**Noname manuscript No.**
(will be inserted by the editor)

# Nice people or potential cooperators when keeping promises

## An experimental and Bayesian account for two explanations

**Said Jiménez · Kamila E. Sip · Roberto E. Mercadillo · Diego Angeles-Valdez · Jairo Muñoz-Delgado · Juan J. Sánchez-Sosa · Eduardo A. Garza-Villarreal ·**

Received: date / Accepted: date

**Abstract** Promises increase cooperation between non-genetically related individuals, whereas lies and betrayals occur daily even in close social relationships. Despite this, the effect of social closeness on the decision to keep or break a promise has not been thoroughly studied. We conducted an experiment in which subjects could freely decide if they broke or kept a promise to three partners with different levels of social closeness: null (computer), low (strange) and high (friend). If subjects kept their word with the computer, it provides evidence in favor of intrinsic motivation (comply to do the correct thing), while extrinsic motivation would be supported (comply to facilitate future cooperation). Using Hierarchical Bayesian Modeling, we found that as social closeness increases, the probability of breaking promises decreases monotonically. The evidence ratio in favor of the hypothesis stranger $<$ computer was 7.47, and friend $<$ stranger was 8.64. Furthermore, using Bayesian Cumulative Modeling we found that subjects were very consistent between their promises and their choices. Although it was more probable that our subjects kept their promises to any of the three partners, there was a proportion of transgressions to the computer, whose inference excludes zero with a 95% posterior probability. Also, our subjects never broke the promise to their friends. Our results suggest that subjects were predominantly "nice people" because they kept their promises even to partners without social closeness. However, if the subjects were not certain that their partner will also be a "potential cooperator", dishonesty emerged in the form of broken promises.

Said Jiménez
Department of Psychology, National University of México (UNAM)
E-mail: said.ejp@gmail.com

Eduardo A. Garza-Villarreal
Institute of Neurobiology, UNAM
E-mail: egarza@gmail.com

**Keywords** Deception · Hierarchical Bayesian Models · Bayesian Cumulative Models · Social closeness · Promises ·

## 1 Introduction

Humans are social animals and we obtain survival advantages from proper socialization. Historically, belonging to a social group has provided protection and guaranteed access to resources such as food, and currently, there is evidence that belonging to a social group has a positive effect on longevity, physical and mental health (Tough, Siegrist, and Fekete 2017; Holt-Lunstad 2018). The interaction between social closeness and cooperation seems to be a key feature to the formation of large groups of individuals without a genetic relationship, and such groups form the basis of communities, societies, and nations, arguably constituting one of the most fundamental conditions for human survival (Fehr and Fischbacher 2003; Fehr and Schurtenberger 2018). A recent meta-analysis showed that communication, regardless of the medium, is the most important element for cooperation between individuals (Balliet 2010). In parallel, there is evidence that following communication sessions between completely unknown individuals, there is an increase in subjective social closeness(Aron et al. 1997). These findings suggest the importance of the relationship between communication, cooperation, and social closeness.

One of the forms of communication that have received the most attention in psychological game theory is "the promise". In an interaction between a sender and a recipient, it is believed that the sender's promises influence the recipient's beliefs, influencing their ability to trust and cooperate (Charness and Dufwenberg 2006; Vanberg 2008). Commonly, however, there are situations where promises are broken, trust is betrayed and people are deceived. The General Social Survey (GSS) in 2018 showed that approximately 11% of people surveyed responded that they have had sexual intercourse with someone different to their partner while they were married (Smith et al. 2018). In economic and social terms, the United Nations Organization has stressed that deception (corruption) has a global cost of at least 5 percent of the annual gross domestic product (GDP), and it is a generator of conflict and instability in all nations (UN 2018). Hence, breaking promises may greatly affect our own social groups and society itself.

Despite the widespread presence of interactions between human belonging to the same group, it is not known how social closeness between participants can affect keeping or breaking promises. This is important because it has been pointed out that lies and betrayals of trust occur daily even in close relationships such as friends, coworkers, or classmates, and their consequences can be serious, from loss of the relationship to work loss and damage to reputation (DePaulo et al. 2004). The two main motivations that have been pointed out in the literature regarding keeping promises are: instrumental and intrinsic (Baumgartner et al. 2009). The instrumental suggests that promises are kept to make future cooperation easier while the intrinsic suggests that the promises are kept to do what is "morally right". In the laboratory, the betrayals of trust have been explicitly studied in experiments where subjects have incentives to lie and break promises. However, the vast majority of studies have been conducted in people without any social closeness (strangers) (Gneezy 2005; Gneezy, Rockenbach, and Serra-Garcia 2013; Baumgartner et al. 2009; Baumgartner, Gianotti, and Knoch 2013; Charness and Dufwenberg 2006; Mazar, Amir, and Ariely 2008; Fischbacher and Föllmi-Heusi 2013). Some studies have addressed deception in close interpersonal relationships using self-report measurements or rather explore the development of deception detection skills in same-sex friends, with clear limitations due to the social desirability bias (DePaulo and Kashy 1998; DePaulo et al. 2004; Anderson, DePaulo, and Ansfield 2002).

In this study, we wanted to understand the effects of social closeness on keeping or breaking promises and cooperation, by means of an experimental trust game with three partners with different levels of social closeness. For this, subjects performed a standard trust game in pairs with three phases: 1) the trustee makes a promise to pay half of his earnings regardless of who his/her investor; 2) the investor receives the promise and decides if he/she invests his/her initial budget; 3) in case the investor has given the budget, the trustee faces the decision to pay or not to pay half of his earnings (keep or break promise). To evaluate the effect of social closeness, our subjects participate in the role of trustee with three partners with different levels of social closeness (zero, low and high) in the role of the investor: a computer, a stranger, and a close friend. Our main pre-registered hypothesis was:

1. Social closeness will reduce the decision of breaking the promise. According to the instrumental motivation, we expect that subjects keep their promises with friends for the purpose of facilitating future cooperation, which probably can be extended even beyond the trials in the experiment. We also expect, although to a lesser extent, that the subjects keep promises to strangers, with the purpose of facilitating cooperation during the experiment. Finally, we anticipate that the participants break the promises to the computer because it is a partner without any social closeness and they could not ensure cooperation in future trials. It should be noted that if the participants keep the promises to the computer, evidence would be given in favor of intrinsic motivation (Pre-registered hypothesis is available here: `https://osf.io/53jqr`).

   Our secondary exploratory hypotheses were:

2. There will be an effect of social closeness on cooperation, regardless of promises. According to the instrumental motivation, the payment will be greater for partners with whom the subject anticipates greater cooperation in the future (friend > stranger > computer).
3. We hypothesize that the subjects will cooperate according to their degree of commitment expressed through promises.

## 2 Methods

### 2.1 Subjects

We recruited 45 subjects (15 male) from the National Autonomous University of Mexico, age range from 19 to 33 years, their minimum educational level were bachelor and right handed (for MRI). Subjects were asked to come to the study with a "close" friend, who fulfilled the following criteria: 1) he/she was matched by sex, 2) did not have a family bond and 3) was not a person with a sentimental or sexual relationship. From this sample (n = 45), 30 subjects (15 men) performed the task in a magnetic resonance imaging (MRI) scanner, however, their image data was not analyzed for the present work.

### 2.2 Task

All subjects performed an adaptation of the trust game with promises, using hypothetical monetary rewards (Baumgartner et al. 2009). The variation with respect to the original task was that, in our experiment, three partners with three levels of social closeness were presented in the role of the investor: computer (no closeness), strange (low closeness) and friend (high closeness), while the participant acted as the trustee. The task was programmed in PsychoPy2 version 1.84.2 (Peirce et al. 2019; Peirce 2008) and consisted of 24 trials between two players: trustee and investor. The trustee originally had 0 Mexican pesos and the investor had 2 pesos, the investor was presented with the opportunity to give his/her money to the trustee or keep it. If the investor gave his/her money, it was multiplied by 5 pesos, so that the trustee had 10 pesos. Finally, the trustee had to decide if he/she wanted to pay half to the investor or keep the 10 pesos. The structure described was repeated in 24 trials, however, in 4 of the trials the trustee could send a promise to the investor, the promises were that he/she would: always, mostly, sometimes, or never pay back. Each promise was valid for three trials, so that 12 of the trials had the effect of the promise and the other 12 did not. Each partner performed the role of investor in 8 of 24 total trials, however, both the promises and social closeness conditions were presented to the trustee in pseudorandom order.

Since our main interest was the trustee's behavior and to avoid unbalanced responses from investors, the investor was always the computer although the trustee was shown to have 3 types of investors, and the investors' decisions were programmed *a priori* to randomly give their amount in 6 trials and withhold the amount in 2 trials. The covert story for all our subjects was that they would be playing in real-time with their friend (whom they brought to the study), the stranger (was told it would be another same-sex unknown person) and the computer. After the study ended, the participants were debriefed about our deception. A diagram of the the experimental task is shown in Figure 1.

Promise phase in part A1 indicated that in the following three trials they could decide without the effect of the promise. In A2, the subject was asked to

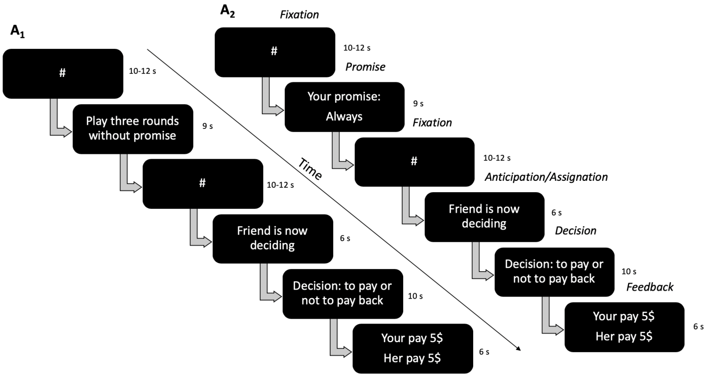


**Fig. 1** Trust game with promises, each box from left to right represents a screen that was shown to the subjects sequentially. Section A1 corresponds to an example of trials without promises and section A2 to trials with promises.

decide between 1) always, 2) mostly, 3) sometimes or 4) never pay back. This pattern was repeated. In the anticipation/assignation phase, subjects were told who their partner was for that trial (computer, stranger or friend) and was given the message that their partner was making their decision. Later in the decision phase, the subject was informed if his partner invested his $2 or not, he was reminded of his promise level (if they were trails with promises) and, in case the partner had invested, he was asked to decide whether to pay back or not. Finally, payments for each trial were shown in the feedback phase and the sequence was repeated until the task finished.

## 2.3 Procedure

Subjects came to the lab with a same-sex friend considered close by him or her. It was emphasized that they must not have a romantic or family relationship with their friends, to try to exclude the effect on cooperation due to a consanguineous relationship or sexual attraction. Also, to ensure that the subjects and their companions had a similar degree of social closeness to each other, both responded the “Inclusion of the Other in Self” IOS scale (Aron et al. 1997) without observing their partner responses. The scale consists of seven pairs of circles that vary in the degree of overlap between them, the respondent must select the pair of circles that best represents the subjective closeness to his partner. Subjects were trained in the main task by two researchers, the game was first explained verbally, and then the computer interface was presented on a portable computer. They were told that they would make their decisions with four buttons on the computer keyboard, which corresponded to the levels of promises and with only two of these to make their decision to pay back or not. Friends heard the explanation, they were told that they would make their decisions in a separate room on another computer even though they were not. Subjects performed between 3 and 6 practice tests together with the researcher who exemplified the course of the game when the investor gave his budget. When the subjects were ready, we proceeded to take them without their companions to another room (or MRI scanner it that was the case) where they would perform the task.The subjects began the task believing that they would really play with the three partners of different social closeness. In the other room, we debriefed the companion about the deception. When the subject finished the task, they were debriefed. Although no participant exercised their right, both subjects and their friends were told that they could withdraw their participation and informed consent if they considered disagreeing with any of the manipulations made by the researchers.

### 2.4 Bayesian Modeling

All models presented below were programmed in R via the **brms** package (Bürkner 2017; R Core Team 2019), which performs the inference using sampling by Markov chain Monte Carlo through **Stan** (Stan Development Team 2018). For each model, the posteriors distributions of all parameters were approximated with four chains of 2000 iterations each, the first 1000 iterations of each chain were discarded (burning period), for a total of 4000 post burning samples. Models convergence was evaluated through visual inspection of the chains and calculation of the $\hat{R}$ statistic, which for all parameters was 1, that can be interpreted as convergence (All data and R scripts are available at `https://github.com/saidejp/motivations-promises`). Vaguely informative prior distributions were used for the parameters of interest in the models, which allows the data to dominate the inference, also assumes as little as possible regarding the nature of the phenomenon, which could be adequate for the current state of evidence in the problem we are studying (McElreath 2018).

## 3 Results

### 3.1 Descriptives

We analyzed 810 decisions about paying back or not (18 per subject), as well as 180 promises (4 per subject). 69% of the decisions, were pay back, while the proportion of the promises selected were never = 5.4%, sometimes = 16.2%, mostly = 43.8% and always = 40.2%. Figure 2 shows the payment proportion depending on the levels of promises and partners. Breaking promises, meaning the subject had promised always and subsequently decides not to pay back, occurred only in 12.4% of trials (19 times they did not pay back of 153 when promised always), of which, 8.50% (13/153) were trials with the computer, 3.92% (6/153) with the stranger and 0% (0/153) with the friend. It is essential to note that breaking the promise was not the most frequent decision, when they promised always and had the computer as a partner, they decided to pay back in 91.5% of trials.

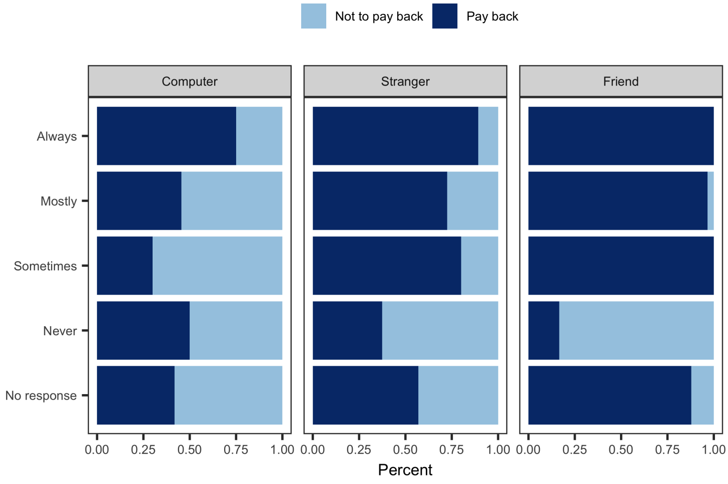


**Fig. 2** Pay back proportion by partners and promises levels, the x-axis represents all decisions made to each partner, the different shades of blue shows the proportions in which the subjects paid back or not. The y-axis presents how these proportions vary according to the promises made.

#### *3.1.1 Social closeness between partners*

To assess social closeness, 30 (15 men) of our subjects and their friends responded to the IOS scale (Aron et al. 1997; Fareri and Delgado 2014; Sip et al. 2015), which consists of seven pairs of circles that vary in the degree of overlap and represent the social closeness that an individual perceives with respect to the other. A Hierarchical and Cumulative Bayesian Model (See section: Commitment expressed in the promises), supports the hypothesis that there are no differences in subjective social closeness between subjects and their friends, with a posterior evidence ratio of 2.74 in its favor.

### 3.2 Hierarchical Bayesian Modeling

#### *3.2.1 Effect of social closeness on breaking promises*

To evaluate the hypothesis that social closeness would reduce the behavior of breaking the promise, we filtered all trials in which subjects promised that they would always pay back and then decided not to pay. Subsequently, Bayesian inference was used to assess the effect of partners on the decision to break the promise at the individual and population level. For this purpose, a Hierarchical Bayesian Model was carried out that assumes that the uncertainty in partners effect on the decision to break the promise varies depending on each individual, however, it also assumes that these variations belong to common population distributions (Gelman and Hill 2006).

$$y_i \sim \text{Bernoulli}(\theta_i)$$
$$\text{logit}(\theta_\text{i}) = \mathbf{X}\beta + \mathbf{Z}u$$

In the previous model, the decision to break the promise $y_i$ comes from a Bernoulli distribution with probability $\theta_i$, the goal of the Hierarchical Model is to predict each decision through the linear combination of the effects of each partner, transformed by its inverse link function logit (Bürkner 2017). In this model, $\beta$ and $u$ are coefficients at the population level and individual level respectively, while $\mathbf{X}$, $\mathbf{Z}$ are their corresponding design matrices. In this case, the population coefficients correspond to the presence of partners with no social closeness $\beta^{computer}$, low $\beta^{stranger}$ and high $\beta^{friend}$.

Figure 3 shows the posterior probability of breaking the promise depending on the partner, circles correspond to the medians of the distributions of the posterior estimates effects, the thick bar, and the thin bar correspond to the interval of 50% and 95% posterior probability, respectively. It is clear that the probability of breaking the promise decreases as social closeness increases, however, we calculate the reasons for evidence for the following hypotheses:

- There is a greater probability to break the promise to the computer than the stranger.
- There is more probability to break the promise to the stranger compared to a friend.

The evidence ratio, which is the ratio between the posterior probability of the mentioned hypotheses and their corresponding alternative hypotheses was 7.47 and 8.64 respectively in favor of the previous hypotheses. As seen in Figure 5 in the estimate for $\beta^{computer}$, no social closeness, we have a 95% posterior probability that the parameter for promise-breaking is between 4-38%, although it is a small proportion, it clearly excludes the probability that promises are not broken.

#### *3.2.2 Effects of promises and social closeness on cooperation*

To evaluate the effect of the experimental conditions on the decision to pay back, we used all trials were partners invest. Again, a Hierarchical Bayesian Model was fit, which assumes that the promises and partners have an effect that varies for each individual, however, it also assumes that these variations belong to common population distributions. The model estimates that the probability $\theta_i$ of the decision to pay back is based on the effect of the presence of promises $\beta^{promise}$, as well as partners with low $\beta^{stranger}$ and high $\beta^{friend}$ social closeness.

Figure 4 shows the posterior distributions of the $\beta$ coefficients in the logit scale, the center line represents the median of the distribution and the shaded area corresponds to the interval of 50% posterior probability. As shown, more than 95% of the posterior density of the population coefficients is greater than 0, which shows strong evidence of the effect of experimental conditions on

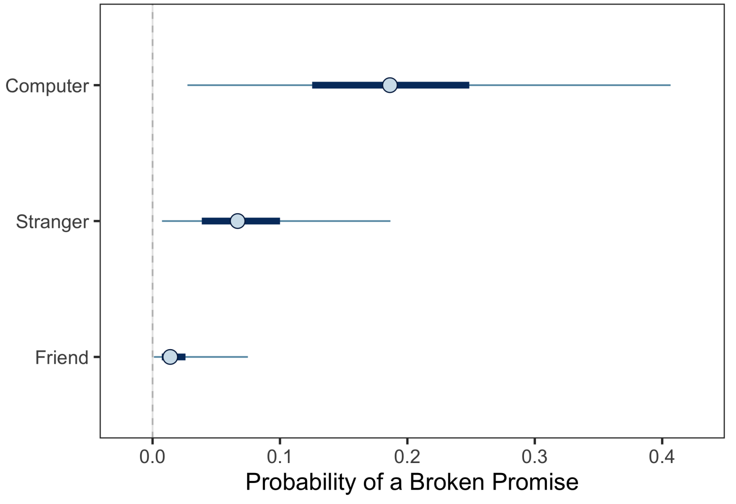


**Fig. 3** Inference regarding the probability of breaking the promise depending on social closeness. The x-axis shows the posterior probability of breaking the promise, and the y-axis indicates the three partners. The dotted line shows the probability that breaking the promise would be zero.

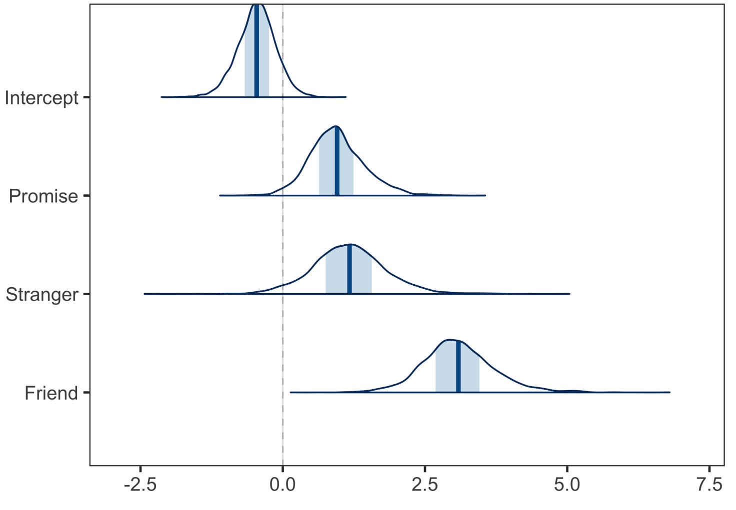


**Fig. 4** Posterior estimates in logit scale. We show posteriors distributions of the hierarchical Bayesian model coefficients for the effect of experimental conditions on the decision to cooperate. The x-axis shows the magnitude of the logit scale coefficients, and the y-axis shows the experimental conditions associated with that effect.

the decision to pay back, although the presence of promises and the stranger clearly increase the probability of paying, the presence of the friend is the condition that has the greatest effect on this behavior. Posterior distributions clearly show that the probability of cooperation increases as social closeness does.

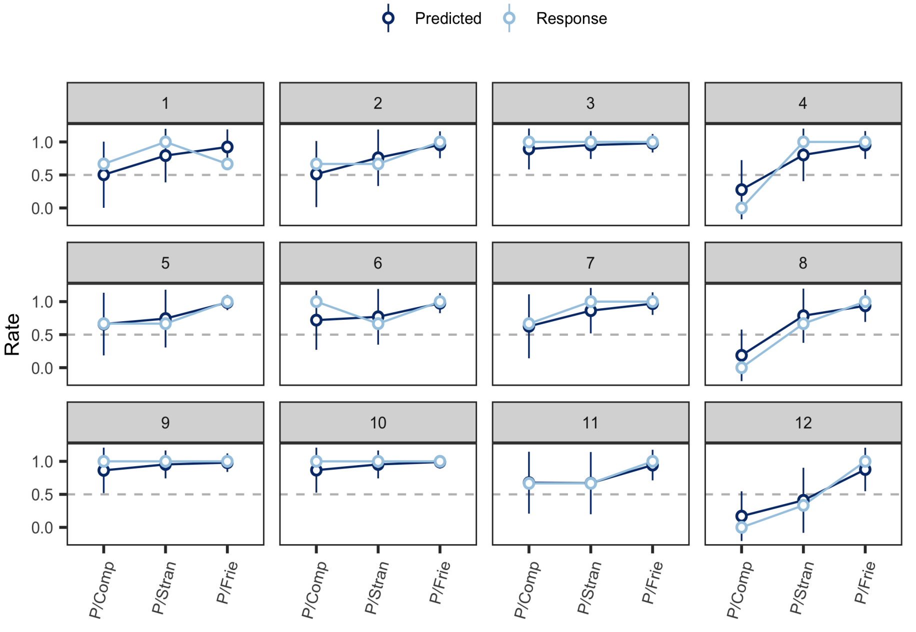


**Fig. 5** Posterior predictive over pay back rate. Correspondence between payment proportions and model predictions depending on the game partner for the first twelve subjects in the sample. The x-axis pointed the three playmates when the subjects had made a promise, while the y-axis indicates the proportion of decisions in which subjects payback.

Figure 5 shows the posterior predictive distribution in contrast to the proportion of decisions to each partner during the promises phase, and each panel corresponds to one of the first twelve subjects. The responses on the dotted line would indicate that the subject paid back random to that partner. Posterior predictive distribution simulates observations of the model and compares them with the actual data, it helps us to identify if the model is sufficiently close to the process that generated the data (Schad, Betancourt, and Vasishth 2019; Lee and Wagenmakers 2014). It can be seen that there is a great correspondence between the subject's responses and predictions of the model. Even in cases where the model is "wrong" (for example, subject 1), the observed response is in the range of a predicted standard deviation, which gives credibility to estimates.

*3.2.3 Effects of social closeness on cooperation varying by promises*

In this model, it is assumed that social closeness has an effect that varies for each level of promise, which implies that there are levels of promise that are more sensitive than others to the effect of the partner's on the decision to pay back. Again, a Hierarchical Bayesian Model was made to estimate the effect of partners at the population level and the variations depending on the promise levels. Table 1 summarizes the subsequent distributions of the model

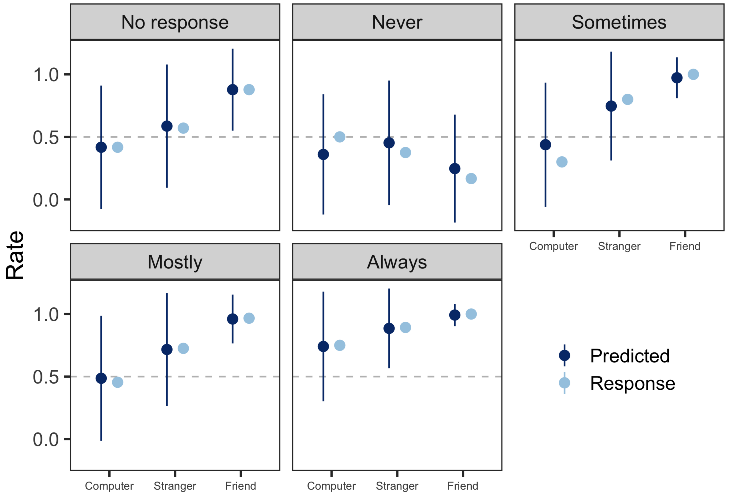


**Fig. 6** Posterior predictive over pay back rate, social closeness varying effects by promise levels. Correspondence between payment proportions and model predictions depending on the game partner at the group level. The x-axis pointed the three playmates when the subjects had made a promise, while the y-axis indicates the proportion of decisions in which subjects payback.

**Table 1** Posterior coefficients estimates

| | | | 95 % CI | |
|---|---|---|---|---|
| Term | Estimate | Est.Error | Lower | Upper |
| Intercept | -0.153 | 0.536 | -1.005 | 0.617 |
| Stranger | 0.829 | 0.410 | 0.128 | 1.470 |
| Friend | 2.084 | 1.248 | -0.151 | 3.975 |

coefficients in the logit scale, including point estimates, standard errors and Bayesian Credibility Intervals of 95%. An estimate similar to the previous models can be observed, with strong evidence of the effect of social closeness on the decision to pay back. Although the credibility range for the friend's effect includes 0, the evidence ratio that the effect is greater than zero is 15.81 with 94% of posterior probability.

On the other hand, Figure 6 shows the posterior predictive distribution, compared to the payment rates of all individuals to the different partners and their variation by the level of promise. With the exception of the promise never, a monotonic positive effect of partners is observed at all levels of promise. However, we also observe how the effect of social closeness varies depending on the strength of the promise, mainly for the decisions towards the computer.

## 3.3 Cumulative Bayesian Modeling

### *3.3.1 Commitment expressed in promises*

In the original study of promises, authors chose to divide their sample into two according to the hierarchical clustering technique with Ward's method (Baumgartner et al. 2009). In this way they obtained two sets of participants that were different in their payment rates, despite the fact that they both made very high promises. For this reason, the authors named the group that paid little as "dishonest" and the group that paid a lot as "honest". In a similar exercise, in the present study we performed the hierarchical grouping technique to obtain a solution for two groups and we found two similar sets in n that we call the group "Low" and "High" (Low = 25, High = 20), with quite different payment ratios (Low = 58%, High = 83%), and a 95% confidence interval of 20% to 32% in the difference favor of the High group.

Although it seems a similar result to that reported in that paper (Baumgartner et al. 2009), we explore the pattern of promises of both groups to determine if it was possible to classify our subjects as honest and dishonest. If we hypothesize that the commitment to pay would be reflected in the level of promises selected, a group of dishonest people could generate in their partners the belief that they will pay by choosing a high level of promises (mostly or always) and, later, betraying that trust when deciding not to pay back. In order to explore whether the groups represent populations that do not differ in the level of commitment expressed in the promises, we use a Cumulative Bayesian Model, which assumes that the levels of promises are an observed ordinal variable $Y$ that originates from the categorization of a continuous latent variable $\tilde{Y}$, for this case, the expressed commitment to pay back (Bürkner and Vuorre 2019). The degree to which the subjects of the High group differ from the Low group, in normal standard deviations ($z$-values), on the latent scale of $\tilde{Y}$, has a point estimate of 0.95, which implies that the High group has 0.95 z-values greater commitment to pay back than the Low group. The 95% Bayesian Credibility Interval indicates that the High group is between 0.40 to 1.51 $z$-values of difference from Low. So we can conclude with at least a 95% probability that people belonging to the High group expressed in their promises a greater commitment to pay back than the subjects of the Low group. If we look at Figure 7, the probability of choosing the different levels of promises varies depending on the group, the one who had the highest percentage of decisions to pay back is more likely to choose always (High group), while the group that had the lowest percentage of decisions to pay back are more likely to choose mostly (Low group). According to our data, we could not justify the classification of our groups according to their honesty or dishonesty, at least not with the technique used in Baumgartner et al. (2009). Since people were quite consistent with keeping what they promised to pay.

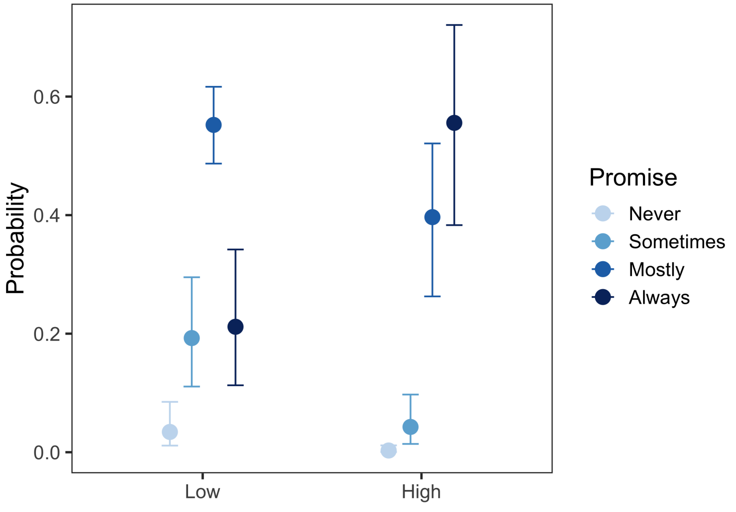


**Fig. 7** Promise choices by group. We show posterior probability estimates regarding the level of promise depending on the group. The Low group (which paid 58/100 of the trials) is more likely to promise mostly, while the High group (which paid 83/100 of the trials) is more likely to promise always.

## 4 Discussion

The goal of this study was to evaluate the effect of social closeness on keeping and breaking promises, as well as on cooperation in a trust game. Our results give evidence that zero social closeness increases the probability of breaking the promise and it decreases monotonically as social closeness increases. Likewise, as social closeness increases, so does cooperation; high social closeness has an effect on decisions that surpasses even those that all other experimental conditions. The most interesting finding was that when subjects expressed high commitment to cooperate (through choosing that they would always pay back), the probability that they comply is also quite high. However, there is a small but consistent proportion of transgressions according to our inferences.

To our knowledge, this is the first experimental study that incorporates social closeness as a predictor of the decision to break promises with socially relevant partners. In previous studies, the participants remain anonymous during the course of the tasks (Baumgartner et al. 2009), their measurements to evaluate transgressions are self-reported (DePaulo and Kashy 1998; DePaulo et al. 2004), they do not directly quantify the breaking of promises (Vanberg 2008; Charness and Dufwenberg 2006), or as we will discuss later, they use heterogeneous measures of social closeness (Glaeser et al. 2000). Our results suggest that the mere fact of considering that subjects play with a human diminishes the probability of breaking the promise. It is more likely to break the promise to the computer than to the stranger, even though that partner was not known. Also, zero social closeness is important to include because it allows us to explore the intrinsic motivation to keep the promises. If humans keep their word because it is morally correct, we would anticipate that the probability of breaking a promise was very low in the three partners and, particularly, the inference regarding the probability of breaking the promise to the computer would include zero. However, we found that, although low, the probability was greater than zero, which provides evidence that seems to contradict, at least in some degree, the hypothesis of intrinsic motivation. This means that we cannot exclude the possibility that the subjects keep promises mainly due to moral motivation because, even in the case of the computer, the subjects kept a large proportion of promises. We can, however, exclude that this is their *only* motivation.

High social closeness allows us to explore the instrumental motivation to keep promises (Baumgartner et al. 2009). In our study, subjects were more likely to break the promise to the stranger than to the friend. If humans keep their word to facilitate cooperation in the future, we would anticipate that there would be less of a chance to break the promise with a partner of high social closeness (with whom they are very likely to cooperate, even after the experiment) and, in fact, the promise to friends was not broken on any occasion. The findings concerning motivations could be explained by the social norm called "conditional cooperation", which indicates that the belief that other people cooperate at high levels also induces high levels of cooperation (Fehr and Schurtenberger 2018). Thus, the uncertainty regarding the computer's decisions could explain the probability of breaking the promise. Similarly, the information that the subjects have regarding their friend's behavior, even before the experiment, could explain the high levels of cooperation and keeping of promises towards the friend.

Cooperation has been studied with diverse experimental manipulations. In one study, the contribution in a "public goods game" was evaluated during several trials in a group made up of the same individuals (partners), and compared to another group of new different subjects for each trial (strangers) (Croson 1996). In the condition of partners, the cooperation was greater because it was a stable group compared to the group of strangers. The mentioned study does not give details regarding the recruitment of participants, thus we could assume that even in the condition of partners these were individuals who did not consider themselves socially close. Another study showed an increase in cooperation in a trust game when individuals were socially close (Glaeser et al. 2000). In this study, subjects knew each other and the researchers carefully measured several variables regarding their social connection such as the number of friends in common and years of relationship. They found greater cooperation depending on the degree of social connection. However, the study did not control for social closeness as some individuals who arrived together were allowed to perform the task together whereas others were paired differently. This limitation allowed social closeness between partners to be heterogeneous, allowed for romantic relationships and assumed that two individuals who arrived together to class were considered close to each other. To try to homogenize social closeness between our subjects and their partners, they performed the task with a same-sex friend considered to be close, emphasizing their companion could not be a romantic or sexual partner. Bayesian analysis of the IOS scale showed evidence in favor that the degree of social closeness between 30 of our 45 subjects and their friends was the same. Although results of the two mentioned studies also found that social closeness enhance cooperation (Croson 1996; Glaeser et al. 2000), it would be valuable to homogenize both the performed tasks and the statistical procedure used to establish similarity in effects magnitudes.

Our methodological contribution to the field is the use of Bayesian inference tools: Hierarchical Models and Cumulative Models. The Hierarchical Models, allowed us to model the effect of the experimental conditions on the individual response. These models assume that lower level observations (e.g. decisions of each individual) are nested in higher level units (e.g. individual subjects). Within-subjects designs have traditionally been analyzed with repeated measures AN(C)OVA, however, Hierarchical Models grant the advantages of naturally dealing with unbalanced data, include categorical and/or numerical predictors, explicitly incorporate individual variability, among others (Vuorre and Bolger 2018; Gelman and Hill 2006; McElreath 2018; Bürkner 2017). In our case, the Hierarchical Model allowed us to capture how the experimental conditions affected the decisions of each subject and the differences between them. For example, there were obviously motivated instrumental subjects such as number 8 or 12 who are very sensitive to the identity of their investors and modify their cooperation under social closeness. At the same time, there were notably intrinsic subjects such as 9 and 10 who cooperate all the time regardless of who their partners are. The Hierarchical Model naturally includes this information for the estimation of population effects, which, if not considered, would lead to inaccurate inferences (Bürkner 2017; Gelman and Hill 2006; McElreath 2018). The Cumulative Model allowed us to capture the strength of the commitment expressed through promises. A very frequent problem that has been pointed out recently is that analyzing ordinal data with methods that assume that observations are metric can lead to serious inference errors (Liddell and Kruschke 2018). The Cumulative Models assumes that the observed responses come from the categorization of a continuous latent variable (Bürkner and Vuorre 2019). In our study, we could identify that groups with different payment proportions, which in other studies have been called "dishonest" and "honest" (Baumgartner et al. 2009; Baumgartner, Gianotti, and Knoch 2013), differ in the commitment they express through their promises. Considering that the tools to perform the Cumulative Models are relatively recent, in future studies, it would be valuable to also use the new alternatives for inference. Although the Hierarchical and Cumulative models are not tools exclusively for Bayesian inference, their application from this approach represents several advantages compared to frequentist statistics. Classic problems such as multiple comparisons, the decision of when to stop collecting subjects (stopping rule), or use of planned comparisons versus post hoc, are not factors that affect the Bayesian approach (Dienes 2011). In Bayesian inference, hypothesis testing makes formal use of probability to express the plausibility of theories, and in our case, we were able to obtain evidence ratios regarding the extent of our data support hypotheses, regardless if these were proposed *a priori* or *post hoc*.

### 4.1 Limitations

A limitation of our study was the use of hypothetical rewards compared to real rewards. A relatively recent study on decision-making reported that there is more loss aversion when subjects have real monetary rewards compared to hypothetical in a risk task (Xu et al. 2016). However, other works do not report differences between the use of both types of rewards in self-control, temporal and social discount tasks (Locey, Jones, and Rachlin 2011; Johnson and Bickel 2006; Rachlin and Jones 2006). Likewise, it can be argued that our results have theoretical congruence (Fehr and Schurtenberger 2018) and are in the same direction as other works that use real money (Croson 1996; Glaeser et al. 2000). Thus, there are not many reasons to expect that other types of rewards would modify our results. Another possible limitation could be the use of multiple trials with each partner. Although the decision to use repeated measures during our design serves the purpose of reducing contaminant sources and increasing internal validity (Maxwell, Delaney, and Ken 2018), the studies could benefit from only including one trial for each partner to have more clarity regarding the difference in the intrinsic and instrumental motivations. Play a one-shot with each partner avoids the possibility that multiple trials could generate the belief that the computer can also vary its behavior according to the decisions of the trustee.

## 5 Conclusions

Subjects in our study seemed to keep their promises by a combination of *intrinsic* and *instrumental* motivations. They can be regarded as predominantly "nice people" because they kept their word even with investors with no social closeness (computer). However, their trustworthiness was far from perfect, since there was a small proportion of betrayals committed towards the computer. If subjects could not be sure their investor would be a "potential cooperator", dishonesty may have emerged in the form of broken promises. Framing our findings in terms of the theoretical predictions of the social norm "conditional cooperation", social closeness decreases the probability of breaking the promise and increases cooperation, and suggests the correspondence between commitment expressed through promises and subsequent behavior.

## Bibliography

Anderson, D. Eric, Bella M. DePaulo, and Matthew E. Ansfield. 2002. "The development of deception detection skill: A longitudinal study of same-sex friends." *Personality and Social Psychology Bulletin* 28 (4): 536–45. https://doi.org/10.1177/0146167202287010.

Aron, Arthur, Edward Melinat, Elaine N. Aron, Robert Darrin Vallone, and Renee J. Bator. 1997. "The experimental generation of interpersonal closeness: A procedure and some preliminary findings." *Personality and Social Psychology Bulletin* 23 (4): 363–77. https://doi.org/10.1177/0146167297234003.

Balliet, Daniel. 2010. "Communication and cooperation in social dilemmas: A meta-analytic review." *Journal of Conflict Resolution* 54 (1): 39–57. https://doi.org/10.1177/0022002709352443.

Baumgartner, Thomas, Urs Fischbacher, Anja Feierabend, Kai Lutz, and Ernst Fehr. 2009. "The Neural Circuitry of a Broken Promise." *Neuron* 64 (5). Elsevier Ltd: 756–70. https://doi.org/10.1016/j.neuron.2009.11.017.

Baumgartner, Thomas, Lorena R R Gianotti, and Daria Knoch. 2013. "Who is honest and why: Baseline activation in anterior insula predicts interindividual differences in deceptive behavior." *Biological Psychology* 94 (1). Elsevier B.V.: 192–97. https://doi.org/10.1016/j.biopsycho.2013.05.018.

Bürkner, Paul-Christian. 2017. "brms : An R Package for Bayesian Multilevel Models Using Stan." *Journal of Statistical Software* 80 (1). https://doi.org/10.18637/jss.v080.i01.

Bürkner, Paul-Christian, and Matti Vuorre. 2019. "Ordinal Regression Models in Psychology: A Tutorial." *Advances in Methods and Practices in Psychological Science* 2 (1): 77–101. https://doi.org/10.1177/2515245918823199.

Charness, Gary, and Martin Dufwenberg. 2006. "Promises and partnership." *Econometrica* 74 (6): 1579–1601. https://doi.org/10.1111/j.1468-0262.2006.00719.x.

Croson, Rachel T.A. 1996. "Partners and strangers revisited." *Economics Letters* 53 (1): 25–32. https://doi.org/10.1016/S0165-1765(97)82136-2.

DePaulo, Bella M, Matthew E Ansfield, Susan E Kirkendol, and Joseph M Boden. 2004. "Serious Lies." *Basic & Applied Social Psychology* 26 (2/3): 147–67. https://doi.org/10.1207/s15324834basp2602&3_4.

DePaulo, Bella M., and Deborah A. Kashy. 1998. "Everyday Lies in Close and Casual Relationships." *Journal of Personality and Social Psychology* 74 (1): 63–79. https://doi.org/10.1037/0022-3514.74.1.63.

Dienes, Zoltan. 2011. "Bayesian versus orthodox statistics: Which side are you on?" *Perspectives on Psychological Science* 6 (3): 274–90. https://doi.org/10.1177/1745691611406920.

Fareri, Dominic S., and Mauricio R. Delgado. 2014. "Differential reward responses during competition against in- and out-of-network others." *Social Cognitive and Affective Neuroscience* 9 (4): 412–20. https://doi.org/10.1093/scan/nst006.

Fehr, Ernst, and Urs Fischbacher. 2003. "The nature of human altruism." *Nature* 425 (6960). Nature Publishing Group: 785–91. https://doi.org/10.1038/nature02043.

Fehr, Ernst, and Ivo Schurtenberger. 2018. "Normative foundations of human cooperation review-article." *Nature Human Behaviour* 2 (7). Springer US: 458–68. https://doi.org/10.1038/s41562-018-0385-5.

Fischbacher, Urs, and Franziska Föllmi-Heusi. 2013. "Lies in disguise-an experimental study on cheating." *Journal of the European Economic Association* 11 (3): 525–47. https://doi.org/10.1111/jeea.12014.

Gelman, Andrew, and Jennifer Hill. 2006. *Data analysis using regression and multilevel/hierarchical models.* https://doi.org/10.2277/0521867061.

Glaeser, Edward, David Laibson, Jose Sceinkman, and Christine Soutter. 2000. "Measuring Trust." *The Quartely Journal of Economics* 115 (3): 811–46.

Gneezy, Uri. 2005. "Deception: The role of consequences." *American Economic Review* 95 (1): 384–94. https://doi.org/10.1257/0002828053828662.

Gneezy, Uri, Bettina Rockenbach, and Marta Serra-Garcia. 2013. "Measuring lying aversion." *Journal of Economic Behavior and Organization* 93. Elsevier B.V.: 293–300. https://doi.org/10.1016/j.jebo.2013.03.025.

Holt-Lunstad, Julianne. 2018. "Why Social Relationships are Important for Physical Health: A Systems Approach to Understanding and Modifying Risk and Protection." *Ssrn*, no. October 2017: 1–22. https://doi.org/10.1146/annurev-psych-122216-011902.

Johnson, Matthew W, and Warren K Bickel. 2006. "Within-subject comparison of real and hypothetical money rewards in delay discounting." *Journal of the Experimental Analysis of Behavior* 77 (2): 129–46. https://doi.org/10.1901/jeab.2002.77-129.

Lee, Michael D., and Eric-Jan Wagenmakers. 2014. *Bayesian Cognitive Modeling.* https://doi.org/10.1017/cbo9781139087759.

Liddell, Torrin M., and John K. Kruschke. 2018. "Analyzing ordinal data with metric models: What could possibly go wrong?" *Journal of Experimental Social Psychology* 79 (November 2017). Elsevier: 328–48. https://doi.org/10.1016/j.jesp.2018.08.009.

Locey, Matthew L, Bryan A Jones, and Howard Rachlin. 2011. "Real and hypothetical rewards in self-control and social discounting." *Judgment and Decision Making* 6 (6): 552–64.

Maxwell, Scott E., Harold D. Delaney, and Kelley Ken. 2018. *Designing experiments and analyzing data: a model comparison perspective.* 3rd edition. New York: Routledge.

Mazar, Nina, On Amir, and Dan Ariely. 2008. "The Dishonesty of Honest People: A Theory of Self-Concept Maintenance." *Journal of Marketing Research* 45 (6): 633–44. https://doi.org/10.1509/jmkr.45.6.633.

McElreath, Richard. 2018. *Statistical rethinking: A bayesian course with examples in R and stan.* https://doi.org/10.1201/9781315372495.

Peirce, Jonathan, Jeremy R. Gray, Sol Simpson, Michael MacAskill, Richard Höchenberger, Hiroyuki Sogo, Erik Kastman, and Jonas Kristoffer Lindeløv. 2019. "PsychoPy2: Experiments in behavior made easy." *Behavior Research Methods* 51 (1): 195–203. https://doi.org/10.3758/s13428-018-01193-y.

Peirce, Jonathan W. 2008. "Generating stimuli for neuroscience using PsychoPy." *Frontiers in Neuroinformatics* 2. https://doi.org/10.3389/neuro.11.010.2008.

Rachlin, Howard, and Bryan Jones. 2006. "Social Discounting." *Psychological Science* 17 (4): 283–86.

R Core Team. 2019. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Schad, Daniel, Michael Betancourt, and Shravan Vasishth. 2019. "Toward a principled Bayesian workflow in cognitive science." osf.io/b2vx9.

Sip, Kamila E., David V. Smith, Anthony J. Porcelli, Kohitij Kar, and Mauricio R. Delgado. 2015. "Social closeness and feedback modulate susceptibility to the framing effect." *Social Neuroscience* 10 (1): 35–45. https://doi.org/10.1080/17470919.2014.944316.

Smith, Tom, Michael Davern, Jeremy Freese, and Michael Hout. 2018. "GSS Data Explorer | NORC at the University of Chicago." https://gssdataexplorer.norc.org/variables/5067/vshow.

Stan Development Team. 2018. "RStan: The R Interface to Stan." http://mc-stan.org/.

Tough, Hannah, Johannes Siegrist, and Christine Fekete. 2017. "Social relationships, mental health and wellbeing in physical disability: A systematic review." *BMC Public Health* 17 (1). BMC Public Health: 1–18. https://doi.org/10.1186/s12889-017-4308-6.

UN. 2018. "Global Cost of Corruption at Least 5 Per Cent of World Gross Domestic Product, Secretary-General Tells Security Council, Citing World Economic Forum Data." https://www.un.org/press/en/2018/sc13493.doc.htm.

Vanberg, Christoph. 2008. "Why Do People Keep Their Promises? An Experimental Test of Two Explanations1." *Econometrica* 76 (6): 1467–80. https://doi.org/10.3982/ECTA7673.

Vuorre, Matti, and Niall Bolger. 2018. "Within-subject mediation analysis for experimental data in cognitive psychology and neuroscience." *Behavior Research Methods* 50 (5): 2125–43. https://doi.org/10.3758/s13428-017-0980-9.

Xu, Sihua, Yu Pan, You Wang, Andrea M Spaeth, Zhe Qu, and Hengyi Rao. 2016. "Real and hypothetical monetary rewards modulate risk taking in the brain." *Scientific Reports* 6. https://doi.org/10.1038/srep29520.