10-2017

Is Cooperative Memory Special? The Role of Costly Errors, Context, and Social Network Size When Remembering Cooperative Actions

Jeffrey R. Stevens

University of Nebraska - Lincoln, jstevens5@unl.edu

Tim Winke

Max Planck Institute for Human Development, tim.winke@gmail.com

Is Cooperative Memory Special? The Role of Costly Errors, Context, and Social Network Size When Remembering Cooperative Actions

Tim Winke[1,2] and Jeffrey R. Stevens[1,3]*

[1] *Center for Adaptive Behavior and Cognition, Max Planck Institute for Human Development, Berlin, Germany,*
[2] *Berlin Graduate School of Social Sciences, Humboldt-Universität zu Berlin, Berlin, Germany,* [3] *Department of Psychology, Center for Brain, Biology & Behavior, University of Nebraska-Lincoln, Lincoln, NE, United States*

Theoretical studies of cooperative behavior have focused on decision strategies, such as tit-for-tat, that depend on remembering a partner's last choices. Yet, an empirical study by Stevens et al. (2011) demonstrated that human memory may not meet the requirements that needed to use these strategies. When asked to recall the previous behavior of simulated partners in a cooperative memory task, participants performed poorly, making errors in 10–24% of the trials. However, we do not know the extent to which this task taps specialized cognition for cooperation. It may be possible to engage participants in more cooperative, strategic thinking, which may improve memory. On the other hand, compared with other situations, a cooperative context may already engage improved memory *via* cheater detection mechanisms. This study investigated the specificity of memory in cooperative contexts by varying (1) the costs of errors in memory by making forgetting defection more costly and (2) whether the recall situation is framed as a cooperative or neutral context. Also, we investigated whether variation in participants' social network size could account for individual differences observed in memory accuracy. We found that neither including differential costs for misremembering defection nor removing the cooperative context influenced memory accuracy for cooperation. Combined, these results suggest that memory accuracy is robust to differences in the cooperative context: Adding more strategic components does not help accuracy, and removing cooperative components does not hurt accuracy. Social network size, however, did correlate with memory accuracy: People with larger networks remembered the events better. These findings suggest that cooperative memory does not seem to be special compared with other forms of memory, which aligns with previous work demonstrating the domain generality of memory. However, the demands of interacting in a large social network may require excellent memory. Thus, modeling the evolution of cooperation requires an understanding of both the social environment in which agents interact and the cognitive capabilities of these agents.

Keywords: cooperation, forgetting, memory, prisoner's dilemma, social network, tit-for-tat

1. INTRODUCTION

Upon entering a bar or restaurant in your home town, you might recognize a friend and purchase a drink for him or her. Yet, when on a layover in an airport far from home, you likely would not join with a stranger and buy a drink for him or her. Why pay a cost to help a friend but not a stranger? More generally, why do we cooperate in some situations but not others? You would likely not continue to cooperate if your friend always expected you to pay for his or her drink. Unilateral cooperation allows cheating by defectors (those who do not cooperate). In evolutionary terms, a population of cooperators would not resist invasion from defectors. Therefore, cooperation must be conditional. It can be conditional on relatedness to your partner (kin selection: Hamilton, 1964), costs imposed on defection (punishment: Boyd and Richerson, 1992), observations from other potential partners (reputation/indirect reciprocity: Boyd and Richerson, 1989), or rates of group fission and extinction (group selection: Traulsen and Nowak, 2006).

In the friend/stranger example, cooperation depends on the likelihood of your partner reciprocating in the future. Such reciprocal altruism or direct reciprocity can allow cooperation to evolve because the costs of cooperating can be recouped in the future, when your partner cooperates in return (Trivers, 1971). The likelihood of a stranger in an airport buying a drink in return is much lower than that of a friend in your home town.

The notion of reciprocity has been formalized as a decision strategy called tit-for-tat (TFT). The TFT strategy cooperates in the first interaction with a partner and, for all subsequent interactions, copies the partner's action in the previous interaction (Rapoport and Chammah, 1965). Computer simulations have shown that TFT outcompetes other strategies in a repeated prisoner's dilemma, a game that captures the dilemma between the selfish advantage of defection and the group advantage of cooperation (Axelrod, 1980a). Analytical work has demonstrated that TFT is not invadable by always defecting (ALLD) if the probability of future encounters is high (Axelrod and Hamilton, 1981). Therefore, with repeated future interactions, reciprocity can maintain cooperation.

1.1. Cognitive Building Blocks

TFT has generated a great deal of theoretical and empirical interest because of its simplicity and intuitive nature. It has become the default model of direct reciprocity (Nowak, 2006). Though being simple and intuitive are desirable characteristics of models, they do not necessarily mean that the models represent how humans or other animals actually make decisions.

Simon (1955) admonished that understanding decision making requires integrating the organism's cognitive capacities into the models. Until recently, researchers had not explored whether TFT was cognitively feasible or whether psychological constraints may prevent implementing it (Hammerstein, 2003; Stevens and Hauser, 2004).

One key cognitive building block for TFT is memory. Because TFT strategists copy their partner's previous action, they must store that action in memory. Theoretical work has shown that, when TFT strategists forget their partner's actions and make errors in their choices, cooperation degrades into defection (Molander, 1985). The detrimental effect of memory errors on TFT's performance has inspired the development of alternative strategies that are more robust to memory errors. For instance, tit-for-two-tats (TF2T) allow two defections from the partner before resorting to retaliatory defection (Axelrod, 1980b). Generous tit-for-tat (GTFT) forgives a single defection with a certain probability (Nowak and Sigmund, 1992). Contrite tit-for-tat (CTFT) recognizes its own mistaken defection and corrects with cooperation (Boyd and Richerson, 1989).

Despite this theoretical interest in the effect of memory errors on reciprocal strategies (Molander, 1985; Stephens et al., 1995; Wu and Axelrod, 1995; Rieskamp and Todd, 2006), the memory error rates were not grounded in any kind of empirical data on forgetting. To estimate empirical memory error rates, Stevens et al. (2011) measured human forgetting rates in a cooperative memory task. In this task, participants viewed actions (cooperate or not cooperate) chosen by a set of simulated partners. After viewing all partners once, they then encountered each partner again in a random order, and they were asked to recall whether that partner cooperated or not. Then, the partner's next action was presented (each action was randomly chosen with equal probability). Participants experienced multiple rounds of these encounters with their partners. Stevens et al. (2011) varied the number of rounds and the number of partners between participants. In the various conditions, participants made errors in 10–24% of the trials. Moreover, error rates increased with more intervening events (number of interactions with other partners between two consecutive interactions with a particular partner). Computer simulations suggested that these error rates would likely preclude the evolution of cooperation (Stevens et al., 2011). It is important to highlight that this study aimed to test the cognitive capacities need for TFT specifically rather than cooperation generally.

1.2. Domain Specificity of Cooperative Memory

Stevens et al. (2011) provided the first empirical estimates of forgetting rates for cooperative events. But they did not address whether cooperative memory is "special," i.e., it remains unclear whether cooperative contexts tap specialized, domain-specific cognitive mechanisms that have evolved to deal with the important adaptive problems of cheater detection (Cosmides and Tooby, 1989). Memory, in particular, seems to be adapted to solve key information processing problems (Bjork and Bjork, 1988; Anderson and Schooler, 1991; Schacter, 1999; Nairne and Pandeirada, 2016). For example, work on "survival memory" indicates that people have better memory for items when primed to think about fitness-relevant contexts compared to fitness-irrelevant contexts (Nairne et al., 2007; Nairne and Pandeirada, 2016). Some have suggested that memory may be enhanced in cooperative situations, as well. Early work in this area showed that people had better recognition memory for cheaters compared with cooperators (Mealey et al., 1996). Subsequent work correcting for biases, however, has

failed to show better recognition memory for cheaters (Barclay, 2008; Volstorf et al., 2011; Bell and Buchner, 2012). Yet, people do have better source memory for cheaters. Cheaters are better remembered when they are smiling or viewed as likeable compared with when they are angry or viewed as unlikeable (Bell et al., 2012a). Thus, certain aspects of the decision context (e.g., emotional incongruency) can enhance memory for cheaters (Bell and Buchner, 2012).

Most work on cooperative memory has focused on whether people preferentially remember one action (cooperate or defect) or one reputation type (cooperator or defector). Although researchers have explored how memory for cooperative reputation may differ from memory of other emotionally arousing reputations (Bell et al., 2012b), we do not know whether memory for cooperative events is enhanced relative to memory for other events. This study aims to investigate whether Stevens et al. (2011) tapped specialized, domain-specific memory abilities for events. We investigated this in two ways. First, we explored whether cooperative situations enhance memory relative to other, neutral situations. Though people remember social information better than non-social information (Mesoudi et al., 2006), it is not clear whether they remember cooperative contexts better than non-cooperative contexts. We created two memory contexts. One context replicated Stevens et al.'s (2011) cooperative memory task, where participants had to recall a partner's cooperation or defection. The other context simply had participants track whether each partner read a newspaper or not each day. Better memory performance in the cooperative context would indicate that cooperative situations trigger domain-specific memory mechanisms.

Second, we explored whether memory performance in Stevens et al. (2011) was rather poor because the cooperative memory task did not properly trigger the relevant cooperative memory mechanisms because there was no strategic component to the task. To enhance the strategic nature of the task, we varied how participants were paid based on whether the correct answer was a "positive" action (cooperate or read newspaper) or "negative" action (not cooperate or not read newspaper). Our standard payoff scheme paid participants 5 cents when they were correct and nothing if they were incorrect, replicating Stevens et al. (2011). Our costly payoff scheme (1) provided differential payoffs depending on whether the correct answer was positive or negative and (2) allowed for losses not just gains. Therefore, in the costly payoff scheme, errors are differentially costly, depending on what is remembered. According to error management theory (Haselton and Buss, 2000), people should minimize the most costly error. Because forgetting defection (and not reading the newspaper) is more costly than forgetting cooperation (and reading the newspaper), this mirrors the strategic aspects of the prisoner's dilemma in which cooperating against a defector yields the lowest payoff in the game. Therefore, in the cooperative but not neutral context, better memory in the costly payoff scheme compared to the standard scheme would indicate that adding a strategic component is needed to trigger domain-specific cooperative memory. If memory is better in the costly compared to standard scheme in both the cooperative and neutral contexts, this would indicate that the costs of mistakes drive performance, not strategic cooperation.

1.3. Memory and Social Networks

A key finding from Stevens et al. (2011) is that memory performance decreased with larger groups of simulated partners. That study, however, used the group's sizes of only 5–15 individuals, much smaller than the estimated sizes of actual social networks (Hill and Dunbar, 2003). The social brain or social intelligence hypothesis predicts that cognitive abilities should reflect the level of social complexity experienced by individuals (Jolly, 1966; Humphrey, 1976; Byrne and Whiten, 1988; Dunbar, 1992). This is typically evaluated by correlating traits across species. For example, Dunbar (1992) showed that primate species that live in larger groups have larger relative brain size (neocortex volume relative to rest of brain volume).

Though the social intelligence hypothesis is primarily an evolutionary hypothesis meant to account for species differences in cognition, it may also apply to individual differences in cognition. This would predict a relationship between an individual's cognitive abilities and his or her functional social network size. Since memory is an important cognitive constraint on cooperation, the social intelligence hypothesis predicts that memory performance should correlate with social network size across individuals. Stiller and Dunbar (2007) found that, in fact, memory accuracy for facts did correlate with social network size. We expect this to carry over to cooperative memory situations, as well. Those individuals who have good memories are able to remember cooperation and defection from a larger group of partners. To investigate this prediction, we measured social network sizes of our participants and correlated them with their memory performance.

1.4. This Study

We investigated the domain specificity of cooperative memory and the social intelligence hypothesis by using a cooperative memory task that replicated the methods of Stevens et al. (2011). We varied the context (cooperation/neutral) and payoff scheme (standard/costly) to explore the domain specificity of cooperative memory. We measured the relationship between memory and social network size to explore the social intelligence hypothesis.

2. MATERIALS AND METHODS

We conducted two experiments for this study. The first experiment varied context and payoff scheme. The second experiment aimed to replicate only the observed correlation between memory and social network size.

2.1. Experiment 1
2.1.1. Participants
We recruited 80 participants (39 males and 41 females) with a mean ± SD age of 25.8 ± 4.2 (range 18–36) years from German Universities *via* the Max Planck Institute for Human Development participant pool from January to February 2011. Due to technical problems, the sample was restricted to 48 participants (see below). Participants earned 5 EUR for showing up plus an average of 8.88 EUR (approximately 12.16 USD in 2011) per person (range = 1.60–15.55 EUR) based on their performance. This project was approved by the Max Planck Institute for Human

Development Ethics Commission, and all participants gave written informed consent in accordance with the Declaration of Helsinki.

2.1.2. Materials and Procedures

This study replicated Stevens et al. (2011), so the same methods and descriptions of methods are given here. All materials were prepared in German (translated into English here). The experiment was programmed in E-prime experimental software (Schneider et al., 2002). The first part of the experiment collected demographic information from the participants (sex, age, educational level, occupation, and college major). For the remainder of the experiment, participants engaged in a memory task in which they viewed simulated partners, and they were asked to recall each partner's previous action (Stevens et al., 2011). Before beginning the experiment, participants received a paper copy of instructions describing the goal of the task: recall the last action for each simulated player (participant instructions are available in the Supplementary Material).

A practice phase familiarized participants with the experiment. The practice phase was identical to the actual experimental session, except the following: (1) it used fewer trials in a fixed order for all participants (three partners with four interactions each and six partners with three interactions each), (2) it included only female partners (the experimental phases included only male partners), and (3) the money earned did not accumulate for the final payment. At the end of the practice session, participants received feedback concerning their success ("You have accomplished the practice session with x out of 21 correct answers.").

Simulated partners included a photograph of an individual and a name (Figure S1 in Supplementary Material). For the photographs of partners, we used images from Ebner (2008) downloaded from the MPI for Human Development FACES Collection (retrieved from http://faces.mpib-berlin.mpg.de on 21 March 2011). We used 9 images of females for the practice phase and 20 images of males for the experimental phase. The depicted persons ranged between 18 and 32 years old, with the same background and color of clothing (Ebner, 2008). For partner names, we used 20 of the most common male German names from 1958 to 2000 (retrieved from http://www.gfds.de/vornamen/beliebteste-vornamen/ on 21 March 2011).

Each session consisted of a series of rounds. Participants met with each of 10 partners once in a randomized order per round. In the initial round, we presented individually for each partner an image, a name, and an action: for instance, "Klaus cooperates" or "Ulrich reads the newspaper" (Figure S1 in Supplementary Material). We randomly assigned partner names and images across participants. Participants viewed each partner's information for 5 s before advancing to the next partner (1 s in between partners). For every trial in the experimental phase, we randomly assigned the partner's action as cooperate or defect, so participants could not associate a pattern of action with each partner and had to track the exact behavior of each partner in the previous round. Due to technical problems, however, partner choices were not completely random for 32 participants. Non-random partner choices could allow participants to perform better than chance just by using

the skewed base rates. Therefore, we removed from the analysis participants whose mean proportion of choices (aggregated across all partners) was lower than 0.4 and higher than 0.6 for either experimental condition. We chose these boundaries (before analyzing the data) because they are reasonably close to the truly random value of 0.5. Also, when viewing the data, there are natural breakpoints in the distributions at these values (Figure S2 in Supplementary Material).

After viewing all members of the group, participants began the retrieval rounds, with a randomized order of partners in each round. We presented the image of the partners, along with the question "What did [name] do last time?" The participant had 10 s to answer by pressing buttons on the keyboard. If they responded within 10 s, they received a feedback screen for 3 s stating whether they were correct, the amount of money they received for that trial (only if they were correct), and an updated total amount received so far in the experiment. If they failed to respond in time, the participant did not receive feedback, only a reminder to respond more quickly next time. After the feedback screen, participants viewed the new action of the current partner for 5 s before advancing to the next partner. In between rounds, participants could pause the program and start a new round at their discretion. Afterward, participants completed a questionnaire asking what kinds of strategies they used to solve the memory task, how often they guessed, and how often they thought the partners cooperated.

2.1.3. Experimental Conditions

We manipulated two within-subjects experimental conditions and one between-subjects experimental condition, leading to a $2 \times 2 \times 2$ experimental design. The first within-subjects manipulation varied the context experienced by the participant. The *cooperative context* condition replicated Stevens et al. (2011) by having partners either cooperate or not cooperate. To introduce the idea of cooperation, participants first read the following text:

> You are supposed to prepare two-person presentations for a high school class together with different partners. In the end, you and your partner will receive the same mark. Each of your partners can now decide: either he will invest time and effort and cooperate or he leaves you alone with the preparations and does not cooperate.

The participant then began the practice trials for that condition, tracking the cooperative actions of their partners. In the *neutral context* condition, rather than tracking cooperation, participants tracked a neutral action that involved no strategic components. In other words, the partner actions did not influence the participant. Before beginning this condition, the participants read the following text:

> You come in contact with different people. They repeatedly decide whether they want to read a newspaper this morning over breakfast or not.

All participants experienced 10 rounds of each condition (one initial round observing actions and nine recall rounds), with

the order counterbalanced across participants (alternating with subject number) and evenly divided between sexes.

The second within-subjects manipulation varied the partner's action. We defined positive actions as cooperating and reading the newspaper and negative actions as not cooperating and not reading the newspaper. For each trial, the computer program randomly chose the positive or negative action with equal probability. Therefore, we could examine whether the type of partner action influenced memory accuracy.

The between-subjects manipulation varied the payoff scheme offered to the participants (**Table 1**). The *standard payoff scheme* was identical to that used in Stevens et al. (2011) in which participants received 5 euro cents if they answered correctly and 0 euro cents if they answered incorrectly, regardless of the correct answer. In the *costly payoff scheme*, the payoffs for correct and incorrect answers depended on the correct answer (**Table 1**). That is, if the partner action was positive (cooperate or read the newspaper), then the participant received 10 euro cents for correct answers and 0 euro cents for incorrect answers. If the partner action was negative (not cooperate or not read the newspaper), then correct answers yielded 5 cents and incorrect answers resulted in losing 5 cents. We counterbalanced which condition participants experienced (alternating with participant number), and evenly divided conditions between sexes.

2.1.4. Questionnaire
After the memory task, participants completed a questionnaire in which we asked them whether they knew any of the people in the photographs, how often they guessed, what types of strategies they used to solve the task, and task motivation. We also asked participants how many friends they had in social networking websites (such as Facebook) and how many family members, close friends, acquaintances, work colleagues, and neighbors they had. We summed all of these contacts as our measure of social network size. The questionnaire is available in the Supplementary Material.

2.2. Experiment 2
2.2.1. Participants
Experiment 2 was a replication study in which we focused on cooperative memory and social network size. We tested 80 participants (30 males, 48 females, and 2 others) with a mean ± SD age of 20.4 ± 1.4 (range 19–25) years from the University of Nebraska-Lincoln Department of Psychology undergraduate participant pool in April 2016. Participants earned course credit

for showing up plus an average of 3.35 USD (range = 2.15–4.65 USD) per person based on their performance. This project was approved by the University of Nebraska-Lincoln Internal Review Board (IRB# 20160316008), and all participants gave written informed consent in accordance with the Declaration of Helsinki.

2.2.2. Materials and Procedures
This experiment replicated the methods of Experiment 1 but only with the cooperation scenario using the standard payoff scheme. Participants experienced 11 interactions with each of 10 hypothetical partners. Because this was an American population, we used 29 of the most popular male names in the U.S. during the 1990s (when the participants were born) from the U.S. Social Security Administration (https://www.ssa.gov/OACT/babynames/ retrieved on 4 April 2016). We also used 29 images of neutral male faces and 3 images of neutral female faces (for practice trials) from the Umeå University Database of Facial Expressions (Samuelsson et al., 2012).

2.2.3. Questionnaire
After the memory task, we asked participants about their social network. Instead of estimating contacts, Facebook friends, etc., we had participants list the initials of everyone with whom they had some kind of social contact in the last 30 days (Lewis et al., 2011). We used this number of contacts as our measure of social network size. We then collected demographic data, such as gender and age.

2.3. Data Analysis
We conducted a binomially distributed generalized linear mixed model (GLMM) with memory accuracy as the binary response variable (0 for incorrect and 1 for correct). We included context (cooperation or neutral) and the partner's previous action (cooperate/read newspaper or defect/did not read newspaper) as within-subjects variables and payoff scheme (standard or costly) as a between-subjects variable. For between-subjects comparisons, we calculated standard 95% confidence intervals; for within-subjects comparisons, we calculated within-subjects 95% confidence intervals (Morey, 2008); and for mixed-effects models, we calculated profile likelihood confidence intervals for coefficients (Venzon and Moolgavkar, 1988). Confidence intervals are presented in brackets after the parameter estimate.

In addition to frequentist statistics, we calculated Bayes factors (*BF*) to provide the weight of evidence for the alternative hypothesis relative to the null hypothesis (Wagenmakers, 2007). For example, *BF* = 10 means that the evidence for the alternative hypothesis is 10 times stronger than the evidence for the null hypothesis. Bayes factors between 0.33 and 3 provide only anecdotal evidence, those between 0.1–0.33 and 3–10 provide moderate evidence, those between 0.01–0.1 and 10–100 provide strong evidence, and those below 0.01 or above 100 provide very strong evidence (Wagenmakers et al., in press). Bayes factors associated with GLMMs were converted from Bayesian Information Criterion (BIC) using $BF = e^{\frac{BIC_{null} - BIC_{alternative}}{2}}$ (Wagenmakers, 2007). Alternative models for main effects included only the main effect of interest and the random subject effect. Alternative models for

TABLE 1 | Payoff schemes.

a. Standard payoff condition

	Positive action	Negative action
Correct response	5 cents	5 cents
Incorrect response	0 cents	0 cents

b. Costly payoff condition

	Positive action	Negative action
Correct response	10 cents	5 cents
Incorrect response	0 cents	−5 cents

interactions included the interaction of the two main effects and the random subject effect. Null models for main effects included only the random subject effect, and null models for interactions included the two main effects (without the interaction) and the random effect of subject. All other Bayes factors were computed using non-informative (Bayesian t-test) or weakly informative (Bayesian regression/ANOVA) priors (Rouder et al., 2009; Rouder and Morey, 2012).

We analyzed the data using R Statistical Software version 3.4.1 (R Core Team, 2017) and the `BayesFactor` (Morey and Rouder, 2015), `cowplot` (Wilke, 2016), `dplyr` (Wickham and Francois, 2016), `ggplot2` (Wickham, 2016), `lme4` (Bates et al., 2015), `papaja` (Aust and Barth, 2017), `plyr` (Wickham, 2011), and `tidyr` (Wickham, 2017) packages. Data, R code, and supplementary figures are available in the Supplementary Material and at the Open Science Framework (https://osf.io/zcv4m/). The manuscript was created using `rmarkdown` (Allaire et al., 2017) and `knitr` (Xie, 2015), and the reproducible research materials are available from author JRS and at https://osf.io/zcv4m/.

3. RESULTS

3.1. Replication of Stevens et al. (2011)

The overall memory accuracies (collapsed across all conditions) from Experiment 1 (mean $\pm$ SD = 81.2 $\pm$ 10.3%) compared with those of Stevens et al. (2011) (77.5 $\pm$ 10.5%) showed no evidence of a difference (mean difference [95% CI] = 3.7% [−1.2, 8.7], Cohen's d = 0.35, BF = 0.64). Experiment 2, however, resulted in lower accuracies (67.1 $\pm$ 12.2%) than Stevens et al. (2011) (mean difference = 10.5% [6.8, 14.1], Cohen's d = 0.86, BF > 100) and Experiment 1 (mean difference = 14.2% [9.1, 19.2], Cohen's d = 1.17, BF > 100). The effect of the number of intervening

FIGURE 1 | Effect of number of intervening events on memory accuracy. Memory accuracy decreased as the number of intervening events increased. Data from Stevens et al. (2011) used comparable parameter values as Experiments 1 and 2 (10 partners and 10 interactions per partner).

events on memory showed similar decreases (with different intercepts) for all data sets (**Figure 1**). Thus, the data from these experiments broadly replicated Stevens et al. (2011).

3.2. Payoff Scheme and Context

For Experiment 1, we conducted a binomial GLMM using memory accuracy as a dependent variable, payoff scheme as a between-subjects variable, and context and partner action as within-subjects variables. Participants remembered 81.5% [77.3, 85.6] of the actions in the standard payoff scheme and 80.8% [76.4, 85.2] in the costly scheme, providing strong evidence that accuracy did not differ between payoff schemes (β = 0.08 [−0.36, 0.52], BF = 0.01; **Figure 2A**).

Participants remembered 80.9% [78.2, 83.6] of the actions in the cooperation context and 81.4% [78.7, 84.0] in the neutral context, providing strong evidence that accuracy did not differ across context (β = 0.26 [0.04, 0.47], BF = 0.01; **Figure 2B**). Therefore, participants remembered partner actions equally independent of payoff scheme or context.
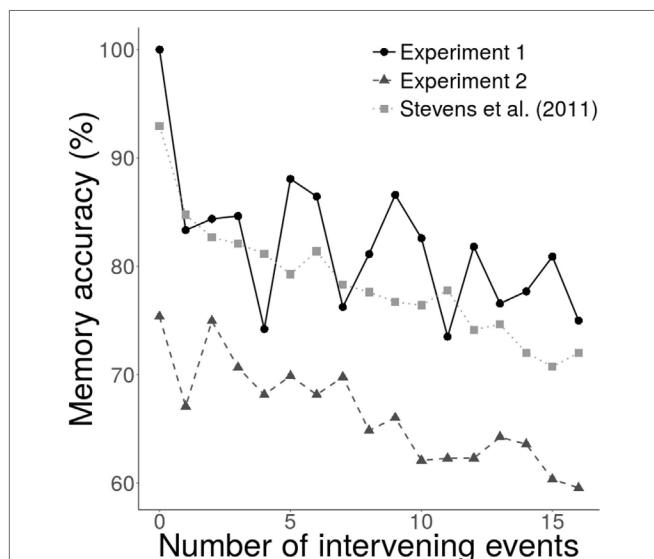
3.3. Partner Action Type

For Experiment 1, participants remembered 79.2% [76.4, 81.9] of the partners' positive actions (cooperate/read newspaper) and 83.0% [80.3, 85.7] of their negative actions (defect/did not read newspaper), which was very strong evidence for a difference between the two partner action types (β = 0.51 [0.28, 0.73], BF > 100; **Figure 3**). There was strong evidence for no interaction between partner action type and payoff scheme (β = −0.15 [−0.47, 0.17], BF = 0.06) and no evidence for an interaction between partner action type and context (β = −0.26 [−0.58, 0.07], BF = 0.56). Therefore, participants remembered defection better than cooperation across the two payoff schemes, but they also remembered not reading a newspaper better than reading a newspaper.
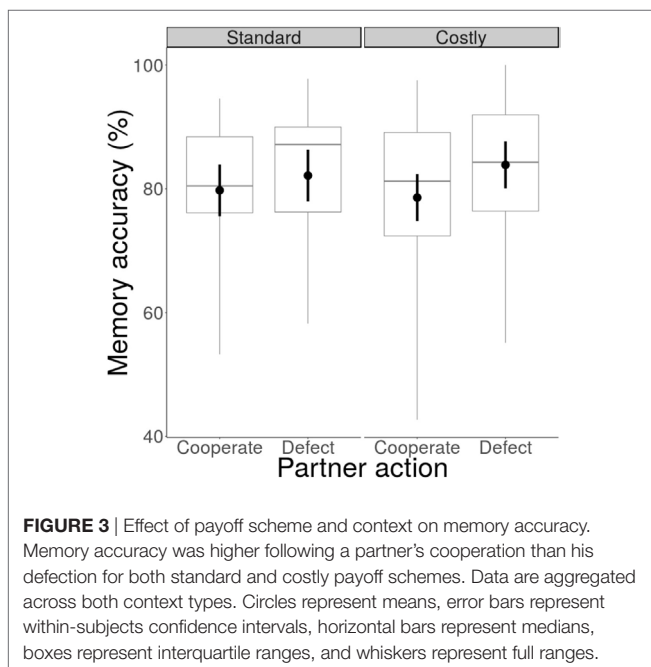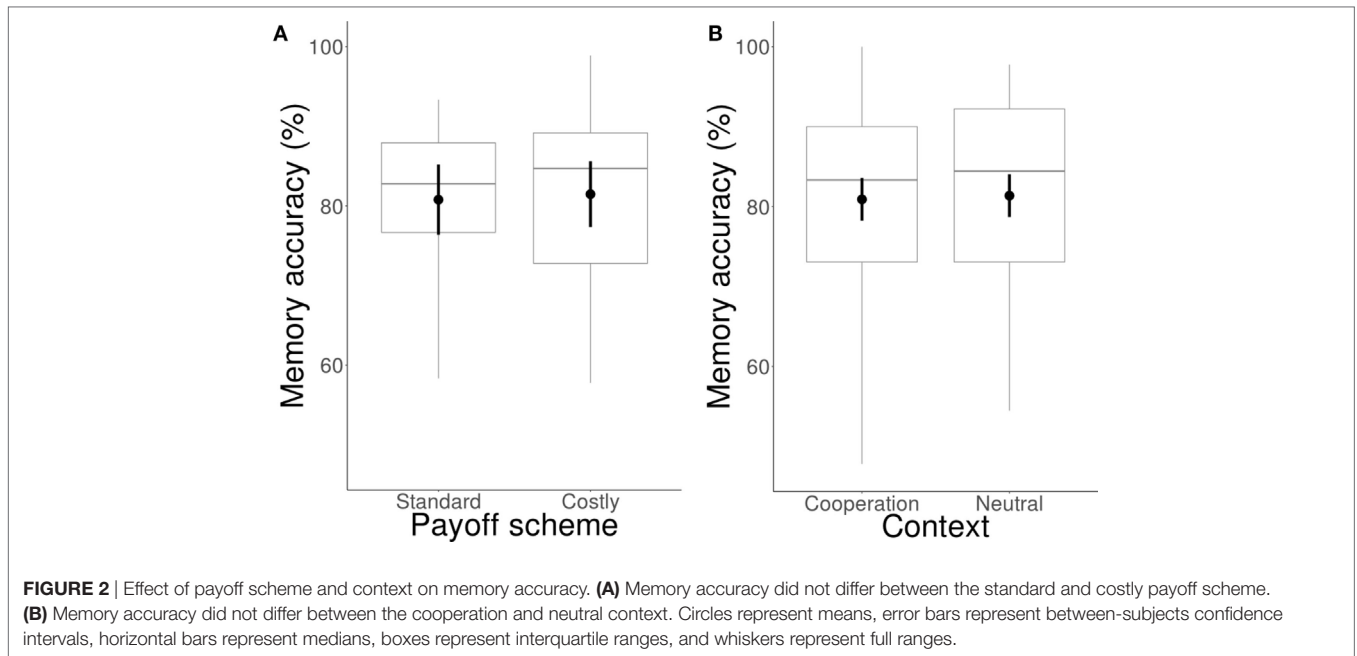
3.4. Social Network Size

For Experiment 1, there was no evidence for memory accuracy correlating with social network size (r_{46} = 0.09, BF = 0.34); however, these data included outliers (**Figure 4A**). As an exploratory analysis, we removed the two data points with high Z-scores (greater than 3.5) and high leverage. The trimmed data showed moderate evidence for a positive relationship between memory accuracy and social network size (r_{44} = 0.35, BF = 3.20; **Figure 4B**).

The replication study (Experiment 2) tested a different group of participants and used an alternative measure of social network size that recorded the number of individuals that participants encountered in the last 30 days. This confirmatory data set replicated the moderate positive relationship between memory accuracy and social network size (r_{78} = 0.32, BF = 9.17; **Figure 4C**). Note that non-informative priors were used to estimate this Bayes factor, so this value is independent of the analysis from Experiment 1 (i.e., we did not use the posterior distribution from Experiment 1 as the priors for Experiment 2). Therefore, individuals with larger social networks remembered better than those with smaller networks.

FIGURE 2 | Effect of payoff scheme and context on memory accuracy. **(A)** Memory accuracy did not differ between the standard and costly payoff scheme. **(B)** Memory accuracy did not differ between the cooperation and neutral context. Circles represent means, error bars represent between-subjects confidence intervals, horizontal bars represent medians, boxes represent interquartile ranges, and whiskers represent full ranges.

FIGURE 3 | Effect of payoff scheme and context on memory accuracy. Memory accuracy was higher following a partner's cooperation than his defection for both standard and costly payoff schemes. Data are aggregated across both context types. Circles represent means, error bars represent within-subjects confidence intervals, horizontal bars represent medians, boxes represent interquartile ranges, and whiskers represent full ranges.
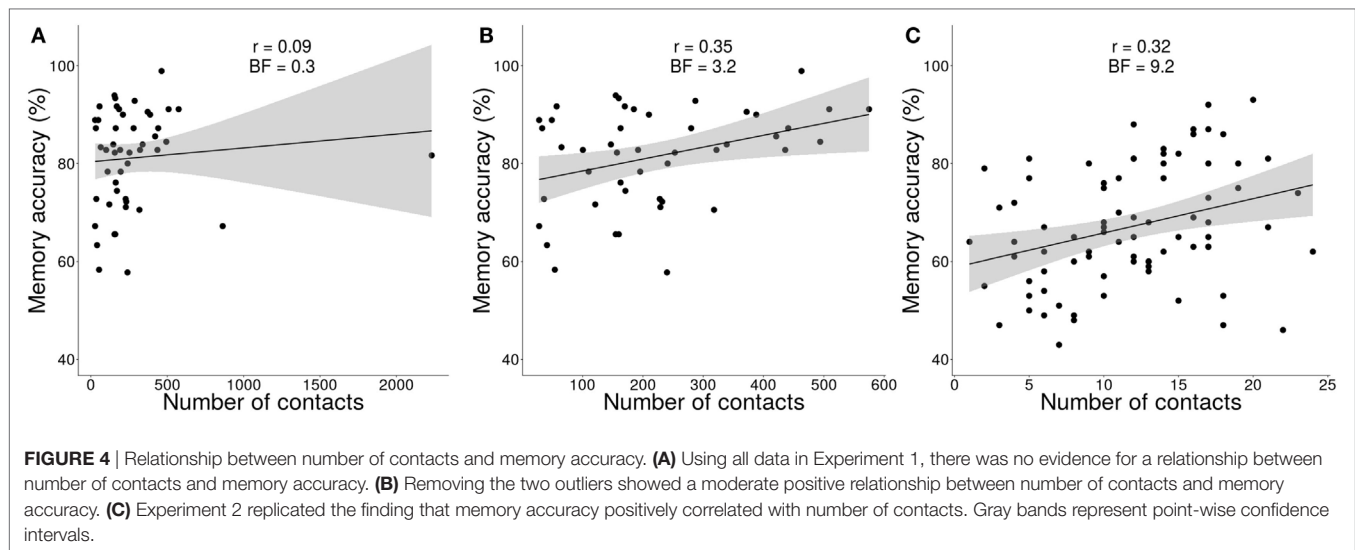
4. DISCUSSION

Our aims in this study were to (1) replicate findings of Stevens et al. (2011), (2) determine whether more strategic payoff situations induce better memory than standard payoffs, (3) determine whether cooperative contexts induce better memory than neutral contexts, and (4) assess whether memory accuracy positively correlates with number of social contacts. In general, our findings replicated those of Stevens et al. (2011) with memory accuracy

decreasing with the number of intervening events (**Figure 1**). To investigate whether a more strategic payoff situation induces better memory than standard payoffs, we made forgetting defection costly and correctly remembering cooperation more valuable than remembering defection (**Table 1**). Despite this difference in payoffs, participants exhibited the same memory accuracy levels (**Figure 2A**). To investigate whether cooperative contexts induce better memory than neutral contexts, we provided contexts framed as either cooperative or not cooperative (recalling whether a partner read a newspaper). Participants did not differ in their memory accuracy levels between the two contexts (**Figure 2B**). Finally, we found that memory accuracy did positively correlate with number of social contacts in two experiments (**Figure 4**). Thus, we did not find evidence of cooperative memory being special, but we did demonstrate a relationship between cooperative memory and social network size.

4.1. Is Cooperation Special?

A potential criticism of Stevens et al.'s (2011) experiment is that the memory task did not engage purported cooperation-specific cognition because there was no strategic component. To address this, we included a condition with a payoff structure that mirrored a prisoner's dilemma payoff scheme, where the reward for correctly remembering cooperation was enhanced and the cost to forgetting defection was enhanced such that the participants lost money. This created a signal detection problem with differential costs for false alarms and misses. If individuals minimize the total number of errors committed, there should be no difference between the payoff schemes. If individuals maximize expected value, there should be no difference between remembering cooperation and defection for the standard payoff scheme, but a bias toward remembering cooperation in the costly scheme. If individuals are minimizing risk (most costly error), there should be no difference

FIGURE 4 | Relationship between number of contacts and memory accuracy. **(A)** Using all data in Experiment 1, there was no evidence for a relationship between number of contacts and memory accuracy. **(B)** Removing the two outliers showed a moderate positive relationship between number of contacts and memory accuracy. **(C)** Experiment 2 replicated the finding that memory accuracy positively correlated with number of contacts. Gray bands represent point-wise confidence intervals.

between remembering cooperation and defection for the standard scheme, but a bias toward remembering defectors in the costly scheme. Our results show a bias for remembering defection better than cooperation, but it applies to both the standard and costly payoff schemes (**Figure 3**). This contradicts the predictions of signal detection theory because the payoff schemes differed in their costs, which should influence performance. Despite contradicting the theories, these results align with many other tests of signal detection theory, where participants ignore costs of errors when detecting signals among noise (Bohil and Maddox, 2001; Lynn and Barrett, 2014).

Signal detection theory does not specify the mechanism by which detection thresholds are set; it only provides optimal solutions to detection problems. The assumption, however, is that individuals are learning about the probability of the different states of the world (in our case, partners cooperating or defecting) and the costs and benefits of errors and accurate judgments. Thus, within a session, participants should adjust to the probabilities and payoffs. We did not find this when aggregating over the entire experimental session. As an exploratory analysis, we tested memory accuracy for cooperation and defection for only the cooperation context as a function of payoff scheme for the last 40 trials per condition but found the same results.

Though a learning-based interpretation of signal detection theory predicts sensitivity to payoffs over the course of an experiment, error management theory (Haselton and Buss, 2000) does not. Rather, error management theory predicts that natural selection is the mechanism that sets detection thresholds. That is, detection thresholds must have some genetic basis, and individuals whose thresholds provide fitness benefits are more likely to pass on those genetic predispositions to their offspring. Therefore, detection thresholds may not be dynamic under this theory. They may be set to values that provide overall benefits. The fact that participants remembered defectors better than cooperators supports the error management theory prediction of thresholds that minimize the most costly error. This result also aligns with other work demonstrating preferential memory for

defectors over cooperation (Buchner et al., 2009). Nevertheless, the lack of a difference between the two payoff schemes suggests that making the Stevens et al. (2011) experimental situation more strategic did not trigger special cognition for remembering cooperative situations.

One possible explanation for the failure of the payoff scheme to influence memory accuracy is that the original task by Stevens et al. (2011) already triggered the cooperation-specific cognitive mechanisms, so accuracy was at ceiling for both payoff schemes. If this is the case, then we would expect participants to perform better in the cooperation context compared to the neutral context. Yet, we found no difference in performance between these contexts. Even the effect of better memory of defection appears in both cooperative and neutral contexts. This is unexpected because there is no *a priori* reason to remember that someone did not read a newspaper better than remembering that they did read a newspaper. One possibility is that people remember negative events (not cooperating and not reading a newspaper) better than positive ones as is observed with words and images (Robinson-Riegler and Winton, 1996; Ito et al., 1998; Baumeister et al., 2001). This negativity bias may apply across domains. Combined with the lack of a difference between payoff schemes, these results indicate that memory for cooperative events is domain general and does not differ from non-cooperative contexts or more strategic cooperative situations. This aligns with previous work demonstrating the generality of memory beyond cooperation (Bell and Buchner, 2010; Bell et al., 2015).

4.2. Social Network Size
We found that individual social network size correlated with memory for actions. This corroborates another finding that social network size correlated with memory for facts (Stiller and Dunbar, 2007). Our results support the social intelligence hypothesis that predicts a relationship between cognition and social group size. Interestingly, these results also suggest that this is not just an evolutionary hypothesis about species differences. Our results indicate that this relationship holds across individuals

within a species, as well. Further, recent work has explored the neural basis for these individual differences and has reported correlations between network size and brain region size/density (Bickart et al., 2011; Lewis et al., 2011; Kanai et al., 2012; Powell et al., 2012), activity (Von Der Heide et al., 2014), and connectivity (Bickart et al., 2012; Hampton et al., 2016).

The social intelligence hypothesis, however, is agnostic to whether group size limits cognition or whether cognition allows expansion of group size. Constraints on cognition for an individual may limit the social network size in which he or she may function. Alternatively, interacting in larger social networks may facilitate enhanced cognition. Though the direction of causation can probably work both ways, there is evidence in rhesus macaques that social network size can drive brain size (Sallet et al., 2011). Researchers experimentally housed macaques in groups of different sizes and found that both gray matter and brain region connectivity increased with group size. This result has important implications for our findings. It implies that, over the course of an individual's lifetime, the exposure to a complicated social environment (large social network) can drive changes in brain region size and connectivity. These neural enhancements may expand cognitive abilities for individuals in large social networks. Memory may not be so much of a constraint on social network size. Rather, the social environment may actually shape cognition. Thus, engaging in social interactions with many partners may improve memory for cooperative actions.

4.3. Limitations and Conclusion

One limitation of our method is that the measures of social network size involve memory (recalling the number of Facebook friends and recalling individual people in their social networks). Thus, to some extent, memory is intertwined with our measure of network size. More direct measures of network size (e.g., experimenters directly viewing Facebook accounts or using experienced-based sampling of social networks) can avoid this issue in future work.

In Experiment 1, our technical problems reduced our sample size from 80 to 48 participants. Though the smaller sample size reduced our power, the sample size was still large enough to provide strong evidence supporting the null hypothesis of no difference between payoff scheme conditions or contexts. Thus, the Bayesian statistics give us confidence to accept the null hypotheses of no difference.

We observed differences between our German and American populations of participants. The error rates observed in our German population matched those of the Stevens et al. (2011) data, which were drawn from a similar German population. The American participants exhibited higher error rates than the German participants. Nevertheless, even with a lower accuracy rate, the American population replicated the positive relationship between number of social contacts and memory accuracy.

In terms of the experimental design, like Stevens et al. (2011), this study presented hypothetical actions of hypothetical partners. Ideally, participants should be engaged in actual cooperative interactions with real partners. Due to the difficulty of simultaneously incentivizing both cooperative interactions and accurate memory recall, we opted to incentivize only memory recall. To make the task more realistic from a strategic perspective, we used the costly payoff scheme to mimic a prisoner's dilemma. Nevertheless, including direct cooperative interactions would further increase the external validity of these results.

This was a fairly fast-paced memory experiment. Participants had 10 s to recall and choose a partner action, 5 s of exposure to new partner actions, and 1 s between partners. Of course, this interaction rate does not match natural interaction patterns that people experience. Decision time, encoding time, and time between interactions were greatly accelerated in this experiment. Similarly, this experiment lacked key emotional components of actual cooperation situations. However, this experiment was not designed to test which cognitive and emotional component are used in actual cooperative behavior. Rather, this experiment and Stevens et al. (2011) were designed to test whether humans have the memory abilities to implement tit-for-tat-like strategies. TFT is agnostic on the pace of interaction or the presence of emotional input. This study in particular aimed to see if the results in Stevens et al. (2011) were robust to changes in the cooperative situation, by attempting to both enhance and reduce memory for cooperation. Both efforts failed, suggesting that cooperative memory is not special. We do not have the memory abilities to implement tit-for-tat-like strategies.

Humans do not accurately remember specific cooperative actions. We do, however, remember overall partner impressions (Volstorf et al., 2011). That is, though we do not track individual actions (cooperating or defecting), we do integrate a series of actions into an overall impression and remember these impressions (cooperator or defector) over time. We argue that focusing on tit-for-tat-like strategies has provided important insights into the evolution of cooperation. But these strategies are simply not cognitively plausible. We recommend that modeling efforts redirect to strategies that more accurately reflect how humans and other animals actually process and remember information. We advocate a switch from action-based to impression-based strategies. These strategies should incorporate not only personal experience but also third-party observation of reputation. The reputation perspective already has a modeling history (Boyd and Richerson, 1989; Nowak and Sigmund, 1998; Leimar and Hammerstein, 2001; Panchanathan and Boyd, 2004; Roberts, 2008), and incorporating personal experience impressions could provide important insights and yield a parsimonious model of cooperation. Critically, this research program takes the psychology of the decision maker seriously, satisfying Simon's (1955) notion of bounded rationality. The modeling world has carefully investigated how the environment influences cooperation. But to truly resolve the puzzle of cooperation, we must understand the cognition of the decision makers.

ETHICS STATEMENT

This study was carried out in accordance with the recommendations of the Max Planck Institute for Human Development Ethics Commission and the University of Nebraska-Lincoln Internal Review Board with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the Max Planck Institute for Human Development Ethics Commission

and the University of Nebraska-Lincoln Internal Review Board (IRB#20160316008).

AUTHOR CONTRIBUTIONS

ACKNOWLEDGMENTS

FUNDING

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at http://www.frontiersin.org/article/10.3389/frobt.2017.00052/full#supplementary-material.

REFERENCES

Allaire, J., Cheng, J., Xie, Y., McPherson, J., Chang, W., Allen, J., et al. (2017). *rmarkdown: Dynamic Documents for R*. Available at: https://CRAN.R-project.org/package=rmarkdown

Anderson, J. R., and Schooler, L. J. (1991). Reflections of the environment in memory. *Psychol. Sci.* 2, 396–408. doi:10.1111/j.1467-9280.1991.tb00174.x

Aust, F., and Barth, M. (2017). *papaja: Create APA Manuscripts with R Markdown*. Available at: https://github.com/crsh/papaja

Axelrod, R. (1980a). Effective choice in the prisoner's dilemma. *J. Confl. Resolut.* 24, 3–25. doi:10.1177/002200278002400101

Axelrod, R. (1980b). More effective choice in the prisoner's dilemma. *J. Confl. Resolut.* 24, 379–403. doi:10.1177/002200278002400301

Axelrod, R., and Hamilton, W. D. (1981). The evolution of cooperation. *Science* 211, 1390–1396. doi:10.1126/science.7466396

Barclay, P. (2008). Enhanced recognition of defectors depends on their rarity. *Cognition* 107, 817–828. doi:10.1016/j.cognition.2007.11.013

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48. doi:10.18637/jss.v067.i01

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., and Vohs, K. D. (2001). Bad is stronger than good. *Rev. Gen. Psychol.* 5, 323–370. doi:10.1037/1089-2680.5.4.323

Bell, R., and Buchner, A. (2010). Valence modulates source memory for faces. *Mem. Cognit.* 38, 29–41. doi:10.3758/MC.38.1.29

Bell, R., and Buchner, A. (2012). How adaptive is memory for cheaters? *Curr. Dir. Psychol. Sci.* 21, 403–408. doi:10.1177/0963721412458525

Bell, R., Buchner, A., Kroneisen, M., and Giang, T. (2012a). On the flexibility of social source memory: a test of the emotional incongruity hypothesis. *J. Exp. Psychol. Learn. Mem. Cogn.* 38, 1512–1529. doi:10.1037/a0028219

Bell, R., Giang, T., and Buchner, A. (2012b). Partial and specific source memory for faces associated to other- and self-relevant negative contexts. *Cogn. Emotion* 26, 1036–1055. doi:10.1080/02699931.2011.633988

Bell, R., Mieth, L., and Buchner, A. (2015). Appearance-based first impressions and person memory. *J. Exp. Psychol. Learn. Mem. Cogn.* 41, 456–472. doi:10.1037/xlm0000034

Bickart, K. C., Hollenbeck, M. C., Barrett, L. F., and Dickerson, B. C. (2012). Intrinsic amygdala-cortical functional connectivity predicts social network size in humans. *J. Neurosci.* 32, 14729–14741. doi:10.1523/JNEUROSCI.1599-12.2012

Bickart, K. C., Wright, C. I., Dautoff, R. J., Dickerson, B. C., and Barrett, L. F. (2011). Amygdala volume and social network size in humans. *Nat. Neurosci.* 14, 163–164. doi:10.1038/nn.2724

Bjork, E. L., and Bjork, R. A. (1988). "On the adaptive aspects of retrieval failure in autobiographical memory," in *Practical Aspects of Memory: Current Research and Issues*, eds M. M. Gruneberg, P. E. Morris, and R. N. Sykes (Oxford, England: John Wiley & Sons), 283–288.

Bohil, C. J., and Maddox, W. T. (2001). Category discriminability, base-rate, and payoff effects in perceptual categorization. *Percept. Psychophys.* 63, 361–376. doi:10.3758/BF03194476

Boyd, R., and Richerson, P. J. (1989). The evolution of indirect reciprocity. *Soc. Networks* 11, 213–236. doi:10.1016/0378-8733(89)90003-8

Boyd, R., and Richerson, P. J. (1992). Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethol. Sociobiol.* 13, 171–195. doi:10.1016/0162-3095(92)90032-Y

Buchner, A., Bell, R., Mehl, B., and Musch, J. (2009). No enhanced recognition memory, but better source memory for faces of cheaters. *Evol. Hum. Behav.* 30, 212–224. doi:10.1016/j.evolhumbehav.2009.01.004

Byrne, R. W., and Whiten, A. (1988). *Machiavellian Intelligence: Social Expertise and the Evolution of Intellect in Monkeys, Apes, and Humans*. Oxford: Oxford University Press.

Cosmides, L., and Tooby, J. (1989). Evolutionary psychology and the generation of culture: II. Case study: a computational theory of social exchange. *Ethol. Sociobiol.* 10, 51–97. doi:10.1016/0162-3095(89)90013-7

Dunbar, R. I. M. (1992). Neocortex size as a constraint on group size in primates. *J. Hum. Evol.* 22, 469–493. doi:10.1016/0047-2484(92)90081-J

Ebner, N. C. (2008). Age of face matters: age-group differences in ratings of young and old faces. *Behav. Res. Methods* 40, 130–136. doi:10.3758/BRM.40.1.130

Hamilton, W. D. (1964). The genetical evolution of social behaviour. I, II. *J. Theor. Biol.* 7, 1–52. doi:10.1016/0022-5193(64)90038-4

Hammerstein, P. (2003). "Why is reciprocity so rare in social animals? A protestant appeal," in *Genetic and Cultural Evolution of Cooperation Dahlem Workshop Reports*, ed. P. Hammerstein (Cambridge, MA: MIT Press), 83–94.

Hampton, W. H., Unger, A., Von Der Heide, R., and Olson, I. R. (2016). Neural connections foster social connections: a diffusion-weighted imaging study of social networks. *Soc. Cogn. Affect. Neurosci.* 11, 721–727. doi:10.1093/scan/nsv153

Haselton, M. G., and Buss, D. M. (2000). Error management theory: a new perspective on biases in cross-sex mind reading. *J. Pers. Soc. Psychol.* 78, 81–91. doi:10.1037/0022-3514.78.1.81

Hill, R. A., and Dunbar, R. I. M. (2003). Social network size in humans. *Hum. Nat.* 14, 53–72. doi:10.1007/s12110-003-1016-y

Humphrey, N. K. (1976). "The social function of intellect," in *Growing Points in Ethology*, eds P. P. G. Bateson and R. A. Hinde (Cambridge: Cambridge University Press), 303–317.

Ito, T. A., Larsen, J. T., Kyle, N., and Cacioppo, J. T. (1998). Negative information weighs more heavily on the brain: the negativity bias in evaluative categorizations. *J. Pers. Soc. Psychol.* 75, 887–900. doi:10.1037/0022-3514.75.4.887

Jolly, A. (1966). Lemur social behavior and primate intelligence. *Science* 153, 501–506. doi:10.1126/science.153.3735.501

Kanai, R., Bahrami, B., Roylance, R., and Rees, G. (2012). Online social network size is reflected in human brain structure. *Proc. R. Soc. Lond. B* 279, 1327–1334. doi:10.1098/rspb.2011.1959

Leimar, O., and Hammerstein, P. (2001). Evolution of cooperation through indirect reciprocity. *Proc. R. Soc. Lond. B* 268, 745–753. doi:10.1098/rspb.2000.1573

Lewis, P. A., Rezaie, R., Brown, R., Roberts, N., and Dunbar, R. I. M. (2011). Ventromedial prefrontal volume predicts understanding of others and social network size. *Neuroimage* 57, 1624–1629. doi:10.1016/j.neuroimage.2011.05.030

Lynn, S. K., and Barrett, L. F. (2014). "Utilizing" signal detection theory. *Psychol. Sci.* 25, 1663–1673. doi:10.1177/0956797614541991

Mealey, L., Daood, C., and Krage, M. (1996). Enhanced memory for faces of cheaters. *Ethol. Sociobiol.* 17, 119–128. doi:10.1016/0162-3095(95)00131-X

Mesoudi, A., Whiten, A., and Dunbar, R. (2006). A bias for social information in human cultural transmission. *Br. J. Psychol.* 97, 405–423. doi:10.1348/000712605X85871

Molander, P. (1985). The optimal level of generosity in a selfish, uncertain environment. *J. Confl. Resolut.* 29, 611–618. doi:10.1177/0022002785029004004

Morey, R. D. (2008). Confidence intervals from normalized data: a correction to Cousineau (2005). *Tutorial Quant. Methods Psychol.* 4, 61–64. doi:10.20982/tqmp.04.2.p061

Morey, R. D., and Rouder, J. N. (2015). *BayesFactor: Computation of Bayes Factors for Common Designs*. Available at: https://CRAN.R-project.org/package=BayesFactor

Nairne, J. S., and Pandeirada, J. N. S. (2016). Adaptive memory: the evolutionary significance of survival processing. *Perspect. Psychol. Sci.* 11, 496–511. doi:10.1177/1745691616635613

Nairne, J. S., Thompson, S. R., and Pandeirada, J. N. S. (2007). Adaptive memory: survival processing enhances retention. *J. Exp. Psychol. Learn. Mem. Cogn.* 33, 263–273. doi:10.1037/0278-7393.33.2.263

Nowak, M. A. (2006). Five rules for the evolution of cooperation. *Science* 314, 1560–1563. doi:10.1126/science.1133755

Nowak, M. A., and Sigmund, K. (1992). Tit-for-tat in heterogeneous populations. *Nature* 355, 250–253. doi:10.1038/355250a0

Nowak, M. A., and Sigmund, K. (1998). Evolution of indirect reciprocity by image scoring. *Nature* 393, 573–577. doi:10.1038/31225

Panchanathan, K., and Boyd, R. (2004). Indirect reciprocity can stabilize cooperation without the second-order free rider problem. *Nature* 432, 499–502. doi:10.1038/nature02978

Powell, J., Lewis, P. A., Roberts, N., García-Fiñana, M., and Dunbar, R. I. M. (2012). Orbital prefrontal cortex volume predicts social network size: an imaging study of individual differences in humans. *Proc. R. Soc. Lond. B* 279, 2157–2162. doi:10.1098/rspb.2011.2574

R Core Team. (2017). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Available at: http://www.R-project.org/

Rapoport, A., and Chammah, A. N. (1965). *Prisoner's Dilemma: A Study in Conflict and Cooperation*. Ann Arbor: University of Michigan Press.

Rieskamp, J., and Todd, P. (2006). The evolution of cooperative strategies for asymmetric social interactions. *Theory Decis.* 60, 69–111. doi:10.1007/s11238-005-6014-6

Roberts, G. (2008). Evolution of direct and indirect reciprocity. *Proc. R. Soc. Lond. B* 275, 173–179. doi:10.1098/rspb.2007.1134

Robinson-Riegler, G. L., and Winton, W. M. (1996). The role of conscious recollection in recognition of affective material: evidence for positive-negative asymmetry. *J. Gen. Psychol.* 123, 93–104. doi:10.1080/00221309.1996.9921263

Rouder, J. N., and Morey, R. D. (2012). Default Bayes factors for model selection in regression. *Multivariate Behav. Res.* 47, 877–903. doi:10.1080/00273171.2012.734737

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., and Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychon. Bull. Rev.* 16, 225–237. doi:10.3758/PBR.16.2.225

Sallet, J., Mars, R. B., Noonan, M. P., Andersson, J. L., O'Reilly, J. X., Jbabdi, S., et al. (2011). Social network size affects neural circuits in macaques. *Science* 334, 697–700. doi:10.1126/science.1210027

Samuelsson, H., Jarnvik, K., Henningsson, H., Andersson, J., and Carlbring, P. (2012). The Umeå university database of facial expressions: a validation study. *J. Med. Internet Res.* 14, e136. doi:10.2196/jmir.2196

Schacter, D. L. (1999). The seven sins of memory: insights from psychology and cognitive neuroscience. *Am. Psychol.* 54, 182–203. doi:10.1037/0003-066X.54.3.182

Schneider, W., Eschmann, A., and Zuccolotto, A. (2002). *E-Prime Reference Guide*. Pittsburgh, PA: Psychology Software Tools, Inc.

Simon, H. A. (1955). A behavioral model of rational choice. *Q. J. Econ.* 69, 99–118. doi:10.2307/1884852

Stephens, D. W., Nishimura, K., and Toyer, K. B. (1995). Error and discounting in the iterated prisoner's dilemma. *J. Theor. Biol.* 176, 457–469. doi:10.1006/jtbi.1995.0213

Stevens, J. R., and Hauser, M. D. (2004). Why be nice? Psychological constraints on the evolution of cooperation. *Trends Cogn. Sci.* 8, 60–65. doi:10.1016/j.tics.2003.12.003

Stevens, J. R., Volstorf, J., Schooler, L. J., and Rieskamp, J. (2011). Forgetting constrains the emergence of cooperative decision strategies. *Front. Psychol.* 1:235. doi:10.3389/fpsyg.2010.00235

Stiller, J., and Dunbar, R. I. M. (2007). Perspective-taking and memory capacity predict social network size. *Soc. Networks* 29, 93–104. doi:10.1016/j.socnet.2006.04.001

Traulsen, A., and Nowak, M. A. (2006). Evolution of cooperation by multilevel selection. *Proc. Natl. Acad. Sci. U.S.A.* 103, 10952–10955. doi:10.1073/pnas.0602530103

Trivers, R. L. (1971). The evolution of reciprocal altruism. *Q. Rev. Biol.* 46, 35–57. doi:10.1086/406755

Venzon, D. J., and Moolgavkar, S. H. (1988). A method for computing profile-likelihood-based confidence intervals. *J. R. Stat. Soc. C Appl. Stat.* 37, 87–94. doi:10.2307/2347496

Volstorf, J., Rieskamp, J., and Stevens, J. R. (2011). The good, the bad, and the rare: memory for partners in social interactions. *PLoS ONE* 6:e18945. doi:10.1371/journal.pone.0018945

Von Der Heide, R., Vyas, G., and Olson, I. R. (2014). The social network-network: size is predicted by brain structure and function in the amygdala and paralimbic regions. *Soc. Cogn. Affect. Neurosci.* 9, 1962–1972. doi:10.1093/scan/nsu009

Wagenmakers, E., Marsman, M., Jamil, T., Ly, A., Verhagen, A., Love, J., et al. (in press). Bayesian inference for psychology: Part I: theoretical advantages and practical ramifications. *Psychon. Bull. Rev.* doi:10.3758/s13423-017-1343-3

Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems of p values. *Psychon. Bull. Rev.* 14, 779–804. doi:10.3758/BF03194105

Wickham, H. (2011). The split-apply-combine strategy for data analysis. *J. Stat. Softw.* 40, 1–29. doi:10.18637/jss.v040.i01

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer.

Wickham, H. (2017). *tidyr: Easily Tidy Data with "Spread()" and "Gather()" Functions*. Available at: https://CRAN.R-project.org/package=tidyr

Wickham, H., and Francois, R. (2016). *dplyr: A Grammar of Data Manipulation*. Available at: https://CRAN.R-project.org/package=dplyr

Wilke, C. O. (2016). *cowplot: Streamlined Plot Theme and Plot Annotations for "ggplot2"*. Available at: https://CRAN.R-project.org/package=cowplot

Wu, J., and Axelrod, R. (1995). How to cope with noise in the iterated prisoner's dilemma. *J. Confl. Resolut.* 39, 183–183. doi:10.1177/0022002795039001008

Xie, Y. (2015). *Dynamic Documents with R and Knitr*, 2nd Edn. Boca Raton, FL: CRC Press.