestricted Communication" treatment, subjects are given time to freely interact face-to-face before they make their decisions. By comparing Baseline and Silent, we can isolate the effect of social identification.[4] A comparison of Silent and Restricted enables us to measure the effect of type detection. A comparison of Restricted and Unrestricted allows us to estimate the effect of commitment value, as well as any additional type detection based on promises.

In line with the existing literature, we find that face-to-face communication drastically increases the cooperation rate. The individual cooperation rate in Baseline is 21%, while it is 77% in Unrestricted, creating a communication gap of 56 percentage points in our experiment. The cooperation rate in Silent is only slightly higher than in the Baseline, and we therefore do not find that social identification is of any importance in our context. We do find an important role for promises, as the cooperation rate in Restricted is 43%, well below the cooperation rate in Unrestricted. Without any restrictions on the communication contents, a majority of subjects makes a promise to cooperate. Most people keep their promise, making it a reliable signal of cooperation. Thus, besides creating a commitment value, promises also facilitate type detection.

Since we also collected participants' beliefs about their opponent's choice, we can further break down the communication gap. If one is willing to make the assumption that beliefs have a causal impact on behavior, then changes in beliefs can be used to estimate the importance of type detection. This is not an innocuous assumption, as the causality may be reversed or there might be an omitted variables problem. With this caveat in mind, we find that type detection is by far the most important driver of the communication gap in our experiment.[5]

We also find evidence that subjects not only *believe* that they are able to predict their partners' decisions in the communication treatments, but are in fact to some extent able to do so. They use several cues that are correlated with behavior, and ignore some other cues that indeed lack predictive power. Interestingly, when we classify subjects as selfish or social on the basis of an independent test, we find that selfish people cooperate much less than social people in Baseline, but this difference is much smaller or even absent in the other treatments.

Existing work mostly focuses on the different factors in isolation. Studies that analyze the content of communication have found that promises can be a very reliable predictor of behavior (Ellingsen and Johannesson 2004b, Charness and Dufwenberg 2006, Belot et al. 2010, Van den Assem et al. 2012). Possible motives for keeping promises include guilt aversion, lying aversion, and shame (e.g., Gneezy 2005, Charness and Dufwenberg 2006, Vanberg 2008, Miettinen and Suetens 2008, Greenberg et al. 2014). Another related strand of literature studies people's ability to predict the behavior of others, i.e., people's ability to detect types. Within the context of social dilemmas, it has been shown that people have some ability to detect cooperators when there is a communication stage (Dawes et al. 1977, Frank et al. 1993b, Brosig 2002, Belot et al. 2012).[6] With the exception of Dawes et al. (1977), those studies do not make a comparison with a treatment without communication, and thus cannot establish the importance of type detection for cooperation rates.[7] The focus of those studies is on people's actual ability to detect types, while the behavior of people depends more on their own perceived ability to detect types. We examine both subjects' actual and perceived ability to detect types.

A few other studies have attempted to disentangle the determinants of the impact of communication. Studies that implemented a treatment in which subjects could only discuss game-irrelevant topics tend to find only a small impact compared to treatments with no communication. For instance, Dawes et al. (1977) find cooperation rates of 35% with restricted communication and 27% with no communication. Communication in the former treatment was restricted to topics unrelated to the game. This does not only exclude making promises, but also other ways of signaling the intention to cooperate, such as discussing fairness norms, and also the intentional exchange of personal information that may be perceived as providing reliable cues, e.g., their study or hobbies. The advantage of our approach is that we only exclude communication that is directly about intentions (see the design section). Frank et al. (1993a) implemented two treatments in which promises were not allowed. The two treatments differed in the length of communication. In this study, it is not clear what counted as a promise and how this was enforced. They compared these two treatments to a treatment with unrestricted communication, in which they explicitly told subjects that they could make a (nonbinding) promise, and found that unrestricted communication enhanced cooperation with 9 and 33 percentage points, respectively. They did not have a treatment without communication.[8] Bohnet and Frey (1999) have no treatment with restricted communication, but they do have treatments with anonymity, mutual identification, and unrestricted communication. They attribute any difference between the mutual identification treatment and the anonymity treatment to a decrease in social distance. Brosig et al. (2003) have treatments with anonymity and mutual identification in a four player public game. They find that visual identification alone has no effect on cooperation. In both Bohnet and Frey (1999) and Brosig et al. (2003), the effect of social identification, if there is any, may be partly driven by reputation concerns. In our experimental design, we control for this potential confound.

In addition, even when the communication is about the game, allowing messages to be free form may be essential. Charness and Dufwenberg (2010) find much less effect of communication when messages are more restricted. In a coordinated resistance game where two responders must jointly challenge the leader to prevent the exploitation of the victim, Cason and Mui (2015) compare a treatment with free-form messages and some treatments with messages containing only intended choices. They find that the possibility of free-form messages is critical for coordinated resistance. It allows people to communicate their social motivations. Brosig et al. (2003) provide further support that rich communication may be a prerequisite for a positive effect on cooperation. In a four player public good game, communication facilitates cooperation only if subjects can visually identify each other while communicating. Similarly, people's (perceived) ability to predict the intentions of others is not confined to face-to-face interactions. People also read cues in written messages (e.g., Charness and Dufwenberg 2006, Chen and Houser 2017) and different types of communication media yield different opportunities to assess other's intentions and consequently result in different cooperation rates (e.g., Brosig et al. 2003).[9]

In addition to the decomposition of the effect of communication, we test some hypotheses that are suggested by the literature. Previous studies that examine the impact of own and opponents' characteristics on cooperation in social dilemmas are somewhat mixed. For instance, both Belot et al. (2012) and Van den Assem et al. (2012) find that females are more likely to be cooperative, while Darai and Grätz (2013) do not find a significant gender effect. Belot et al. (2012) do not find evidence that a subject's own attractiveness or the other's attractiveness is predictive of cooperation, but Darai and Grätz (2013) find that the other's attractiveness increases cooperation in mixed-gender pairs. We are not aware of any other studies that included risk attitudes or social as controls. Another question that received attention in the literature is if selfish and social types differ in their ability to judge the person with whom they play the game. Brosig (2002) hypothesizes that conditional cooperators should be better at identifying the other's willingness to cooperate and finds that cooperative individuals are somewhat more accurate in their beliefs. Like these studies, in our analysis we investigate the roles that risk attitudes, gender, and attractiveness play in people's cooperation decision, and we investigate if conditional cooperators are better at predicting the other's decision.[10]

The rest of the paper is organized as follows. We describe the experimental setup in Section 2. Section 3 reports the results. Finally, Section 4 concludes.

2. Experimental Setup

2.1. Treatments Design

Table 1 presents the monetary payoffs of the specific prisoner's dilemma that we use in the experiment. Each of the two subjects makes a choice between X (cooperate) and Y (defect). The game is played only once. The experiment consists of three parts. Subjects receive the instructions of later parts only after finishing a part. In part 1, they are informed that earnings in the remainder of the experiment are completely independent of the decisions in part 1.

Part 1 consists of a communication phase and a choice making stage. The communication stage is payoff irrelevant. Our four treatments vary in the extent to which they allow subjects to communicate. In all treatments, subjects at some point meet the person with whom they are playing the game. In the Baseline treatment B, a short silent meeting of 10 seconds takes place *after* the subjects have chosen between X and Y. The other treatments allow subjects to communicate in varying degrees *before* they make a choice. In the Silent treatment S, subjects meet in silence for 10 seconds. In the silent meeting, subjects are not allowed to communicate in any way.[11] In the Restricted treatment R, subjects are allowed to talk face-to-face for two minutes. This communication is free form, except for the following restriction: subjects are not allowed to make a statement that would become a lie for any of the two choices. As an example, we inform subjects that they cannot promise to choose X, because that would become a lie if they would then choose Y instead. We also stress that no statement is allowed that *could* become a lie for any of the two choices, even if they are planning not to lie. In the Unrestricted treatment U, they are allowed to communicate with one another for two minutes without any restriction.

After subjects are informed of the specific game they are going to play, but before they have met the person with whom they are playing the game, subjects

Table 1. Payoff Matrix

	Other's decision	
	X	Y
Your decision		
X	8, 8	0, 12
Y	12, 0	4, 4

are asked to predict the likelihood that the other will choose option X, on a scale from 0 to 100 ("beliefs before").[12] Subjects are again asked to predict the choice of the other after they have met the other person and have made their decision ("beliefs after"). We chose not to incentivize these beliefs, because we felt that subjects would be intrinsically motivated to take the task seriously in such a short experiment. By incentivizing beliefs, we risked that subjects would hedge across the payoffs of the different tasks. For a discussion of the circumstances under which incentivized beliefs are desirable, see Schlag et al. (2014).

We use the subjects' beliefs to study their ability to detect the type of the other player. In the experiment, we asked to predict the other's behavior instead of the other's type, because we considered this question to be the clearest. The prediction of other's behavior is based on a combined judgment of the other player's type and of the other player's belief. Although it would have been interesting to retrieve information on these concepts separately, we refrained from doing it because we feared that this would become too subtle.

A comparison of the Baseline and Silent treatments allows us to assess the effect of social identification, controlling for potential reputation or image concerns that our subjects may experience. That is, in the Baseline treatment, subjects are aware that they will meet the other subject, just like in the Silent treatment. The comparison of the Silent and Restricted treatments enables us to measure a potential additional effect of type detection and the comparison of the Restricted and Unrestricted treatments allows us to judge a further effect of type detection, as well as an incremental effect of commitment value. Table 2 provides a summary of treatments.

Part 2 consists of two tasks. First, a version of the social value orientation test (e.g., Offerman et al. 1996) is conducted to acquire a measure of a subject's social preferences. We use the standard decomposed game method as developed by Griesinger and Livingston (1973) and Liebrand (1984). In this method, each subject makes 24 choices. In each of these choices, they choose between two "own-other" payoff vectors. Each of these payoff vectors assigns a certain amount of money to the subject and another amount to another subject. The payoff vectors are located at 24 equally spaced points around a circle when mapped in a two-dimensional own-other payoff space. Based on the choices, a subject is classified as aggressive, competitive, individualistic, cooperative, or altruistic, following (e.g., Griesinger and Livingston 1973, Liebrand 1984). Because some categories only contain a few observations, we group all subjects into two groups: *selfish* (aggressive, competitive, and individualistic subjects) and *social* (cooperative and altruistic subjects).

Second, risk attitudes are elicited using a similar task as Eckel and Grossman (2008), adding one option to capture risk-seeking behavior. In this method, a subject makes a choice among six 50–50 gambles that vary in the degree of risk and expected value. Gamble 1 is optimal for risk-seeking subjects, gamble 2 for risk-neutral subjects, and gambles 3–6 for subjects who are characterized by enhancing levels of risk aversion.

In part 3 of the experiment a questionnaire is administered that includes some questions on background information. Only at the very end of the experiment, the outcome of the prisoner's dilemma game is revealed. One of the two tasks of part 2 is randomly selected for payment, and this payment is added to the payment of part 1.

Table 2. Summary of Treatments

Treatment (label)	Timing of choice X, Y	Length of meeting (s)	Communication restrictions	No. of subjects
Baseline (B)	Before meeting	10	Silent	56
Silent (S)	After meeting	10	Silent	98
Restricted (R)	After meeting	120	No promises	100
Unrestricted (U)	After meeting	120	None	80

Notes. In Restricted, any statement that would be a lie for any of the two choices was not allowed. The number in the last column is the total number of participated subjects before excluding subjects from the analysis (see Section 2.4).

2.2. Procedures

The experiment was conducted at the University of Amsterdam. Dutch subjects were recruited from the Center for Research in Experimental Economics and Political Decision Making database. Subjects communicated in their native language. We ran two to four sessions each day, and the experiment series lasted 12 days in total. Treatments were randomized at the session level. Each subject participated in one treatment only.

Upon arrival, each subject was directed to one of two separate rooms according to their random online recruitment assignment. Before the experiment started, subjects were informed that part of the experiment would be recorded on video, and that they could participate if they gave their consent or leave otherwise. The experiment only started when the two rooms had the same number of subjects. Each subject was randomly paired with one subject from the other room. Depending on how many people showed up, the number of subjects per session varied between 4 and 12.[13] In each room, an experimenter read aloud the main instructions. The instructions are provided as part of the online supplementary materials.

In the communication stage, subjects were called one by one into another room where they met the person with whom they were matched. They did not have the possibility to communicate while walking to the meeting room. The subjects were seated by the experimenter and were then left to communicate by themselves. An experimenter was always present in the L-shaped meeting room, but remained out of sight during the communication stage. Afterward, they returned without any further communication one by one to the rooms where they originally came from. An experimenter was present when subjects returned to their rooms.

At the payment stage, each subject was paid in private in the meeting room. Subjects left the room one by one. The average payoff was €14.60, including a fixed show up fee of €6. The experiment took between 30 and 60 minutes and was conducted with paper and pencil.

2.3. Coding of Variables

Four research assistants independently coded the recorded conversations on several dimensions. Online Appendix C contains the list of the variables that were coded. Online Appendix B provides the list of variables that were used in the analysis. To code whether or not a subject made a promise to cooperate, we instructed the coders to use the same definition as was used in the instructions for the subjects (any statement that would be a lie for some choice is classified as a promise). Even though our coders tend to agree on their ratings, there are some inconsistencies. In 41 out of 74 cases (55%) all four coders agreed. In another 30 cases (38%), three out of four coders agreed. In 5 cases (7%) the coders were divided, with two out of four coders classifying the subjects' statement as a promise.[14] We classified promises into three categories: no promise (if at most one coder classified the statement as a promise), *weak promise* (if the coders were divided), and *strong promise* (if at least three coders classified the statement as a promise).

2.4. Observations and Descriptive Data

In total, we ran 38 sessions with 334 subjects (about the maximum number of Dutch speaking subjects that could be recruited). Subjects' choices are independent on the individual level in the Baseline treatment and on the pair level in the other treatments. Therefore, we allocated relatively fewer subjects to the Baseline treatment to balance the independent observations per treatment.

In the analysis, we exclude subjects who violated the instructions. In the Silent treatment, one pair of subjects talked to each other, and eight other pairs communicated by means of nonverbal signs (all eight cases happened before we revised the instructions; see Endnote 10). Another three pairs were excluded because they exchanged their identity cards, thereby creating an external commitment device. Regarding the Restricted treatment, we dropped 13 pairs who self-reported to have made promises in the communication phase. These self-reported results were consistent with independent coding results, suggesting that no more pairs violated the instructions.[15] Finally, we excluded four more pairs because they reported to be friends.[16]

After excluding these subjects, we are left with 276 subjects. Some characteristics of these subjects are listed in Table 3. The Kruskal–Wallis test results show that the randomization procedure was successful on these characteristics.

Table 3. Descriptive Statistics

	Treatment				
	B	S	R	U	p-value
Female, %	50.0 (6.74)	54.1 (5.83)	37.5 (5.75)	43.2 (5.80)	0.204
Age (years)	22.3 (0.408)	21.7 (0.263)	21.8 (0.366)	23.1 (0.575)	0.348
Attractiveness (1–7)	4.2 (0.163)	4.3 (0.122)	4.0 (0.148)	4.1 (0.138)	0.661
Risk aversion (1–6)	3.5 (0.202)	3.2 (0.169)	3.4 (0.186)	3.0 (0.165)	0.267
Social types[a] (%)	26.8 (5.97)	18.9 (4.58)	27.8 (5.32)	31.1 (5.42)	0.386
N	56	74	72	74	

Notes. Standard errors in parentheses. In the last column, p-values refer to Kruskal–Wallis tests of equality-of-populations between treatments. The number of subjects refers to the number after excluding some subjects (see Section 2.4).
[a]Social types according to the social value orientation test.

3. Results

In our analysis, cooperation rates and beliefs play an important role. We define the cooperation rate as the fraction of subjects who choose strategy X. We distinguish beliefs according to the moment at which they were elicited: beliefs elicited before the meeting are referred to as beliefs before and beliefs elicited after the meeting as beliefs after.

Table 4 lists the cooperation rates and the beliefs in each treatment. Consistent with the existing evidence, adding free format face-to-face communication increases the cooperation rate substantially, in our case from 0.21 to 0.77.[17] To understand the sources of this increase, we now take a close look at the role of social identification, commitment value, and type detection, respectively.

3.1. Social Identification

Existing studies do not tease out the effects of social identification from reputation effects. If there is a chance of future interaction (outside the experiment), subjects may behave nicer to the identifiable other to protect their own reputation. We allow our subjects to identify each other in both the Baseline and the Silent treatments, keeping reputation concerns constant. These two treatments only differ in the timing

Table 4. Cooperation Rates and Beliefs

	Treatment			
	B	S	R	U
	Choices			
Cooperation rate (fraction choosing X)	0.21	0.24	0.43	0.77
Test equal to B		(0.812)	(0.020)	(<0.001)
Test equal to S			(0.029)	(<0.001)
Test equal to R				(<0.001)
Coordination rate (fractions)				
Both subjects in pair cooperate (X/X)	0.04	0.08	0.22	0.68
Both subjects in pair defect (Y/Y)	0.61	0.60	0.36	0.14
Expected coordination if choices are independent	0.67	0.64	0.51	0.65
Test actual = Expected coordination (p-value)	(0.776)	(0.399)	(0.238)	(0.003)
	Beliefs			
Beliefs before (1–100)	37.5	40.4	44.5	51.9
Test equal to B		(0.486)	(0.168)	(<0.001)
Test equal to S			(0.391)	(0.001)
Test equal to R				(0.059)
Beliefs after (1–100)	34.4	37.1	55.0	70.2
Test equal to B		(0.683)	(0.001)	(<0.001)
Test equal to S			(0.002)	(<0.001)
Test equal to R				(0.008)

Notes. Statistical tests report p-values of two-sided Mann–Whitney tests (cooperation rates and beliefs) or Chi-square tests (coordination rates). For cooperation rates and beliefs after, the independent unit of observation is the mean over a pair of subjects.

of the meeting—whether it is before or after choosing in the prisoner's dilemma. A difference in behavior between these two treatments can be attributed to social identification.

We do not find any role for social identification. The cooperation rate in the Silent treatment is not statistically different, and only slightly higher, from that in the Baseline treatment. There are also no systematic differences in reported beliefs between these two treatments. It is still possible that social identification plays some role, insofar as subjects in the Baseline treatment correctly *anticipate* the effects of social identification. Our results show that there is no effect from *experiencing* social identification compared to any effects of anticipated social identification.

This result differs from the findings of Bohnet and Frey (1999). They have documented a difference between a treatment with mutual identification and a treatment with anonymous interactions in a prisoner's dilemma. Their design did not control for reputation effects: subjects in the anonymous treatments never identified each other, even not after making their decisions. Our results suggest that the effect they found might at least partly be due to this reputation effect, and not to social identification.[18] This is consistent with their findings from a dictator game, where they find that solidarity rates increase much more with two-way identification than one-way identification (in which reputation effects are absent).[19] This result is also in line with the findings in the four player public good game of Brosig et al. (2003). Without controlling for reputation, they found no difference in cooperation between a treatment with visual identification and a treatment with anonymity.

3.2. Promises

Communication gives people the opportunity to convey their own intentions and to learn about the intentions of others. Among all kinds of statements to express intentions, promises are particularly powerful. A distinctive feature of promises is that they relate intentions directly to actions. Unlike other statements, promises may create commitment value to the promise maker, and the partner may understand the commitment value and become more trusting as a result. For instance, if the promise maker is lying averse, it will be costly for him to break his promise to cooperate. Of course, promises may also facilitate the process of recognizing types. If others perceive promises as a credible sign of cooperation, even when promises are not credible, they may decide to cooperate as well.

We find clear evidence for the positive effect of promises on the cooperation rate. Table 4 shows that the cooperation rate in the Restricted treatment, when promises are not allowed, is well below the cooperation rate in the Unrestricted treatment. In addition, after communicating, subjects are much more optimistic about the behavior of their partner in the Unrestricted treatment. Interestingly, the positive effect of the possibility to make promises is already anticipated by subjects before communicating. At first glance, the option to make promises seems to account for a substantial part of the communication gap.[20]

In the Unrestricted treatment, subjects eagerly use the possibility to make promises. Only 13 out of 74 subjects make no promises at all. Among the subjects who do make promises, 5 make weak promises and 56 make strong promises. If promises create commitment value, one would expect that subjects who make promises are more likely to cooperate. This turns out to be the case. Among subjects who do not make promises, only three (23%) cooperate. By contrast, those who make promises cooperate much more often: 4 out of 5 (80%) for weak promises and 50 out of 56 (89%) for strong promises. Such large effects are consistent with findings of other studies (e.g., Ellingsen and Johannesson 2004b, Charness and Dufwenberg 2006, Belot et al. 2010, Van den Assem et al. 2012). Of course, these differences should not be interpreted as being necessarily causal; we did not exogenously vary whether or not a promise was made, but only the *option* to make promises.[21]

3.3. Commitment and Type Detection

Identifying or communicating with the other subject can help people to assess the intentions of others. Conditional cooperators may adjust their behavior according to their assessments. To form these assessments, people may rely on cues such as promises and gender, attractiveness, etc.[22] However, since those cues may fail to predict the actual behavior, we make a distinction between perceived and actual cues. If in the communication process conditional cooperators perceive positive cues that the other will cooperate, they may have more confidence in the other and cooperate more. This process may be an important source for the positive effect of communication.

A comparison between Silent and Restricted provides evidence that type detection is important. The difference in cooperation is 19 percentage points. This cannot be attributed to commitment or social identification. Type detection can also play a role in other treatments. Because we collected data on beliefs, we can study the effect of type detection in more detail. We present the analysis in two subsections. In Section 3.3.1, we check whether or not people change their belief on the basis of how they communicated; and if they do, how much it affects their own decision. We contrast the perceived ability to detect types with the actual ability to detect types, and we also investigate how beliefs map into decisions. In Section 3.3.2, we identify which part of the communication gap is due to commitment and which part to type detection. In Section 3.3.3 we show which cues subjects use when they change their beliefs, and we compare them to the cues that actually predict when a partner will cooperate.

In the analysis that follows, we assume that the elicited beliefs have a causal impact on behavior. This is not necessarily the case. There can be omitted variables and the causality may be reversed. We come back to this important point at the end of Section 3.3.2, where we discuss the potential consequences of this.

3.3.1. Changes in Beliefs and Behavior. Table 5 shows that subjects change their beliefs in all treatments. A revision of beliefs after meeting the matched partner suggests that subjects do rely on cues. In the treatments where verbal communication is allowed, the largest changes in beliefs are observed. The average absolute change in beliefs in those treatments is 27 percentage points, compared to 16 and 17 percentage points in the Baseline and Silent treatments, respectively. They are also more likely to change their beliefs with more than 10 percentage points when verbal communication is possible. In the communication treatments, 73% of the subjects change their beliefs with more than 10 percentage points, against 52% in the other treatments. In the final two columns we distinguish between pairs who made a mutual promise and those who did not in the Unrestricted treatment. Although pairs who made

Table 5. Updating of Beliefs

	Treatment				U mutual	U no mutual
	B	S	R	U		
Average absolute change in beliefs	15.5 (2.0)	16.6 (1.9)	27.0 (2.3)	27.4 (2.7)	28.9 (3.18)	20.7 (3.12)
Proportion of subjects with large changes in beliefs (%)	50.0	52.7	75.0	71.6	71.7	71.4

Notes. Change in beliefs is *beliefs after/beliefs before*. Standard errors in parentheses. A large change in beliefs occurs when a subject changes his or her beliefs by more than 10 percentage points in absolute terms. The column "U mutual" reports the changes in beliefs for the 30 pairs who made a mutual promise in U. The column "U no mutual" reports the changes in beliefs for the other 7 pairs.

a mutual promise on average changed their beliefs to a larger extent, the difference is not significant (Mann–Whitney, $p = 0.48$).

If subjects are able to update their beliefs about each other in the right direction, one expects that the choices of their partners are positively correlated when communication is possible.[23] In the first row of Table 6, the correlation between choices is presented for each treatment separately. In the Baseline treatments, subjects make their decisions prior to meeting the other, and indeed their choices are not correlated. In the Silent treatment, the correlation is small and not significantly different from zero. In the Restricted treatment, the correlation is slightly stronger but still not significant. In the Unrestricted treatment, the correlation is sizable and significant.[24]

It is possible that subjects are accurate when they change their beliefs, but nevertheless do not adjust their own decisions correspondingly. A selfish person, for instance, might never cooperate independent of his beliefs about the other. As a result, the correlation between choices may be an underestimate of the true ability to assess the intentions of others. We therefore examine if subjects' beliefs after meeting vary

Table 6. Ability to Detect Types

	Treatment			
	B	S	R	U
Correlation between choices in pairs (p-value, Pearson's corr. coefficient)	−0.06 (0.657)	0.12 (0.312)	0.15 (0.208)	0.47 (<0.001)
Accuracy of beliefs after meeting (α) (p-value, Mann–Whitney test)	5.90 (0.738)	1.75 (0.865)	16.09 (0.066)	36.54 (<0.001)
Accuracy of beliefs after meeting (β) (p-value, Binomial test)	3.50 (0.662)	4.23 (0.537)	10.60 (0.104)	18.65 (<0.001)

Notes. See the main text for definitions of α and β. All statistical test are based on two-sided tests.

systematically with the actual choice of the other. Our measure of the accuracy of beliefs is given by

$$\alpha = \bar{p}^a(c) - \bar{p}^a(d),$$

where $\bar{p}^a(c)$ is the average beliefs about others who cooperate, and $\bar{p}^a(d)$ is the average beliefs about others who defect. The measure α varies between −100 and 100, where 100 would reflect a perfect ability to predict the other's choice, 0 would reflect random guessing, and −100 would reflect that people are completely wrong in telling the other's choice.

Values of α are presented in Table 6. We find evidence that subjects' beliefs have predictive value about the other's decision when they can communicate (Restricted and Unrestricted treatments), but not otherwise. In the Baseline and Silent treatments, the values of α are 5.9 and 1.75, respectively, and we cannot reject random guessing (Mann–Whitney test). In the Restricted treatment, α equals 16 and is significantly different from zero. Finally, in the Unrestricted treatment, α equals 36 and is again significantly different from zero.[25]

We think that our accuracy measure is very sensible as long as there is a sufficient number of people who cooperate and a sufficient number of people who defect. If, for instance, there is only one subject in a sample who defects, then the measure will be very sensitive to the report made by the person who was matched with this subject However, this problem did not materialize in our sample. In each of our treatments, at least 12 subjects defect or cooperate.

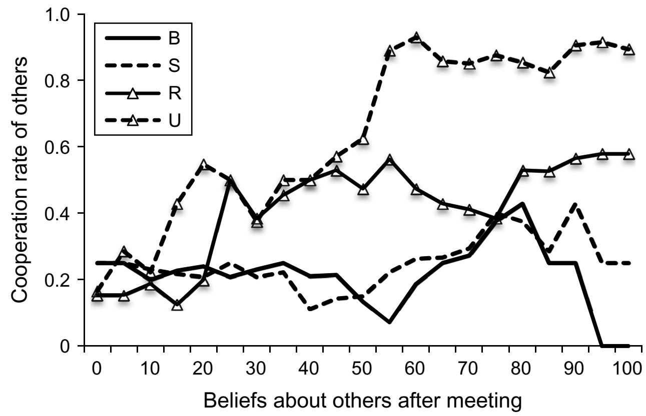
To investigate the robustness of the result, we include an additional measure of accuracy (β), which is also used by, e.g., Brosig (2002), Dawes et al. (1977), and Frank et al. (1993b). To construct this measure, we classify beliefs as correct if they are above 50% and the other cooperated, or if they are below 50% and the other defected. We then compare the fraction of correct beliefs ($\hat{p}$) to a benchmark of random guessing ($\hat{q}$), determined as $\hat{q} = pq + (1-p)(1-q)$, where p is the fraction of times that cooperation is predicted (i.e., beliefs above 50%) and q is the actual fraction of cooperators. The measure of accuracy is then

$$\beta = 100(\hat{p} - \hat{q}).$$

We use a binomial test to test if β differs significantly from 0. Note that, unlike the other measure α, this alternative measure is sensitive to the fraction of cooperators. In treatments in which the fraction of cooperators differs a lot from 0.5, there is less scope to predict better than the benchmark of random guessing.

The results are shown in Table 6. The results are qualitatively similar to the other measure of accuracy. Subjects' predictions are not distinguishable from random guessing in treatments Baseline and Silent. The accuracy is somewhat higher in Restricted (with $p = 0.104$). Again, only the Unrestricted treatment provides strong results.

Figure 1. Average Cooperation Rate of Others by Beliefs After Meeting



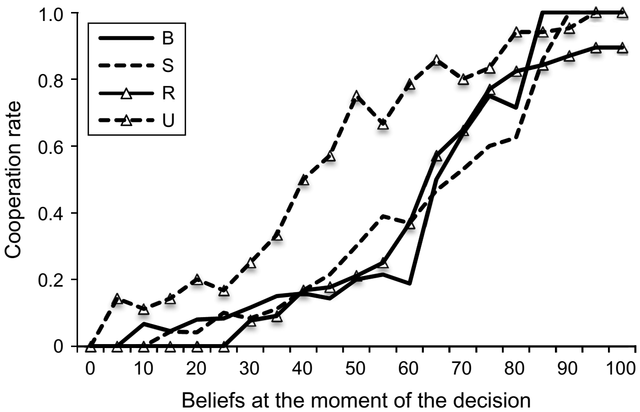
Note. The plotted lines are moving averages, where for each belief we compute the uniformly weighted average over all beliefs within a distance of 10 percentage points of that belief.

There is another way to examine the accurateness of the reported beliefs. Figure 1 plots for each belief after meeting the corresponding cooperation rate of the other. The flat trend in the Baseline and Silent treatments provides further support for the claim that subjects are not able to accurately predict the behavior of the other there. There is some increasing trend in the Restricted treatment and a clear increasing trend in the Unrestricted treatment, illustrating that subjects predict better in these two treatments.

3.3.2. Identifying the Effects of Commitment and Type Detection. Type detection can explain part of the communication gap if the resulting changes in beliefs affect the propensity to cooperate. To estimate this effect, we construct for each treatment a "reaction function" that maps subjects' beliefs about the likelihood that the paired subject will cooperate into (population average) cooperation rates. Figure 2 plots the smoothed reaction functions (see Online Appendix A for a detailed description of the smoothing). In each treatment, the reaction function is increasing, reflecting that subjects are conditional cooperators, and that they are more inclined to cooperate when they are more optimistic that the other will cooperate. If type detection makes subjects on average more optimistic, this results in higher cooperation levels. Additionally, with a convex reaction function, cooperation increases when type detection results in more dispersed beliefs, even when on average beliefs stay the same.[26]

We use the reaction functions to estimate the effects of commitment value and type detection. If communication creates a commitment value, this would manifest itself as an upward shift *of* the reaction function;

Figure 2. Average Cooperation Rate of Subjects as a Function of Beliefs at the Moment of the Decision



Notes. These are beliefs before meeting in Baseline and beliefs after meeting in the other treatments. B, S, R, and U refer to the Baseline, Silent, Restricted, and Unrestricted treatments, respectively. The plotted lines are moving averages, where for each belief we compute the uniformly weighted average over all beliefs within a distance of 10 percentage points of that belief.

given any beliefs about the intentions of the paired subject, subjects become more willing to cooperate. If, on the other hand, people change their beliefs as a consequence of type detection, this would result in changes *along* the reaction function. Figure 2 shows that the reaction function of Unrestricted is indeed above that of the treatments in which subjects were not allowed to make promises, suggesting that commitment value plays some role.

We isolate the effect of commitment value by constructing the counterfactual cooperation rate for subjects in Unrestricted had they not been allowed to make promises. We do this by using those subjects' beliefs and calculating the cooperation rate using the reaction function of the Restricted treatment, in which there is no commitment value. This gives an estimate of the expected cooperation rate if subjects have beliefs as in Unrestricted but would behave as subjects in Restricted. The counterfactual cooperation rate is 62.4, which is 14.6 percentage points below the actual cooperation rate in Unrestricted. Thus, we attribute 14.6 percentage points of the communication gap to a commitment value.

We attribute the remaining part of the communication gap (41.0) to type detection, i.e., changes in beliefs. Note that some of the changes in beliefs are indirectly caused by the commitment value if promises are perceived as a reliable signal of cooperation. We can thus distinguish between (i) type detection due to a commitment value and (ii) type detection due to other factors. Since the cooperation rate in Restricted is 43.1%, while it would have been 62.4% if subjects had beliefs as in Unrestricted, the difference in beliefs between the two treatments yields an estimated difference in cooperation rates of 19.3 percentage points. This means that

Table 7. Decomposition of the Communication Gap

Commitment value	14.6 (26.3%)
Type detection	41.0 (73.7%)
Due to promises	19.3 (34.7%)
Due to other factors	21.7 (39.0%)
Total communication gap	55.6

19.3 of the 41.0 percentage points can be attributed to type detection as a result of a commitment value, and the remaining 21.7 to type detection as a result of other factors. This is summarized in Table 7. Of course, we do not want to give the suggestion that these exact numbers will generalize to other settings. The main point of these calculations is to illustrate that in our experiment type, detection is the most important factor.

If our decomposition technique is correct, one should expect to find no commitment value among participants that did not make a promise in Unrestricted. This is indeed the case. If we compute the counterfactual cooperation rate for this subsample, we find that it is very close to their actual cooperation rate (23.3 versus 23.1, respectively). It is also comforting to observe that (at least visually), the reaction functions of Baseline, Silent, and Restricted are very similar, consistent with the idea that a commitment value is absent in those treatments. In Online Appendix A, we report some more robustness tests, and find that our estimates are not sensitive to alternative assumptions.

Before proceeding, we emphasize that our decomposition analysis rests on the presumption that we can give a causal interpretation to the estimated reaction function, where beliefs determine actions. As mentioned before, this is not necessarily the case. There can be omitted variables and the causality may be reversed. According to the consensus effect, people's beliefs are biased toward their own behavior (e.g., Blanco et al. 2014). Additionally, subjects who choose not to cooperate may report pessimistic beliefs to justify their behavior toward the experimenters or even toward themselves. Such behavior would result in an underestimation of the commitment effect: the commitment effect induces people to cooperate, and a justification of behavior would result in more optimistic beliefs, which we would then incorrectly attribute to type detection.[27] We cannot give a precise quantitative estimate of any bias, but some other studies have examined the causal role of beliefs on actions. Blanco et al. (2014) study the consensus effect in a sequential prisoner's dilemma. They find that the best-response function based on subjects' reported beliefs differs from the best-response function when subjects are given objective probabilities. They attribute the difference to the consensus effect. However, although the best-response function based on objective probabilities is substantially flatter around the sample mean, the average slope is very similar, so that the average bias may be small. Moreover, even with the objective probabilities, there is a strong correlation between beliefs and actions. Within the context of a trust game, Costa-Gomes et al. (2014) use an instrumental variable approach to detect a causal link from beliefs to actions. They find strong evidence of a causal effect, and their results indicate that there is no strong or significant omitted-variable problem. Of course, we cannot directly apply those results to our setting, but two features of our data suggest that also in our setting there is some link between behavior and beliefs. First, the fact that the measure of accuracy (α) is significantly different from zero in some of the treatments supports the idea that beliefs have a causal impact on behavior: we would not find a positive correlation between beliefs and the other's decision if beliefs would only reflect own intended behavior (reverse causality), or if some omitted variables would cause beliefs to differ between subjects. Second, the upward shift in the response function is consistent with the idea that commitment value plays a role in Unrestricted, but not in the other treatments. If subjects were just biasing their beliefs toward their actions, we would not expect to see a shift of the response function, but merely a movement along the response function.

3.3.3. Perceived and Actual Cues. To find out the role of *perceived* cues, we investigate the influence of observable cues on the beliefs that subjects report after the meeting. If subjects think that certain cues signal that their partners are going to cooperate, more optimistic beliefs are expected when such cues are present. We use ordinary least squares (OLS) regressions to determine the effects of some observable cues.[28] The dependent variable is the belief reported after the meeting. Table 8 shows the independent variables together with the regression results.

In Table 8, the control variables include observable characteristics such as gender, age, and attractiveness of the opponents. These features are easily observed when subjects meet. The regression results suggest that our subjects do not rely on any of these characteristics to update their beliefs. Another control variable is the promise that the partner may make, which only takes place in the Unrestricted treatment. Since promises are highly correlated within a pair, we cannot distinguish if a subject's own promise matters or that of the partner. We therefore construct the dummy variable *Promise* that is 1 if both subjects make a promise and 0 otherwise. We combine weak and strong promises, as their effect is nearly identical. A promise has a large and significant positive effect on the reported belief. We also include the level of risk aversion and the variable social (based on the social value orientation test) as controls. Those are not directly observed by the subjects, but conceivably people can form an impression about someone's risk attitude or social value orientation. We

Table 8. Beliefs After and Cues

	Treatement				
	(1) All	(2) B	(3) S	(4) R	(5) U
	Opponent's characteristics				
Female	1.213	−2.123	10.598	3.425	−3.660
	(4.163)	(8.463)	(6.528)	(10.976)	(6.039)
Attractiveness (1–7)	0.715	1.768	−0.374	0.401	0.456
	(1.632)	(2.944)	(3.559)	(4.000)	(3.036)
Risk aversion (1–6)	−1.363	−2.698	−2.442	−0.809	1.390
	(1.399)	(3.123)	(3.385)	(3.128)	(1.811)
Social types[a]	−0.656	1.386	14.749	−0.376	−5.810
	(4.252)	(10.532)	(9.644)	(6.859)	(5.961)
Cooperates	25.701***	4.108	0.188	17.965*	20.013*
	(4.660)	(13.908)	(10.720)	(9.750)	(11.450)
Promise					34.297***
					(12.462)
Constant	40.822***	36.364**	38.418**	49.066**	24.006
	(8.135)	(17.481)	(15.218)	(18.204)	(18.060)
Observations	262	54	69	66	73
R-squared	0.156	0.046	0.065	0.084	0.433

Notes. OLS estimates. Dependent variable: Beliefs after (0–100). Robust standard errors in parentheses, clustered at the pair level.
[a]Social types according to the social value orientation test.
$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

do not find that they influence beliefs though. Finally, we include the actual cooperation decision of the partner as a control variable. It has a positive effect on beliefs in the communication treatments, which suggests that our subjects in these treatments use some reliable cues that are not captured by any of the other control variables.

To determine the *actual* cues that influence the choice to cooperate, we use OLS regressions in which the dependent variable is the choice to cooperate. In the regression, the choice equals 0 when the subject defects and 100 when the subject cooperates. The independent variables consist of both own characteristics and the partner's characteristics.[29] We do not include the risk aversion and social type of the opponent, as in the previous analysis we did not find evidence that these variables affect beliefs.

Table 9 shows the regression results. Some own observable characteristics affect the decisions. Women tend to cooperate more than men when subjects do not meet before the decision. The gender coefficient in Restricted and Unrestricted is still fairly large, but ceases to have a significant effect. In Silent, the gender coefficient is even negative, though also not significant. We find some surprising results for the variable social. Subjects who are classified as social according to the social value orientation test are more likely to be cooperative, but the effect is much larger in the baseline than in the other treatments. It is also interesting to observe that risk-averse people are much less likely to

Table 9. Cooperation and Cues

	Treatement				
	(1) All	(2) B	(3) S	(4) R	(5) U
	Own characteristics				
Female	6.195	23.556*	−13.803	10.834	12.995
	(6.998)	(12.275)	(9.878)	(15.864)	(10.987)
Attractiveness (1–7)	0.415	−1.560	−0.889	3.734	−0.016
	(2.751)	(3.685)	(6.221)	(7.846)	(3.878)
Risk aversion (1–6)	−2.172	−10.942**	4.128	3.224	0.704
	(2.101)	(4.040)	(5.796)	(3.555)	(3.679)
Social types[a]	17.598**	40.935***	18.299	−9.021	13.798*
	(6.983)	(13.917)	(16.687)	(12.821)	(7.448)
	Opponent's characteristics				
Female			0.199	16.387	−6.176
			(12.881)	(14.459)	(10.166)
Attractiveness (1–7)			−0.268	−5.622	2.035
			(6.288)	(6.048)	(4.198)
Cooperates			−3.909	21.185	24.583
			(19.586)	(19.683)	(21.086)
	Joint characteristics				
Promise					44.995***
					(16.056)
Constant	41.169***	43.472**	21.291	22.760	3.485
	(13.782)	(20.283)	(37.796)	(36.348)	(35.918)
Observations	264	56	65	60	72
R-squared	0.028	0.259	0.067	0.124	0.399

Notes. OLS estimates. Dependent variable: Cooperates (0 or 100). Robust standard errors in parentheses, clustered at the pair level.
[a]Social types according to the social value orientation test.
*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

cooperate than others in the baseline treatment, but the effect disappears in the other treatments. Possibly the communication process diminishes the perception that cooperation is a risky decision.[30]

We again include the variable promise as a joint characteristic (because we cannot distinguish the effect of a subject's own promise from the promise of the partner). A promise is a very reliable predictor of cooperation: when promises are made within a pair, subjects are 45 percentage points more likely to cooperate.

Besides the variables included in Tables 8 and 9, we asked our coders to code the variables listed in Online Appendix C. Several of those variables (roll-call, threat, trust, risk, justice, and caring) were very infrequently used (in less than 10% of the dialogues) and so we do not have enough observations to reliably identify any effects. For some other variables (first promise, discuss structure of the game) the interrater reliability was questionable or poor (see Table B1 in Online Appendix B for the values of Cronbach's alpha). In Tables B2 and B3 of Online Appendix B, we reproduce the regressions of Tables 8 and 9 when we include the variables for which the interrater reliability was acceptable (Cronbach's alpha of 0.7 or higher). Because we have missing observations for some of those variables, and because the sample size is relatively small, we included the variables one by one.[31] In most cases, the inclusion of those variables does not affect the estimated size and significance of the coefficients of the other variables. One exception is the coefficient of female, which becomes larger in some specifications although it never becomes significant and it is not consistent in sign across specifications. Perhaps the most interesting finding is that shaking hands before talking to one another decreases the beliefs that the matched participant will cooperate in Restricted, while shaking hands after talking has a positive effect on cooperation rates in Unrestricted. These results are consistent with the findings of Darai and Grätz (2010). Only one pair of participants shook hands at the end of the dialogue in treatment Restricted, while this happened relatively frequently in treatment Unrestricted. As a matter of fact, in all cases that they shook hands they also made a promise and they always kept their promise. It thus seems that participants shook hands on their promises, thereby "sealing the deal."

A striking result is that while selfish subjects are much less likely than social subjects to cooperate in the anonymous baseline treatment, selfish subjects catch up with social subjects in terms of the cooperation rate when they meet their partner. This cannot be explained by beliefs, as average beliefs are the same for both types in all treatments. This result is further illustrated in Table 10, which reports the correlation between own beliefs and own cooperation rates. Social types tend to behave as conditional cooperators in all treatments, and there is no noticeable trend when the possibilities for communication are enhanced. In contrast, for selfish types the correlation between own beliefs and own behavior is smaller in the Baseline treatment than in the other treatments. Selfish people behave in a more selfish way behind the veil of anonymity, but once the veil is lifted, they start responding to their beliefs like other people. In the Unrestricted treatment, 86% of the selfish and 74% of the social subjects make a weak or strong promise. A promise has remarkable

Table 10. Correlation Between Cooperation Rates and Own Beliefs by Type and Treatment

	Treatement			
	B	S	R	U
Selfish types	0.41	0.62	0.71	0.62
	(0.008)	(<0.001)	(<0.001)	(<0.001)
Social types	0.56	0.67	0.77	0.66
	(0.030)	(0.009)	(<0.001)	(<0.001)

Notes. Spearman rank correlations. *p*-values testing coefficient equals zero in parentheses. Classification of selfish and social types based on social value orientation test.

Table 11. Ability to Detect Types by Own Type

		B	S	R	U
Accuracy of beliefs	Selfish	14.6	2.6	15.3	31.4***
after meeting (α)	Social	−18.3	−1.6	21.8	46.9**
Accuracy of beliefs	Selfish	5.7	5.6	11.3	15.5**
after meeting (β)	Social	−4.2	−2.0	5.6	25.7***
Differences in beliefs	If other defects	−15.9*	−2.9	−1.5	9.6
selfish and social	If other cooperates	17.1	1.3	−7.9	−5.9

Notes. The upper and middle panels present accuracy levels α and β by treatment and type. Types classified based on social value orientation test. The lower panel presents differences in beliefs between the selfish and social types, conditional on the choice of the other. In both panels, Mann–Whitney tests are used to test equality of coefficients to zero.

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

commitment power even for selfish types. They keep their promise in 84% of the cases, while social subjects always keep their promise. One could also imagine that social subjects are more likely to make a promise first, and that selfish subjects follow the example.[32] This is related to the findings of Cason and Mui (2015), who find that potential victims are more likely than potential beneficiaries to start a conversation and to appeal to fairness arguments. We do not find evidence for this in our context, as social and selfish subjects are equally likely to make a promise first (51% of selfish and 47% of social subjects make a promise first), although we emphasize that the interrater reliability of "first promise" is too low to derive reliable conclusions (Cronbach's alpha 0.45). Furthermore, even in pairs where both subjects are classified as selfish, the percentage of promises is very high (85%).

Finally, we investigate if selfish and social types differ in their ability to judge the person with whom they play the game. Table 11 shows that, based on the measure α, selfish subjects predict slightly better than social subjects in the treatments without face-to-face communication. However, this result is reversed in the treatment with unrestricted communication. When face-to-face communication is present and selfish subjects start behaving more as social subjects, social subjects form more accurate beliefs. However, conditional on the choice of the other person, there are no significant differences in the beliefs of the two types at the 5% level. Using the measure β also does not reveal systematic differences between selfish and social types.

4. Conclusion

Communication has a profound effect on how people behave in social dilemmas. The main goal of our paper is to investigate in one setup why communication has such a strong effect on people's behavior in social dilemmas. Previous papers have reported effects of the various factors influencing cooperation in isolation. We find that perceived type detection is the most important factor, while the remainder of the gap can be attributed to a commitment value. We find no evidence for a positive effect of social identification. Promises stand out as a powerful predictor for both the belief that the other will cooperate and the actual decision to cooperate. In fact, the option to make promises not only creates commitment value, but also further facilitates the process of recognizing types. In our study, we find that approximately half of the effect of type detection is due to promises. We expect our results to be robust qualitatively. We would not want to assign too much importance to the precise quantitative effects that we obtain though, as these may vary when the payoffs of the prisoners' dilemma are changed, and they may also change if beliefs are (partly) affected by subjects' own behavior.

What matters for subjects' decision is their *perceived* ability to gauge the type of the subject with whom they are matched, provided that beliefs have a causal impact on behavior. Interestingly, we find that perceived type detection is supported by an ability to *actually* detect types if subjects are allowed to communicate face-to-face. The evidence in favor of actual type detection is strong when subjects are able to make promises. But even when their talk is restricted and they cannot make promises, there is some degree to which they can predict the behavior of their partner (Table 6).

We think that our results are important for understanding managerial processes. Our study not only contributes to understanding why communication works, but it may also help managers to pay attention to cues that actually work (like explicit promises) and to avoid cues that have feeble predictive value.

In this paper, we assume that reputational concerns are constant across the four treatments. An alternative possibility is that treatment effects are partly driven by different strengths of reputational concerns. If a person defects after a restricted conversation she may perceive more damage for her reputation than when she defects after visual identification. Similarly, breaking a promise after face-to-face communication may be perceived to damage a person's own reputation more than to defect after restricted talk. Figure 2 showed that the "response function" is almost identical in the Baseline, Silent, and Restricted treatments. This suggests that the reputational effect does not differ between these treatments. When we compare Restricted and Unrestricted, we do observe a shift in the response function. We expect that the shift is mainly caused by the fact that subjects can make promises in the latter treatment, but we cannot exclude that it is partly due to differential reputational concerns. We leave it to future studies to distinguish between the two.

It is likely that type detection also plays an important role in many other settings, including coordination games, competitions, and bargaining situations.

In all these settings, people will assess the intentions and characteristics of others, and condition their own behavior on these assessments. We see it as an important avenue for future research to study the process of type detection in such contexts.

Acknowledgments

The authors are very grateful to three anonymous referees, the associate editor, and department editor Uri Gneezy for valuable comments and suggestions. The paper benefited from valuable comments by Dan Houser, Ragan Petrie, and seminar and conference participants at Interdisciplinary Center for Economic Science (George Mason University), the 2013 Amsterdam Symposium on Behavioral and Experimental Economics meeting (University of Amsterdam), the First International Meeting on Experimental and Behavioral Social Sciences conference (University of Oxford), the 16th Conference on the Foundations of Utility and Risk meeting (Erasmus University Rotterdam), the 2014 CREED-CEDEX-CBESS meeting (University of Nottingham), the 2014 European Meeting of the Economic Science Association (Charles University in Prague), and the 2015 Morality, Incentives, and Unethical Behavior Conference (University of San Diego). The authors also thank Harrie Beek, Aviva Heijmans, Joris Korenromp, and Jacco van Mourik for helping to run the experiments and code the communication.

Endnotes

[1] Many interactions within and across organizations have the character of a social dilemma. For instance, two employees who are engaged in team production may have an incentive to shirk if individual efforts cannot be monitored. Similarly, different divisions within a firm may to some degree be engaged in a competition with each other, even though this is not in the best interest of the firm as a whole. Or firms may try to find ways to cooperate with other firms instead of competing with them (e.g., on prices), but nevertheless have incentives to undercut their rivals. It is therefore of interest to managers to know how they can encourage cooperation in such situations.

[2] These cues can be very subtle and brief. Ekman (2009) argues that micro expressions often display emotions that people try to conceal, and that spotting such expressions can help to recognize intentions.

[3] There might be other factors that explain the importance of communication in real life. For instance, incurring large travel costs to meet a client may act as a signal of the value that the person attaches to a meeting, and being present forces the other to pay attention to you. De Haan et al. (2015) provide evidence that messages are more credible when senders choose to incur a cost to communicate them, which is understood by receivers who anticipate that costly communication is more informative.

[4] We adopt the definition of social distance used by Bohnet and Frey (1999). Social distance is reduced if the other person has been identified. Note that there may also be an effect of an anticipated reduction in social distance, if subjects know that they will meet each other after deciding. We cannot rule this out, but, as our empirical analysis shows, the effect of social distance is in any case likely to be very modest.

[5] Evidence from the existing literature suggests that there is indeed a strong impact of beliefs on actions. We discuss this issue in Section 3.3.2.

[6] The ability to predict others' behavior above chance levels has now also been established in other games, such as the trust game and bargaining games. See, for instance, Van Leeuwen et al. (2014) for a list of studies.

[7] In psychology, social dilemmas are often studied with groups of four players or more. This makes the study of type recognition more complicated, as subjects will have to take into account the characteristics of all other group members in forming their beliefs.

[8] Other studies have used a restricted communication treatment that was similar to the one of Dawes et al. (1977). These studies find little impact of irrelevant communication on cooperation (Bouas and Komorita 1996, Mulford et al. 2008, Bicchieri et al. 2010, Ismayilov and Potters 2014).

[9] Bicchieri and Lev-On (2007) provide a more exhaustive discussion of the effect of the medium of communication on cooperation in social dilemmas.

[10] There is also a stream of literature that studies communication in a variety of other games, including coordination games, cheap talk games, and hold-up problems (e.g., Brandts et al. 2016; Charness and Dufwenberg 2011; Cooper et al. 1992; Ellingsen and Johannesson 2004a, b, among others). Communication in those studies is typically in the form of messages.

[11] The instructions of the Silent treatment were revised once. In early sessions, some subjects communicated by means of nonverbal signs (like making the sign "X" with their hands, to signal their intentions). After observing this, we adjusted the instructions by including an explicit statement that no type of communication was allowed. In sessions with the revised instructions, only one pair of subjects violated the instructions. We ran three sessions with 30 subjects with the original version of the instructions, and eight sessions with 68 subjects with the revised version. We excluded subject pairs who used nonverbal signs in the first version from our analysis.

[12] In the Baseline treatment, this question was asked before subjects made their choice between X and Y.

[13] There was one session with 4 subjects and five sessions with 6 subjects. All other sessions were conducted with 8 to 12 subjects.

[14] Cronbach's alpha (a measure of interrater reliability) for the variable promise is 0.75, which is usually considered as acceptable.

[15] In two instances, one of the four coders coded the conversation as violating the instruction, whereas the participants themselves did not report a violation. Dropping those observations does not affect our results.

[16] In the questionnaire, subjects were asked to rate their relation with their partner *before* the experiment. One of the options was "I consider the other person to be a friend." If subjects stated to be friends, they were excluded from the analysis. Their payments were unaffected by the answers in the questionnaire.

[17] Our results remain almost the same if we do not exclude the subjects for the various reasons mentioned in Section 2.4. Then, the cooperation rates are 0.21 (B), 0.33 (S), 0.40 (R), and 0.79 (U). If we exclude the subjects in the Silent treatment before we revised the instructions, the cooperation rate in Silent is 0.25.

[18] Since their study is a classroom experiment, future interactions are likely and reputation effects are therefore plausible. Bohnet and Frey (1999) acknowledge this potential drawback of their design (see their footnote 8). It is also possible that some subjects knew each other, and that their treatment difference is caused by friendships.

[19] Bohnet and Frey (1999) do find a substantial effect of one-way identification on solidarity rates if the dictator also learns some background information about the potential recipient, such as their name and major.

[20] Almost all subjects who expressed their intention stated that they would cooperate; only one subject mentioned the intention to defect (according to more than one coder).

[21] An ingenious study by Ismayilov and Potters (2014) suggests the relationship is not causal. In a trust game setting, they exogenously

vary whether a written promise is delivered to the matched subject. Even if subjects know that their message is not delivered, they tend to keep their promise. However, in a control treatment with irrelevant communication, they find similar levels of trustworthiness. They conclude that it is not promises per se that affect trustworthiness. Rather, trustworthy people are more likely to make a promise. By contrast, we do find a difference between Restricted and Unrestricted, suggesting at least some causality from promises on behavior.

[22] We measured attractiveness in the post experimental questionnaire. There, we asked the question, "How would you rate the beauty of the person with whom you were matched on a 1–7 scale (1 is ugly; 7 is beautiful)?"

[23] A priori, this is not clear though. Belot et al. (2010) argue that altruism or efficiency concerns can result in a negative correlation. Since surplus is destroyed when no one cooperates, noncooperative behavior becomes unattractive if an altruistic or efficiency minded subject expects the other to be noncooperative. In an earlier working paper version, they analyze the game in a Bayesian framework with incomplete information about the cooperativeness of other people and show that an increase in a subject's characteristic that makes the subject more likely to cooperate decreases the opponent's equilibrium probability of cooperating.

[24] Using data from game shows, both Belot et al. (2010) and Van den Assem et al. (2012) do not find evidence of a correlation in decisions, despite a free-format communication stage.

[25] We find that same-gender match improves accuracy: α equals 55.2 for the same-gender pairs, and α equals 20.0 for the mixed-gender pairs. In a similar spirit, Coffman and Niehaus (2014) find that similarity between seller and buyer matters. In their paper, homophily facilitates persuasion by the sellers, who gain substantially more when the buyer has the same gender.

[26] Hugh-Jones and Reinstein (2012) also discuss how information about others' types can increase cooperation rates. In their setup, people's types are revealed by their actions instead of the communication process.

[27] We thank an anonymous referee for pointing this out.

[28] We cluster the standard errors at the subject-pair level to account for interdependency within a pair. Estimations based on seemingly unrelated regressions yield similar results.

[29] The other's characteristics are not included in the regressions for the Baseline treatment and the combined treatments. In the Baseline treatment, subjects decide before meeting the person with whom they play the game. Therefore, the subjects are not influenced by the other's characteristics at the moment of decision.

[30] To assess the possibility of multicollinearity, we calculated the *variance inflation factor*. This factor provides an index that measures how much the variance of an estimated regression coefficient is increased because of collinearity. The regressions in Tables 8 and 9 have modest variance inflation factors between 1 and 2. These factors are far lower than 10, which is regarded as the threshold for (severe) multicollinearity.

[31] If three or four of the raters agreed then we coded a variable in accordance with the majority of the raters. If the raters were split, we coded the variable as a missing observation.

[32] We thank a referee for suggesting this.

References

Balliet D (2009) Communication and cooperation in social dilemmas: A meta-analytic review. *J. Conflict Resolution* 54(1):39–57.

Belot M, Bhaskar V, Van de Ven J (2010) Promises and cooperation: Evidence from a TV game show. *J. Econom. Behav. Organ.* 73(3):396–405.

Belot M, Bhaskar V, Van de Ven J (2012) Can observers predict trustworthiness? *Rev. Econom. Statist.* 94(1):246–259.

Bicchieri C, Lev-On A (2007) Computer-mediated communication and cooperation in social dilemmas: An experimental analysis. *Politics, Philosophy Econom.* 6(2):139–168.

Bicchieri C, Lev-on A, Chavez A (2010) The medium or the message? Communication relevance and richness in trust games. *Synthese* 176(1):125–147.

Blanco M, Engelmann D, Koch AK, Normann HT (2014) Preferences and beliefs in a sequential social dilemma: A within-subjects analysis. *Games Econom. Behav.* 87(September): 122–135.

Bohnet I, Frey BS (1999) The sound of silence in prisoner's dilemma and dictator games. *J. Econom. Behav. Organ.* 38(1):43–57.

Bouas KS, Komorita SS (1996) Group discussion and cooperation in social dilemmas. *Personality Soc. Psych. Bull.* 22(11):1144–1150.

Brandts J, Ellman M, Charness G (2016) Let's talk: How communication affects contract design. *J. European Econom. Assoc.* 14(4): 943–974.

Brosig J (2002) Identifying cooperative behavior: Some experimental results in a prisoner's dilemma game. *J. Econom. Behav. Organ.* 47(3):275–290.

Brosig J, Weimann J, Ockenfels A (2003) The effect of communication media on cooperation. *German Econom. Rev.* 4(2):217–241.

Cason TN, Mui VL (2015) Rich communication, social motivations, and coordinated resistance against divide-and-conquer: A laboratory investigation. *Eur. J. Political Econom.* 37(March): 146–159.

Charness G, Dufwenberg M (2006) Promises and partnership. *Econometrica* 74(6):1579–1601.

Charness G, Dufwenberg M (2010) Bare promises: An experiment. *Econom. Lett.* 107(2):281–283.

Charness G, Dufwenberg M (2011) Participation. *Amer. Econom. Rev.* 101(4):1211–1237.

Chen J, Houser D (2017) Promises and lies: Can observers detect deception in written messages. *Experiment. Econom.* 20(2): 396–419.

Coffman L, Niehaus P (2014) Pathways of Persuasion. Mimeo, Ohio State University, Columbus.

Cooper R, DeJong DV, Forsythe R, Ross TW (1992) Communication in coordination games. *Quart. J. Econom.* 107(2):739–771.

Costa-Gomes MA, Huck S, Weizsäcker G (2014) Beliefs and actions in the trust game: Creating instrumental variables to estimate the causal effect. *Games Econom. Behav.* 88(November):298–309.

Darai D, Grätz S (2010) Determinants of successful cooperation in a face-to-face social dilemma. SOI Working Paper 1006, University of Zurich, Zurich.

Darai D, Grätz S (2013) Attraction and cooperative behavior. Working Paper 82, University of Zurich, Zurich.

Dawes RM, McTavish J, Shaklee H (1977) Behavior, communication, and assumptions about other people's behavior in a commons dilemma situation. *J. Personality Soc. Psych.* 35(1):1–11.

de Haan T, Offerman T, Sloof R (2015) Money talks? An experimental investigation of cheap talk and burned money. *Internat. Econom. Rev.* 56(4):1385–1426.

Eckel CC, Grossman P (2008) Forecasting risk attitudes: An experimental study using actual and forecast gamble choices. *J. Econom. Behav. Organ.* 68(1):1–17.

Eckel CC, Petrie R (2011) Face value. *Amer. Econom. Rev.* 101(4): 1497–1513.

Ekman P (2009) *Telling Lies: Clues to Deceit in the Marketplace, Politics, and Marriage*, Revised ed. (WW Norton & Company, New York).

Ellingsen T, Johannesson M (2004a) Is there a hold-up problem? *Scandinavian J. Econom.* 106(3):475–494.

Ellingsen T, Johannesson M (2004b) Promises, threats and fairness. *Econom. J.* 114(495):397–420.

Fischbacher U, Gächter S, Fehr E (2001) Are people conditionally cooperative? Evidence from a public goods experiment. *Econom. Lett.* 71(3):397–404.

Forbes Insight (2009) Business meetings: The case for face-to-face. Report, Forbes Insights, New York, http://www.forbes.com/forbesinsights/Business_Meetings_FaceToFace/.

Frank RH, Gilovich T, Regan DT (1993a) Does studying economics inhibit cooperation? *J. Econom. Perspect.* 7(2):159–171.

Frank RH, Gilovich T, Regan DT (1993b) The evolution of one-shot cooperation: An experiment. *Ethology and Sociobiology* 14(4):247–256.

Gneezy U (2005) Deception: The role of consequences. *Amer. Econom. Rev.* 95(1):384–394.

Greenberg AE, Smeets P, Zhurakhovska L (2014) Promoting truthful communication through ex-post disclosure. Working paper, University of California, Los Angeles, Los Angeles.

Griesinger DW, Livingston JW (1973) Toward a model of interpersonal motivation in experimental games. *Behavioral Sci.* 18(3):173–188.

Hoffman E, McCabe K, Smith VL (1996) Social distance and other-regarding behavior in dictator games. *Amer. Econom. Rev.* 86(3):653–660.

Hugh-Jones D, Reinstein D (2012) Anonymous rituals. *J. Econom. Behav. Organ.* 81(2):478–489.

Ismayilov H, Potters J (2014) Promises as commitments. Working paper, Tilburg University, Tilburg, Netherlands.

Kerr NL, Kaufman-Gilliland CM (1994) Communication, commitment, and cooperation in social dilemma. *J. Personality Soc. Psych.* 66(3):513–529.

Liebrand WB (1984) The effect of social motives, communication and group size on behaviour in an *N*-person multi-stage mixed-motive game. *Eur. J. Soc. Psych.* 14(3):239–264.

Miettinen T, Suetens S (2008) Communication and guilt in a prisoner's dilemma. *J. Conflict Resolution* 52(6):945–960.

Mulford M, Jackson J, Svedsäter H (2008) Encouraging cooperation: Revisiting solidarity and commitment effects in prisoner's dilemma games. *J. Appl. Soc. Psych.* 38(12):2964–2989.

Offerman T, Sonnemans J, Schram A (1996) Value orientations, expectations and voluntary contributions in public goods. *Econom. J.* 106(437):817–845.

Sally D (1995) Conversation and cooperation in social dilemmas: A meta-analysis of experiments from 1958 to 1992. *Rationality Soc.* 7(1):58–92.

Schlag KH, Tremewan J, Van der Weele JJ (2014) A penny for your thoughts: A survey of methods for eliciting beliefs. *Experiment. Econom.* 18(3):457–490.

Vanberg C (2008) Why do people keep their promises? An experimental test of two explanations. *Econometrica* 76(6):1467–1480.

Van den Assem MJ, Van Dolder D, Thaler RH (2012) Split or steal? Cooperative behavior when the stakes are large. *Management Sci.* 58(1):2–20.

Van Leeuwen B, Noussair CN, Offerman TJS, Suetens S, Van Veelen M, Van de Ven J (2014) Predictably angry: Facial cues provide a credible signal of destructive behavior. Working paper, Institute for Advanced Study in Toulouse, Toulouse, France.