Journal of the Economic Science Association (2023) 9:51–65
https://doi.org/10.1007/s40881-023-00129-3

ORIGINAL PAPER

Reciprocity with stochastic loss

Nathan W. Chan[1] · **Leonard Wolk**[2]

Received: 17 January 2022 / Revised: 23 February 2023 / Accepted: 25 February 2023 /
Published online: 8 April 2023

Abstract

We introduce stochastic loss into a gift exchange game to study how information on intentions affects reciprocity. In one treatment, the respondent observes the amount received and whether a loss occurred, so both the consequential outcome and the sender's original intention are known. In the other two treatments, information about whether a loss occurred is hidden, and the respondent is only informed of the amount received (outcome) or the amount initially sent (intention). Using both regression-based approaches and non-parametric tests, we find greater reciprocity in the two treatments that reveal intentions. These differences arise even in a simple one-shot setting without reputational benefits and are economically meaningful; they are similar in magnitude to the difference attributable to a full point reduction in the amount received. Our findings show the impact of the information environment on reciprocity in settings with uncertainty and suggest that transparency is important to reciprocity.

Keywords Reciprocity · Gift exchange · Uncertainty · Information

JEL Classification C72 · C91 · D80 · D91

✉ Leonard Wolk
l.wolk@vu.nl

Nathan W. Chan
nchan@umass.edu

[1] Stockbridge Hall, 80 Campus Center Way, Resource Economics, University of Massachusetts Amherst, Amherst MA 01003, USA

[2] Department of Finance, Vrije Universiteit Amsterdam, De Boelelaan 1105, NL-1081HV Amsterdam, The Netherlands

1 Introduction

Reciprocity is central to economic relationships and has particular importance given the prevalence of incomplete contracts (Fehr and Gächter 2000). Intentions, or perceived intentions, may affect agents' decisions, even when observable outcomes are indistinguishable (McCabe et al. 2003; Dufwenberg and Kirchsteiger 2004; Charness et al. 2007; Stanca et al. 2009; Stanca 2010; Sebald 2010). Yet assessing intentions can be difficult, particularly when they are shrouded by uncertainty.

We ask: how does positive reciprocity unfold when outcomes are determined by both intentions and chance? We posit that transparency of intentions may enhance trust in and of itself. This issue is germane to many settings that involve both trust and chance. For example, consider taxi services. A foreign visitor may be unfamiliar with optimal routes and traffic, making it difficult to ascertain whether a circuitous and lengthy journey resulted from the driver's intentions or from stochastic factors. Absent such knowledge, the foreign visitor may be hesitant to offer a generous tip, whereas a local passenger will have more cause to trust. In politics, voters may reward or punish politicians based on policy outcomes, even though those outcomes are not entirely under the politician's control. Transparency from the politician may modulate trust and support at the ballot box, even conditional on the same policies being delivered. Intentions, and observability of intentions, may similarly matter for customers and service providers for car and home repairs. Stochastic factors can affect the success of diagnosis and repair, and the customer's reciprocity (e.g., online review, willingness to rehire) may depend upon the transparency of the provider's intentions. In all of these examples, observability of intentions can affect underlying behaviors as well as the welfare of the involved parties.

To study these settings, we modify the classic gift-exchange game by making it possible for the first-mover's gift to experience a partial loss. In one treatment (*Intentions and Outcomes*, or *I+O*), the respondent observes the amount received and whether a loss occurred, making the first-mover's intention perfectly observable. We contrast this with two treatments where information about the loss is hidden. In the first such treatment (*Outcomes*), the respondent can only observe the amount received, leaving uncertain the first-mover's intention. In the other (*Intentions*), the respondent observes the amount sent but not how much has actually been received.

We find greater reciprocity when intentions are known than when they are not. Differences are economically meaningful: they are roughly the same magnitude as the difference attributable to a full point reduction in the amount received. We find these results in a simple one-shot setting without reputational benefits, suggesting that transparency is a virtue in itself for reciprocity, even while ruling out dynamic strategic considerations. Our experimental setup and results are unique from the extant literature, which has focused primarily on observable intentions and observable outcomes.

In reciprocal relationships, one must assess the kindness (intentions) of a partner's actions (Rabin 1993; Dufwenberg and Kirchsteiger 2004; Falk and

Fischbacher 2006), and chance can furthermore influence perceptions of kindness (Sebald 2010). Several studies incorporate chance in experimental settings with reciprocity. Closely related to our work is that of Charness et al. (2007). Their gift exchange game includes a random variable that can increase or decrease the first-mover's gift. Conditional on the same amount received, respondents are more generous when good intentions (and bad luck) give rise to a given result than when good luck augmented an otherwise small gift. Cushman et al. (2009) use a different experimental design but likewise allow different choice/chance combinations to reach the same consequential outcomes. These studies demonstrate the importance of both intentions and outcomes, with strong parallels to our *I+O* treatment. However, our additional *Intentions* and *Outcomes* treatments shed unique light on the nature of reciprocity when risk and incomplete information are involved.

Rubin and Sheremeta (2016) implement random shocks to effort in a three-stage principal-agent game. First, the principal offers a suggested wage and effort combination; second, the agent chooses effort; and third, the principal observes the outcome and can reward or punish the agent. The authors vary whether a random shock alters the agent's effort in the second stage, and they also vary whether the principal can observe this random shock (in addition to the realized outcome) in the third stage. Adding random shocks to the effort leads to lower effort by agents and lower payoffs for the principal but, interestingly, not for the agents. In Rubin and Sheremeta (2016), the principal can communicate expectations in the first round and use that as a reference point in the third stage; in contrast, our focus is on situations where an individual cannot offer a contract or signal expectations.

Other papers also explore related questions in different settings. Rand et al. (2015) use a repeated prisoner's dilemma and report that observable intentions lead to more cooperation. However, they study cooperativeness in an infinitely repeated prisoner's dilemma game, which is distinct from reciprocity in a one-shot game like ours. In spite of the one-shot setting, we continue to find that observability of intentions matters for reciprocity. Falk et al. (2008) implement a sequential game to show that intention-based models are limited and that preferences for fairness remain important for the respondent. Gago (2021) uses a dictator game with punishment opportunities, finding that unkind intentions trigger punishments, even if the realized outcome is not a bad one. His subjects are fully informed about intentions, whereas in our experiment, we directly compare behavior when varying the observability of intentions.

In a closely related paper, Toussaert (2017) studies the role of intentions using a noisy binary trust game in which the first-mover's decision can be replaced by a random decision by a computer with some probability. The second-mover faces uncertainty about the likelihood that this happens and is unaware of the true state they are in.[1] The findings show that trust relationships are less likely to occur when the

[1] Because of how noise is incorporated, the first-mover's intentions are never observable or even inferrable by the second mover (except in the edge case with zero noise). She asks: "When actions are noisy signals of trust intentions, how much noise becomes too much noise for the trust-reciprocity outcome to emerge?" (p.141).

probability of computer involvement is large, suggesting that players are more prosocial when trust intentions can be more credibly signaled. Building on the insights of Toussaert (2017), we directly vary the observability of intentions, outcomes, or both in this paper. In our case, the respondent always knows that the intention is at least as kind as the outcome. Furthermore, the action sets for both the first- and second-mover are discrete rather than binary in our study. This has the advantage that the second-mover can adjust their response to match their beliefs more closely than when only having two actions (e.g. trust or don't trust) at their disposal. This makes our game relevant for questions such as: "When something bad happens in a trust relationship, how much reciprocity will we see? How does the level of reciprocity depend on the information available to the respondent?" A further difference is that Toussaert (2017) uses the strategy method; doing so can dampen emotional responses, which may be important for pro-social behavior.

In related work by Friehe and Utikal (2018), a player chooses between a probability distribution that will favor herself or a different distribution that will (in expectation) provide better benefits to her partner. The partner only observes the final outcome, but there is some probability that the first-mover's choice will be revealed. The authors find that intentions and outcomes both matter, but furthermore, the respondent will punish more strongly if the first-mover attempts to conceal the original choice. Their findings suggest that hiding intentions is viewed as unfair. Unlike (Friehe and Utikal 2018), we focus on the (exogenous) observability of intentions rather than (endogenous) concealment of intentions, and we do so in a setting with positive rather than negative reciprocity.

Taken together, our study is unique in identifying the dual roles of uncertainty and the overarching information environment. Our work provides new insights on settings with incomplete or unenforceable contracts by investigating how information—or lack of information—on the realization of uncertainty shapes reciprocal relationships. These findings are relevant for wide classes of problems in the real-world where uncertainty may shroud the link between intentions and outcomes.

2 Experimental design and procedure

2.1 Design

Subjects were randomly assigned to pairs to play an anonymous one-shot game based on the classic gift-exchange game. The game involved two roles—a first-mover (P1, "she") and a respondent (P2, "he")—and each subject was randomly assigned to one of these roles.

Each subject was endowed with 5 points. P1 was then asked to allocate some of her endowment to send to the P2, keeping the remainder of her endowment in her private account. The amount sent would be multiplied by 3, and that total would be

entered into P2's account. P2 would then decide how much of his initial endowment (from 0 to 5) to send back to P1. The amount sent would be multiplied by the same factor of 3, and the total would be entered into P1's account.

We compare three experimental conditions with visibility of Intentions and Outcomes (*I+O*), Outcomes only (*Outcomes*), and Intentions only (*Intentions*). In all treatments, there was a 50% chance that one unit would be lost from the first-mover's gift *before* being multiplied by the multiplication factor.[2] For *I+O*, P2 could observe both the outcome and whether there was a loss. The *Outcomes* and *Intentions* treatments were identical to *I+O*, except P2 could not observe whether a loss occurred or not. In *Outcomes*, the respondent only observed the amount received; in *Intentions*, only the amount sent was shown.

2.2 Procedure

We ran the experiment online with participants recruited via Prolific (https://www.prolific.ac/), a platform for online studies.[3] All participants had to be (1) above the age of eighteen and (2) fluent in English to ensure that they understood the task. We recruited 468 participants for the study, for a total of 234 pairs. The data for *I+O* and *Outcomes* were collected across multiple sessions between 2018 and 2020, and the data for *Intentions* were collected in September and October of 2021. The experimental software was implemented in Python running the oTree server (Chen et al. 2016). We obtained ethical approval from Vrije University's Research Ethics Review Board at the School of Business and Economics and the Institutional Review Board at the University of Massachusetts Amherst.

Subjects began by providing informed consent. Those giving consent were randomly assigned into pairs for the one-shot game. They proceeded to read instructions that outlined the rules of the game, information about payoffs, and an example scenario. Subjects were then asked to complete two unincentivized quiz questions to check their comprehension. The first control question (true/false) was answered correctly by 420 out of 468 subjects in a single try and the second control question (calculation) was answered correctly by 129 subjects on their first and only try. For the second control question, another 130 subjects gave an answer that was within two points of the correct answer. This error is equivalent to forgetting to deduct the cost of their own contribution from the payoff. After each answer, they received feedback on their responses, including detailed information on how to solve the problems.[4]

[2] If P1's gift was initially zero or one, a loss would reduce the gift to zero (i.e., gifts could not be made negative by the loss). This design choice was made for practical reasons, as we did not want to take money for participants.

[3] Prolific has reliable data quality and has been used in prior experimental studies. For direct studies of Prolific data quality, see (Palan and Schitter 2018), Arechar et al. (2018), and Peer et al. (2017).

[4] Relatively few participants got the second control question correct, which may raise concerns about whether comprehension affects our experimental findings. We address discuss these concerns in Sect. 4.

Table 1 Participant characteristics by treatment. The Prolific score is a rating of the user's quality as assessed by Prolific with a maximum of 100

	I+O	Intentions	Outcomes
Age	32.0	25.1	29.94
Female	48.9%	62.4%	45.2%
Student	37.4%	60.1%	45.2%
Prolific score	99.2	99.6	99.1

Table 2 Summary statistics split by treatment

	I+O	Intentions	Outcomes
Sent	3.37 (1.27)	3.33 (1.16)	3.37 (1.17)
Received	2.84 (1.35)	2.85 (1.36)	2.87 (1.21)
Returned	3.24 (1.31)	3.1 (1.13)	2.67 (1.3)
Loss	0.53 (0.50)	0.48 (0.50)	0.49 (0.50)
N	92	79	63

Variable means are shown as the primary entries; standard deviations are in parentheses

From there, subjects played the one-shot game. Upon completion, subjects were paid via the Prolific system using a conversion rate of 1 point = \$0.30 USD. In general, once a subject entered the experiment on Prolific, it took only a few minutes to complete. Screenshots of the instructions and decisions screens for the *I+O* treatment are available in Figs. 1, 2, 3 and 4 in Appendix 1.

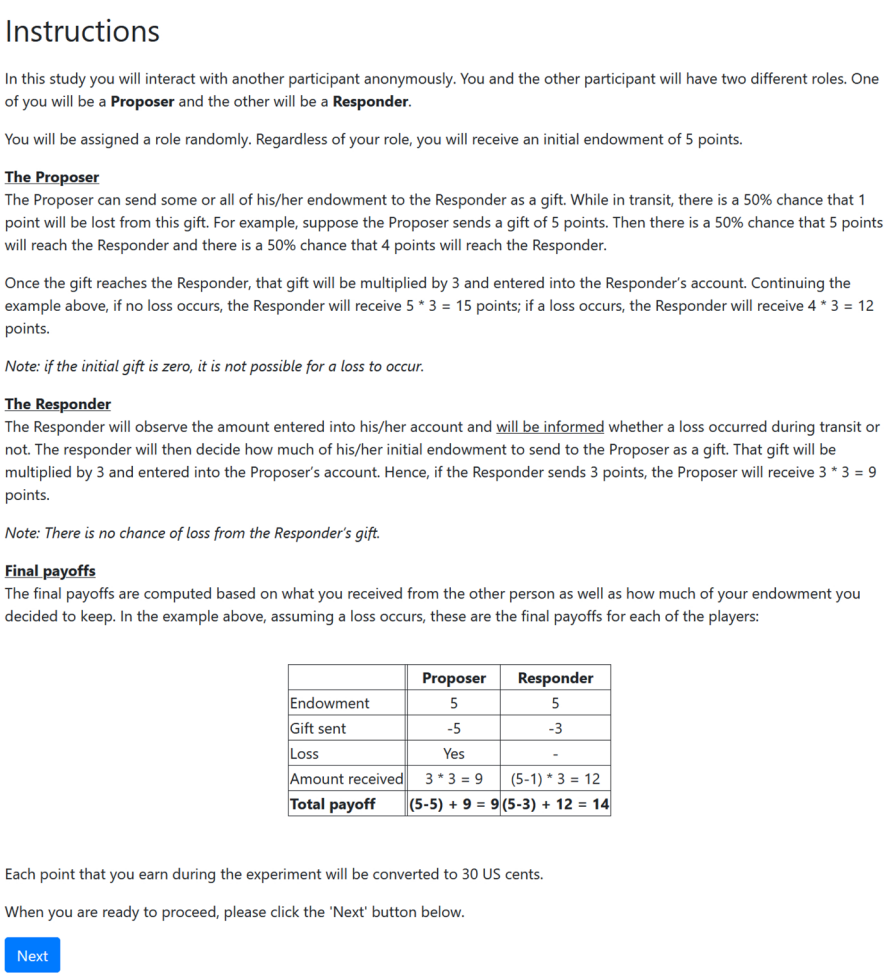
3 Results

3.1 Sample composition

Prolific provides information on participant characteristics, which we can use to assess balance across treatments. In Table 1, we observe only minor differences between *I+O* and *Outcomes*; the third treatment, *Intentions*, however, shows meaningful differences from the others. Due to this, we control for participant characteristics in subsequent analysis.[5] However, as we will show, our results are very similar regardless of whether we use non-parametric tests or regressions that include or exclude controls for these individual characteristics.

Table 2 presents summary statistics for each treatment, with averages for the original amount sent by the first-mover (Sent), the amount received by the respondent (Received), the amount the respondent returned to the first-mover (Returned), and a binary variable indicating whether a loss occurred (Loss). There are no significant

[5] As mentioned in Sect. 2.2, data for *Intentions* were collected at a later date in which time Prolific's subject pool may have changed.



Instructions

In this study you will interact with another participant anonymously. You and the other participant will have two different roles. One of you will be a **Proposer** and the other will be a **Responder**.

You will be assigned a role randomly. Regardless of your role, you will receive an initial endowment of 5 points.

The Proposer

The Proposer can send some or all of his/her endowment to the Responder as a gift. While in transit, there is a 50% chance that 1 point will be lost from this gift. For example, suppose the Proposer sends a gift of 5 points. Then there is a 50% chance that 5 points will reach the Responder and there is a 50% chance that 4 points will reach the Responder.

Once the gift reaches the Responder, that gift will be multiplied by 3 and entered into the Responder's account. Continuing the example above, if no loss occurs, the Responder will receive 5 * 3 = 15 points; if a loss occurs, the Responder will receive 4 * 3 = 12 points.

Note: if the initial gift is zero, it is not possible for a loss to occur.

The Responder

The Responder will observe the amount entered into his/her account and will be informed whether a loss occurred during transit or not. The responder will then decide how much of his/her initial endowment to send to the Proposer as a gift. That gift will be multiplied by 3 and entered into the Proposer's account. Hence, if the Responder sends 3 points, the Proposer will receive 3 * 3 = 9 points.

Note: There is no chance of loss from the Responder's gift.

Final payoffs

The final payoffs are computed based on what you received from the other person as well as how much of your endowment you decided to keep. In the example above, assuming a loss occurs, these are the final payoffs for each of the players:

	Proposer	Responder
Endowment	5	5
Gift sent	-5	-3
Loss	Yes	-
Amount received	3 * 3 = 9	(5-1) * 3 = 12
Total payoff	**(5-5) + 9 = 9**	**(5-3) + 12 = 14**

Each point that you earn during the experiment will be converted to 30 US cents.

When you are ready to proceed, please click the 'Next' button below.

Next

Fig. 1 Instructions

differences across treatments, except for the amount returned in *Outcomes*, which is lower than in *Intentions* and *I+O*.

In what follows, we will provide a closer investigation of participant behavior. Throughout, the generic regression framework we use is:

$$y_i = \alpha + \beta treat_i + \gamma controls_i + \varepsilon_i, \tag{1}$$

where y_i is the outcome of interest (e.g., amount sent or amount returned), $treat_i$ is a categorical variable denoting the treatment that player i is assigned to, and $controls_i$ is a vector of individual characteristics.

In all, we had 234 pairs, each of which comprises a unique data point. Sample sizes for each treatment are: 92 in *I+O*, 63 in *Outcomes*, and 79 in *Intentions*. The subsequent regressions use these sample sizes, pooling across treatments; in some

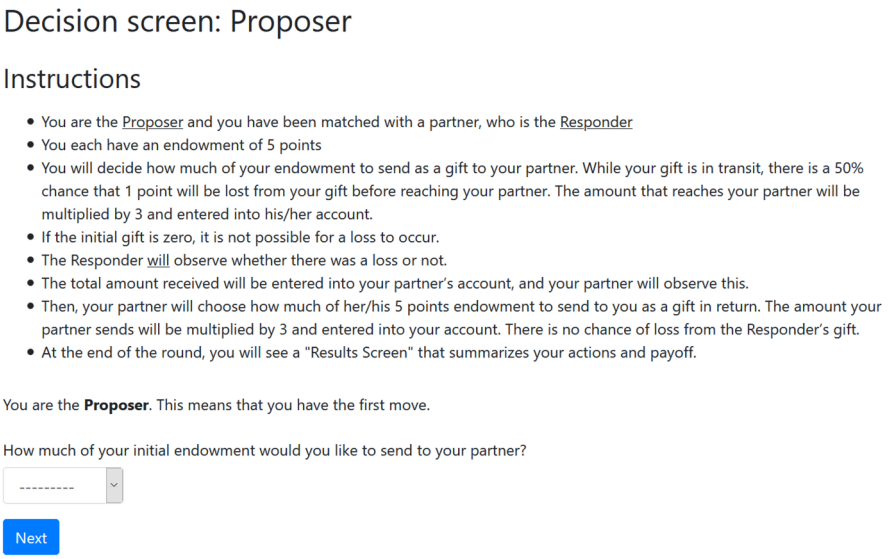


Fig. 2 First-mover task

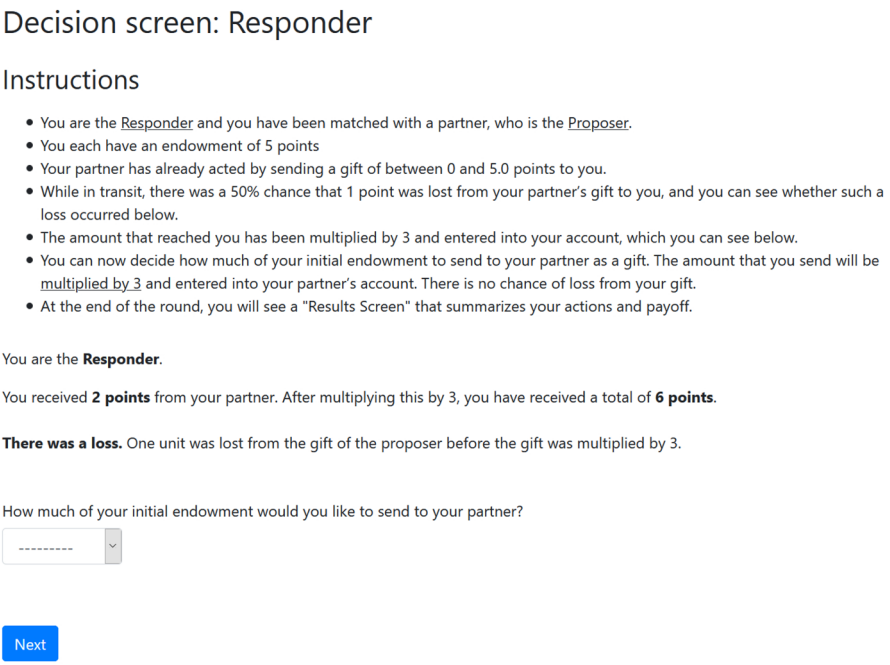


Fig. 3 Respondent task

Fig. 4 Result screen

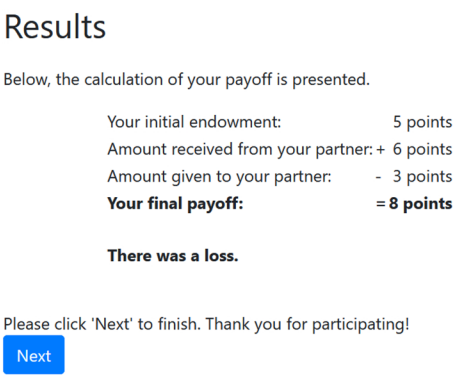


Table 3 Regression for P1 amount sent, with *Outcomes* as the baseline treatment

	Sent	
	(1)	(2)
I+O	0.00	−0.05
	(0.20)	(0.21)
Intentions	−0.04	0.06
	(0.20)	(0.21)
Student (=1)		0.08
		(0.19)
Female (=1)		−0.22
		(0.17)
Prolific score		0.14***
		(0.04)
Age (years)		0.01
		(0.01)
Constant	3.37***	−10.80**
	(0.15)	(4.18)
Observations	234	218
R^2	0.00	0.04

Robust standard errors in parentheses. *, **, and *** indicate significance at the 0.10, 0.05, and 0.01 levels, respectively. Testing for equality between *I+O* and *Intentions* gives *p*-values of 0.982 and 0.798, respectively

specifications, the sample sizes are slightly smaller because several participants are missing data for control variables. We included the most participants in the *I+O* treatment, as this would allow us to examine whether reciprocity is affected by the observation of a loss.

Table 4 Regression analysis of P2 reciprocity. The reference treatment is denoted with a '•'. A '–' indicates that the treatment group was excluded from the regression

	Dependent variable: Returned								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Intentions	0.43**	0.52**	0.51**	–	–	−0.12	−0.06	–	•
	(0.21)	(0.22)	(0.22)			(0.17)	(0.20)		
Outcomes	•	•	•	−0.59***	−0.54***	–	–	•	–
				(0.19)	(0.19)				
I+O	–	–	–	•	•	•	•	0.49**	0.35
								(0.23)	(0.24)
I+O × loss								0.09	−0.50**
								(0.24)	(0.25)
Sent		0.30***				0.40***	0.40***		0.41***
		(0.09)				(0.07)	(0.08)		(0.08)
Received			0.26***	0.50***	0.49***			0.50***	
			(0.08)	(0.08)	(0.09)			(0.09)	
Student (=1)		0.05	0.08		0.11		−0.23	0.11	−0.27
		(0.26)	(0.26)		(0.24)		(0.22)	(0.24)	(0.21)
Female (=1)		0.03	−0.01		0.02		−0.19	0.02	−0.23
		(0.20)	(0.21)		(0.19)		(0.18)	(0.19)	(0.18)
Prolific score		−0.06	−0.05		−0.03		−0.02	−0.03	−0.04
		(0.04)	(0.04)		(0.04)		(0.06)	(0.04)	(0.06)
Age (years)		0.01	0.01		0.01		−0.00	0.01	−0.01
		(0.01)	(0.01)		(0.01)		(0.01)	(0.01)	(0.01)
Constant	2.67***	7.10*	6.38*	1.83***	4.66	1.88***	4.11	3.94	5.93
	(0.16)	(3.68)	(3.69)	(0.26)	(4.21)	(0.28)	(6.52)	(4.18)	(6.17)
I+O + I+O × loss = 0 *(p-value)*								<0.01	0.50
Observations	142	134	134	155	150	171	160	150	160
R-squared	0.03	0.12	0.11	0.28	0.28	0.16	0.17	0.28	0.19

Robust standard errors in parentheses. *, **, and *** indicate significance at the 0.10, 0.05, and 0.01 levels, respectively

3.2 First-mover behavior

Before examining respondent reciprocity, we first establish whether first-movers' actions differ across treatments. Because different information will be revealed to the respondent, this may in turn affect first-mover behavior. We regress the amount sent by P1 on treatment categories and present results in Table 3; *Outcomes* is the omitted reference category.

For both regressions, we find no significant differences in P1 behavior across treatments, and the inclusion/exclusion of participant characteristics does not meaningfully affect between-treatment differences. Moreover, we verify using

t-tests and Mann-Whitney U tests that there are no significant differences in the average amount sent by P1 or the amount received by P2 across treatments. Overall, this implies that the amount sent by P1 is not significantly driven by the information that we provide to P2. On this evidence, we feel confident to proceed in analyzing P2's reciprocity—which is the object of primary interest—using both non-parametric and regression-based tests.

3.3 Reciprocity from the respondent

We find higher reciprocity when respondents can perfectly infer intentions of the first-mover. We present several regressions to support this claim in Table 4.

First, consider the *Outcomes* and *Intentions* treatments. Comparing only the amount returned (Column 1 in Table 4) reveals significant differences in respondent behavior: respondents in *Intentions* return on average 0.43 units more, simply from revealing intentions of the first-mover instead of outcomes. This finding is robust to controlling for respondent characteristics, the amount sent, and the amount received (Columns 2 and 3).[6] In short, columns 1–3 offer the same fundamental insight: return amounts are 0.43−0.52 points (16%−19.5%) higher in *Intentions* than *Outcomes*. These findings are all significant at the 5% level and the magnitudes are comparable across specifications.

However, simply comparing *Intentions* and *Outcomes* does not provide a complete picture because they have two degrees of difference: *Intentions* provides information on intentions but not outcomes, while *Outcomes* does the reverse. To address this issue, we bring in the *I+O* treatment, which is only one degree of difference from each. Both *I+O* and *Outcomes* reveal outcomes, but *I+O* also reveals intentions. Similarly, *I+O* and *Intentions* both reveal intentions, but *I+O* additionally reveals outcomes.

I+O features higher reciprocity than *Outcomes*, as seen in Table 4, Columns 4 and 5. Both regressions condition on the P2 outcome (i.e., amount received, which is observable in both treatments), and the latter also controls for P2 individual characteristics. We find that P2 returns 0.54−0.59 points (20%–22%) more in *I+O* than in *Outcomes*. The coefficient on amount received is also sensible, indicating that receiving more leads P2 to return more to P1.

Next, we compare *I+O* and *Intentions* in Table 4, Columns 6 and 7. Here, both regressions condition on the P1 intentions, and the latter also controls for P2 individual characteristics. We find no significant difference in reciprocity between *I+O* and *Intentions* in either specification, while reciprocity tends to increase in the P1 intention (amount sent), as one might expect.

Lastly, let us consider how reciprocity differs in *I+O* under different loss realizations, relative to *Outcomes* (Column 8) and *Intentions* (Column 9). In both cases, we include a treatment dummy for *I+O*, but we additionally include an interaction term

[6] For the comparison between *Intentions* and *Outcomes*, controlling for amount sent by P1 or amount received by P2 is not entirely sensible, as P2 cannot observe the amount sent in *Outcomes* or the amount received in *Intentions*. We implement the *I+O* treatment to alleviate any concerns about the comparability between *Intentions* and *Outcomes*.

for *I+O*×Loss, which indicates how much more or less was returned when a loss was realized in *I+O*. This variable is sensible to condition on because it is visible to the respondent in the *I+O* treatment. However, we should note that the interpretation of this variable differs across columns because we condition on amount received in Column 8 (since this is visible in both *I+O* and *Outcomes*) and amount sent in Column 9 (since this is visible in both *I+O* and *Intentions*).

In Column 8, the coefficient on amount received is positive and significant, indicating that the outcome matters. However, the more interesting observation is that return amounts are higher in *I+O*, regardless of whether a loss was realized or not. Because *I+O* and *Outcomes* differ only on whether the amount sent is knowable, we attribute this difference to the observability of intentions.

In Column 9, the coefficient on amount sent is positive and significant, indicating that intentions matter. More interestingly, the coefficients on *I+O* and *I+O*×Loss have opposite signs and comparable magnitudes. The point estimates suggest that return amounts are slightly higher in *I+O* when no loss occurs (+0.35), and slightly lower when loss occurs (net effect: $0.35 - 0.50 = -0.15$, *p*-value: 0.50). Given that we condition on the amount sent, both observations are consistent with the fact that outcomes matter; return amounts are higher when the amount received is higher (i.e., no loss) than when the amount received is lower (i.e., loss). Behavior in *Intentions* is essentially a convex combination of behavior in *I+O* across loss states, which is consistent with the respondent acting on expected outcomes when they do not know about losses in *Intentions*. Thus, we conclude that communicating the intentions, with or without communicating outcomes, leads to similar P2 behavior. Any differences between the two can be attributed to differences in outcomes, with higher returns in *I+O* without loss and lower returns in *I+O* with loss.

In sum, all of these findings align with our overarching claim: knowledge of intentions affects reciprocity. *I+O* and *Intentions* feature similar behavior, with both providing information about the first-mover's intentions. Meanwhile, both of these treatments differ from *Outcomes*, where intentions are unknown. Notably, subjects reciprocate more in *I+O* than *Outcomes*, regardless of whether they experienced a loss or not in *I+O*. Moreover, our regression-based results can be corroborated with non-parametric tests. Comparing return amounts across treatments, we find significantly higher return amounts in *I+O* than in *Outcomes* (Mann–Whitney *U* test *p*-value: 0.008) and likewise higher return amounts in *Intentions* than in *Outcomes* (Mann–Whitney *U* test *p*-value: 0.059). We do not find significant differences between *I+O* and *Intentions* (Mann–Whitney *U* test *p*-value: 0.301).

4 Discussion

Extensive research has investigated the relative import of intentions and outcomes in social interactions. However, the majority of such work has focused on *known* intentions and *known* outcomes. We investigate how the overarching information

environment—i.e., whether intentions can be gleaned or not—shapes these interactions. To this end, we introduce uncertainty into the classic gift exchange game and vary the availability of information.

We find that reciprocity is higher when intentions are transparent than when they are hidden. Based on our regressions, these discrepancies in generosity between *I+O* and *Outcomes* (0.54) and *Intentions* and *Outcomes* (0.51) are of comparable or larger in magnitude than the coefficients on amount sent (0.30) or amount received (0.49). Thus, the consequences of having intentions visible or not are at least as large as the impact of increasing or reducing the first-mover's gift by a full point.

These results speak to the importance of information availability, and they would be difficult to rationalize through other channels. Besides transparency of intentions, what else might explain why *Outcomes* has uniquely low return amounts? These discrepancies are not explainable through differences in P1's generosity or differential levels of loss, as the amount sent by P1 and loss rates are comparable across treatments. Moreover, we directly control for amount sent (intentions), amount received (outcomes), and participant demographics across regression specifications, and the sizable gap in P2 reciprocity is robust to such variations.

One potential explanation is motivated reasoning.[7] In *Outcomes*, if P2 believes (or decides to believe) that no loss occurred, then he can return less to P1. Doing so would improve his own payoffs without harming his self-image or social-image. He can reasonably believe he is reciprocating appropriately given unobservability of P1's initial intention, and P1 cannot distinguish whether P2 is being unreciprocal or if P2 simply believes that no loss occurred. Such reasoning would not be possible in *I+O* and *Intentions*, where P1's intentions are known.[8]

Could weaknesses in the experimental implementation confound our interpretation? Perhaps there are concerns about anchoring and salience. *Intentions* and *Outcomes* make different values salient (amount sent and amount received, respectively). Mechanically, the amount sent will be weakly larger than the amount received, so if P2 is thinking uncritically, he may judge P1's actions to be less generous in the latter case. Along similar lines, it may be possible that P2 anchors to the value presented to him, which is the amount sent in *Intentions* and the amount received in *Outcomes*, and P2's subsequent choice may be affected by this anchoring. Thus, the *Intentions* vs. *Outcomes* comparison may be confounded by these behavioral biases. However, our findings from the *I+O* treatment help rule out both of these possibilities. *I+O*, like *Outcomes*, places greater attention and salience on the outcome. In *I+O*, respondents are informed about how much they received (exactly as in *Outcomes*) and whether a loss occurred. Hence,

[7] We thank an anonymous referee for this suggestion.

[8] Could our results be driven by inequality aversion rather than reciprocity? In our setting, concerns about reciprocity and inequality aversion would both drive P2 to return more when P1 sends more, so we cannot distinguish between these explanations. However, prior work by Engelmann and Strobel (2010) shows that behavior in games like these is better captured by reciprocity than inequality aversion.

if salience or anchoring were driving our results, then we would expect *I+O* and *Outcomes* to look similar to one another and *Intentions* to be unique among the three.[9]

Taken together, our experiment yields robust results on the importance of the overarching information environment. Future work may consider whether transparency and knowability of intentions play a similar role in settings with negative reciprocity. It will also be interesting to investigate how our findings change when uncertainty (i.e., the probability of loss) or stakes (i.e., the size of loss) change.

Appendix 1: Screenshots

In this section, we show screenshots, including instructions and tasks, for the *I+O* treatment.

Acknowledgements We thank the participants at the New England Experimental Economics Workshop at Middlebury College for valuable feedback and we are grateful to Ella Gu for excellent research assistance. We also acknowledge helpful suggestions from our editor and anonymous referees. This study was approved by the Institutional Review Board at the University of Massachusetts Amherst as well as the Research Ethics Review Board of the School of Business and Economics at VU Amsterdam. A pilot study was conducted with funding from a Colby College Social Science Grant and with approval from the Institutional Review Board at Colby College.

Declarations

Conflict of interest The authors have no competing interests to declare.

[9] Another possibility: might differences in comprehension drive differences across treatments? To probe this question, we re-run our analyses on two alternative subsamples: (i) only respondents who were correct on both control questions and (ii) only respondents who were correct on the first question and who have the correct or 'almost' correct answer for the second question, as described before. For the second criteria, the significance of results in Table 4 are retained, and many of the point estimates actually grow in magnitude. For the first criteria, the differences between *Intentions* and *Outcomes* (columns 1 to 3 in Table 4) become insignificant. Yet, in all such cases, the point estimates grow in magnitude. Crucially, we retain significance in the post-regression test ($I{+}O + I{+}O \times \text{loss} = 0$) in column 8 for both subsamples (p-values for sub-samples: (i) 0.06 (ii) 0.03). This suggests that our results are not only about P2 repaying kind intentions of P1.

References

Arechar, A. A., Gächter, S., & Molleman, L. (2018). *Conducting interactive experiments online. Experimental economics, 21*(1), 99–131.

Charness, G., & Levine, D. I. (2007). Intention and Stochastic Outcomes: An Experimental study. *The Economic Journal, 117*(522), 1051–1072.

Chen, D. L., Schonger, M., & Wickens, C. (2016). oTree-An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance, 9*, 88–97.

Cushman, F., Dreber, A., Wang, Y., & Costa, J. (2009). Accidental Outcomes Guide Punishment in a "Trembling Hand Game''. *PLOS ONE, 4*(8), 1–7.

Dufwenberg, M., & Kirchsteiger, G. (2004). A theory of sequential reciprocity. *Games and Economic Behavior, 47*(2), 268–298.

Engelmann, D., & Strobel, M. (2010). Inequality Aversion and Reciprocity in Moonlighting Games. *Games, 1*(4), 459–477.

Falk, A., & Fischbacher, U. (2006). A theory of reciprocity. *Games and Economic Behavior, 54*(2), 293–315.

Falk, A., Fehr, E., & Fischbacher, U. (2008). Testing theories of fairness-Intentions matter. *Games and Economic Behavior, 62*(1), 287–303.

Fehr, E., & Gächter, S. (2000). Fairness and Retaliation: The Economics of Reciprocity. *Journal of Economic Perspectives, 14*(3), 159–181.

Friehe, T., & Utikal, V. (2018). Intentions under cover - Hiding intentions is considered unfair. *Journal of Behavioral and Experimental Economics, 73*, 11–21.

Gago, A. (2021). Reciprocity and uncertainty: When do people forgive? *Journal of Economic Psychology, 84*, 102362.

McCabe, K. A., Rigdon, M. L., & Smith, V. L. (2003). Positive reciprocity and intentions in trust games. *Journal of Economic Behavior & Organization, 52*(2), 267–275.

Palan, S., & Schitter, C. (2018). Prolific.ac-A subject pool for online experiments. *Journal of Behavioral and Experimental Finance, 17*, 22–27.

Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology, 70*, 153–163.

Rabin, M. (1993). Incorporating Fairness into Game Theory and Economics. *The American Economic Review, 83*(5), 1281–1302.

Rand, D. G., Fudenberg, D., & Dreber, A. (2015). It's the thought that counts: The role of intentions in noisy repeated games. *Journal of Economic Behavior & Organization, 116*, 481–499.

Rubin, J., & Sheremeta, R. (2016). Principal-Agent Settings with Random Shocks. *Management Science, 62*(4), 985–999.

Sebald, A. (2010). Attribution and reciprocity. *Games and Economic Behavior, 68*(1), 339–352.

Stanca, L. (2010). How to be kind? Outcomes versus intentions as determinants of fairness. *Economics Letters, 106*(1), 19–21.

Stanca, B. L., & Corazzini, L. (2009). Testing theories of reciprocity: Do motivations matter? *Journal of Economic Behavior & Organization, 71*(2), 233–245.

Toussaert, S. (2017). Intention-based reciprocity and signaling of intentions. *Journal of Economic Behavior & Organization, 137*, 132–144.