Reciprocity explains food sharing in humans and other primates independent of kin selection and tolerated scrounging: a phylogenetic meta-analysis

Adrian V. Jaeggi¹,² and Michael Gurven¹

1 Integrative Anthropological Sciences, Department of Anthropology, and ²Sage Center for the Study of the Mind, University of California Santa Barbara, Santa Barbara, CA 93106-3210, USA


Received: 21 June 2013
Accepted: 17 July 2013

Subject Areas:
behaviour

Keywords:
altruism, cooperation, food sharing, reciprocity, kin selection, meta-analysis

Author for correspondence:
Adrian V. Jaeggi
e-mail: ajaeggi@anth.ucsb.edu

1. Introduction

Helping, i.e. behaviour increasing the fitness of others, can evolve when directed towards kin or reciprocating partners. These predictions have been tested in the context of food sharing both in human foragers and non-human primates. Here, we performed quantitative meta-analyses on 32 independent study populations to (i) test for overall effects of reciprocity on food sharing while controlling for alternative explanations, methodological biases, publication bias and phylogeny and (ii) compare the relative effects of reciprocity, kinship and tolerated scrounging, i.e. sharing owing to costs imposed by others. We found a significant overall weighted effect size for reciprocity of $r = 0.20–0.48$ for the most and least conservative measure, respectively. Effect sizes did not differ between humans and other primates, although there were species differences in in-kind reciprocity and trade. The relative effect of reciprocity in sharing was similar to those of kinship and tolerated scrounging. These results indicate a significant independent contribution of reciprocity to human and primate helping behaviour. Furthermore, similar effect sizes in humans and primates speak against cognitive constraints on reciprocity. This study is the first to use meta-analyses to quantify these effects on human helping and to directly compare humans and other primates.
[11] finding independent effects of reciprocity even when controlling for such factors do suggest the possibility of score-keeping. That animals may act on the basis of interactions experienced over long timescales is known in other contexts such as dominance and aggression, where histories of past interactions compressed into statistical representations are more predictive of behaviour than short-term fluctuations [14]. In the context of helping then, statistical representations of relative partner value, rather than detailed accounts of benefits given and received could function as score-keeping and be recruited in decisions on help allocation [15,16]. Allocating help to the most valuable partners is more akin to market models than traditional tit-for-tat models of reciprocity [2–5] but should similarly result in fitness benefits to both partners and thus be under positive selection.

In summary, we emphasize that help received over long timescales is more predictive of help given than recent interactions, and acts as key evidence of reciprocity in a number of species [4,5,9,10]. Evidence of score-keeping requires showing such effects independent of other factors [11,13]. Score-keeping could come about through computational processes accessible to many species [14] suggesting that species differences in reciprocity should be due to potential fitness benefits of reciprocal helping rather than cognitive constraints. Here, we investigate these questions in the context of food sharing among primates and human foragers, where evidence for reciprocity is also controversial [17,18]. In addition to the general critiques of reciprocity mentioned earlier, it is particularly debated whether food can be controlled enough to be preferentially shared with kin or reciprocating partners [19–21], or whether sharing results from a lack of such control, i.e. tolerated scrounging [22,23].

Food sharing (henceforth sharing) can be defined as the un-resisted transfer of food from one individual to another [17]. Although food production can be mutualistic [24], any benefits to sharing are likely to be delayed and therefore not explained by mutualism [25]. Sharing often occurs in situations conducive to reciprocity as food possessors can provide large benefits to non-possessors at small marginal cost due to diminishing returns of consumption [23] and roles are frequently reversed, at least in human foragers [18]. An important function of reciprocal sharing among human foragers is to smooth consumption by reducing the risk of shortfalls and by increasing production efficiency; and reciprocity may thus be an essential component of the human foraging niche [18]. Furthermore, dominants or high producers with greater access to food may trade it for services provided by others such as mating or coalitionary support [26–29]. As different commodities are exchanged among the same partners, the combination of in-kind reciprocity and trade should result in long-term correlations between giving food and receiving food and/or any other commodities, as is commonly found among primates [7,8,26,29].

Alternatives to reciprocal sharing include nepotistic sharing with kin or tolerated scrounging [22,23] (or harassment [30]). With tolerated scrounging, sharing occurs because of possessors’ inability to monopolize food due to costs imposed by non-possessors, and food given is therefore predicted by factors operationalizing these costs such as spatial proximity, relative need, dominance rank or solicitation pressure [20,31–33] rather than food or other benefits received. Nepotistic sharing and tolerated scrounging should lead to imbalances in exchange in favour of close kin, neighbours, low producers or dominants. As with the classic example of sharing among vampire bats [13], solid evidence for reciprocity in humans and primates therefore needs to show balanced exchange after controlling for sharing due to nepotism and tolerated scrounging. Evidence for reciprocity independent of these other factors is consistent with some form of score-keeping [11].

Following from this introduction, we address a number of questions:

- Is reciprocity a commonly observed feature of human and primate food sharing, i.e. are there significant effects of receiving food and/or other commodities on giving food?
- Are there differences in the strength of reciprocity between taxonomic subgroups, in particular between humans, apes and monkeys?
- Does the effect of reciprocity remain even after controlling for kinship, tolerated scrounging, methodological differences and phylogenetic non-independence?
- What are the relative effects of reciprocity, kinship and tolerated scrounging?

Following previous work on grooming reciprocity in primates [9–11], we used meta-analyses to address these questions. Modern meta-analyses statistically combine results from all available quantitative studies to get overall weighted effect sizes and are therefore more objective than simple vote-counting or narrative reviews [34]. Furthermore, they allow controlling for various confounding variables such as methodological differences and publication bias.

2. Material and methods

(a) Compiling the dataset

We gathered all quantitative studies that tested for a statistical relationship between food given and food and/or other commodities received in human foragers and primates, building on previous comprehensive review work [17,18,26,35] as well as searches of subsequent literature. This search identified 25 studies, all of which focused on sharing among adult individuals. Different species or different groups in the same study were treated as independent study populations and therefore separate data points [10]. Results from the same group observed at different time periods however are not independent, and we therefore restricted our analysis to the most comprehensive study, i.e. the largest number of dyads sampled in order to avoid pseudoreplication, or averaged results across time periods if reported in the same study [36]. Similarly, results reported on partially overlapping subsets of subjects, e.g. male and female recipients [20,37], or obtained from different experiments on the same subjects [9] were averaged. Only study populations with more than three dyads could be considered. This resulted in 23 studies on 32 independent study populations, as listed in electronic supplementary material, table S1.

(b) Obtaining and combining effect sizes

In order to statistically combine studies, we used Pearson’s correlation coefficient $r$ as a measure of effect size [34]. Effect sizes were transformed into Fisher’s $z$ and combined into an overall weighted $r$ using random-effects meta-analyses with inverse-variance weighting [34]. Thus, studies sampling a larger number of dyads were weighted more heavily [10]. In addition, study period length and the average number of interactions per
dyad are expected to influence overall effect size [26] and were available for a subset of studies. We obtained $r$ directly from reported primary analyses or after transforming $p$-values into $z$-scores following standard meta-analytical procedures [34], either from the reported primary analyses or after running generalized linear mixed-effects models (GLMMs) on raw data reported or obtained from first authors (see electronic supplementary material, table S1). If no exact $p$-value was reported in the primary analysis (e.g. n.s. or $p > 0.05$), we emailed the respective first author to enquire about exact values. For significant $p$-values reported as below a certain threshold (e.g. $p < 0.05$) we conservatively used the upper boundary ($p = 0.05$). We only included continuous measures of sharing but not binary measures such as presence or absence of sharing relationships. Furthermore, we only included direct reciprocity but not generalized reciprocity, i.e. correlations between giving to all others and receiving from all others.

Gaussian or Poisson GLMMs were fitted on raw data depending on the distribution of the dependent variable, and included donor and recipient ID as random factors as well as additional controls, if available, as fixed factors. Each dyad was included only once (i.e. A-B but not B-A, with random ordering) and dyads that never shared food were excluded, as correlations between zeroes due to incomplete sampling would inflate $r$ [36]. Even though zeroes may also represent a form of reciprocal interaction as individuals could mutually choose not to interact, their exclusion ensures conservative results. Zeroes were not excluded in all published studies, however, this practice largely overlaps with the use of analyses controlling for repeated measures such that we did not code exclusion of zeroes separately.

Studies differ in the measures of sharing analysed, with no one measure being available for all studies (see electronic supplementary material, table S1). Whenever multiple sharing measures were available, such as frequency, quantities, marginal value or proportion of production shared, we conservatively used the lowest resulting $r$. However, we also report the overall effect obtained from the higher possible $r$, i.e. the least conservative measure in the subset of studies for which multiple measures were available [38].

Studies of human foragers typically consider only food-for-food exchange, whereas studies on primates commonly analyse the effect of receiving multiple different commodities such as grooming, coalitionary support or mating on sharing [17]. We therefore mainly focus on the combined effect of in-kind reciprocity and trade [26]; however, separate effects are reported and discussed, with detailed results in electronic supplementary material. To combine in-kind reciprocity and trade, we averaged effect sizes for multiple commodities obtained from the same study unit but separate analyses (e.g. food-food, food–grooming, food–coalitions) using the method provided by Corey et al. [39], and added effect sizes obtained from the same analysis (e.g. food–food + grooming + coalitions). Relationship quality, a composite factor including grooming, coalitionary support, aggression and/or time spent in proximity was treated as a measure of trade [8,31,32]. Interaction terms could not be taken into account. Partial effect sizes for kinship and tolerated scrounging were obtained in the same way [11].

(c) Testing moderators of effect size

In order to get the most robust measure of overall effect size, meta-analyses test for the effect of moderators, i.e. factors that predictably increase or decrease effect size across studies [34]. Here, we coded six potential moderators (table 1) and tested for their effect on reciprocity using meta-regressions [43]. First, effect sizes are likely to vary across studies depending on the statistical procedures employed. In particular, studies that did not control for repeated measures of the same individuals are subject to pseudo-replication and therefore likely to overestimate effect size [10]. Thus, we emailed all authors who reported analyses without controlling for repeated measures and solicited raw data in order to re-run the analyses as described earlier. If no raw data were obtained, we coded the study as not controlling for repeated measures (table 1 and electronic supplementary material, table S1).

<table>
<thead>
<tr>
<th>variable</th>
<th>definition</th>
<th>effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>main effect</td>
<td>statistical relationship between giving food and receiving food and/or other commodities (grooming, mating, support)</td>
<td>$0.20$ (95% CI 0.14–0.26)</td>
</tr>
<tr>
<td>moderators</td>
<td>binary, did the study control for repeated measures of the same individuals in multiple dyads?</td>
<td>$-0.11$ ($\pm 0.08$), $z = -1.45$, $p = 0.15$</td>
</tr>
<tr>
<td>repeated measures</td>
<td>binary, did the study control for relatedness (as defined by original studies)?</td>
<td>$-0.03$ ($\pm 0.07$), $z = -0.52$, $p = 0.60$</td>
</tr>
<tr>
<td>kinship</td>
<td>binary, did the study control for kinship (as operationalized by original studies)?</td>
<td>$-0.08$ ($\pm 0.06$), $z = -1.32$, $p = 0.19$</td>
</tr>
<tr>
<td>tolerated scrounging</td>
<td>binary, did the study control for tolerated scrounging (as operationalized by original studies)?</td>
<td>$-0.07$ ($\pm 0.08$), $z = -0.84$, $p = 0.40$</td>
</tr>
<tr>
<td>captivity</td>
<td>binary, were primates observed in captivity?</td>
<td>$-0.01$ ($\pm 0.07$), $z = -0.14$, $p = 0.89$</td>
</tr>
<tr>
<td>observation bias</td>
<td>binary, was the study subject to unequal sampling?</td>
<td>$-0.10$ ($\pm 0.06$), $z = -1.64$, $p = 0.10$</td>
</tr>
</tbody>
</table>

*aSee text for how multiple measures were combined.

bExample includes spatial proximity, relative need, dominance rank and solicitation pressure [20,31–33].

cObservation bias applies to groups with fluid association [21] (unless association is controlled for [29]), or with unequal sampling methods [40,41]. It does not apply to studies of captive primates or more sedentary human groups with equally distributed sampling.

dProduction bias applies when some individuals consistently had more access to food but can be controlled for by correlating the proportions of production shared [38] or the proportions of interactions resulting in sharing [42] rather than amounts or number of times shared.

Table 1. Definitions for the main effect of reciprocity and potential moderator variables, as well as an overview of the overall weighted effect size and the influence of individual moderators as obtained through meta-regressions.
Furthermore, we coded whether studies controlled for kinship or tolerated scrounging as operationalized by the original authors. We also compared studies on captive and wild primates. Positive correlations between giving and receiving could also arise from biases in observation or production even when food is shared indiscriminately, as argued by Hawkes et al. [19] with the help of simulations. They showed that simply observing A and B more often together than other dyads can result in positive correlations. This observation bias (table 1) is problematic for groups with fluid association [21], unless association is controlled for [29]. Hawkes et al. [19] also showed that positive correlations can come about by both A and B being high producers and therefore more often getting shares from each other. This production bias (table 1) can be controlled for by taking the proportion of A’s production that is shared with B rather than the number of transfers or amount transferred [38,40]. Similarly, in primates, food possession and therefore the opportunity to share may be biased towards dominants but can be controlled for by measuring the proportion of B’s food requests to A that resulted in transfer [42]. Whenever calculating r from raw data, we therefore controlled for production bias by calculating r from proportions as well as total amounts shared and, again, conservatively used the lower resulting r in the main analysis. In addition, we coded whether observation or production bias applied to a given study given its methodology and tested for differences between studies with or without these biases (table 1).

(d) Controlling for phylogeny and publication bias

Meta-analyses across species and populations are potentially affected by phylogenetic non-independence [44]. In order to account for this, we tested whether the strength of the phylogenetic signal λ in the effect sizes was significantly greater than zero. After constructing a phylogenetic tree of our study populations based on genetic information (see electronic supplementary material, methods and figure S1), we estimated λ with a maximum-likelihood model and compared the fit of this model to one where λ is set to zero using likelihood ratio tests [45]. However, there was no evidence of phylogenetic signal in the data as the estimated λ was not significantly different from zero (all λ < 0.0001, χ² < -0.0001, d.f. = 1, p = 1) and we therefore ignored phylogeny in subsequent analyses. This is not surprising given how flexibly sharing is adjusted to local ecology [46].

Lastly, we used two methods to estimate and control for publication bias, i.e. the tendency for positive results to get published more than negative or null results [34]. First, we calculated weighted fail-safe numbers, a qualitative measure of robustness, following Rosenberg [47]: N₁ is the sample size of a single study of effect size zero, and Nₖ the number of studies of average sample size and effect size zero that would be necessary to eliminate an overall significant effect. Rosenberg’s fail-safe numbers were obtained using a downloadable calculator (http://www.rosenberglab.net/software.php?failsafe). Second, we used Egger et al.’s [48] test for asymmetries in a funnel plot to detect publication bias (cf. electronic supplementary material, figure S2), and the trim-and-fill method [49] to identify missing studies and adjust the overall effect size. All analyses and graphics were done in R v. 2.15.2. [50] using the packages meta [51], lmtest [52], and ggplot2 [53]. For each meta-analysis, we report the weighted effect size r with 95% confidence intervals (CIs), the number of study populations k, the fail-safe numbers N₁ and Nₖ, as well as tests for differences between taxonomic subgroups, i.e. monkeys, apes and humans.

3. Results

(a) Independent effect of reciprocity across species

The effect of receiving food and/or other commodities on giving food was tested in 32 independent study populations (eight monkeys, eight apes and 16 humans; figure 1 and electronic supplementary material, table S1). The overall weighted effect size was r = 0.20 (95% CI = 0.14–0.26, k = 32) which was significantly greater than zero (z = 6.22, p < 0.001), and the large fail-safe numbers (N₁ = 1301, Nₖ = 1406) indicate substantial robustness of this effect. There were no significant differences in effect size between monkeys (r = 0.12, 95% CI = −0.04–0.27, k = 8), apes (r = 0.20, 95% CI = 0.09–0.30, k = 8) and humans (r = 0.22, 95% CI = 0.13–0.31, k = 16; test for subgroup difference: Q = 1.23, d.f. = 2, p = 0.54). This is shown in figure 1. Egger et al.’s test for asymmetry in the funnel plot indicated no publication bias (t = 1.61, d.f. = 30, p = 0.12; see electronic supplementary material, figure S2) and the trim-and-fill method detected no missing studies, leaving r unchanged. None of the individual moderators significantly changed overall effect size (table 1), indicating that the effect of reciprocity was robust across diverse methodologies.

The most and least conservative measures of effect size were compared in the 11 study populations (all humans) for which multiple sharing measures were available, such as frequency, quantities, marginal value or proportion of production shared (see electronic supplementary material, table S1). In this subsample, the overall weighted effect of the most conservative r’s, i.e. the ones used in the above analysis was r = 0.26 (95% CI = 0.10–0.40, k = 11, z = 3.18, p = 0.002), only slightly higher than the complete sample for humans. However, the overall weighted effect of the least conservative measures was almost twice as high, r = 0.48 (95% CI = 0.31–0.61, k = 11, z = 5.17, p < 0.001; see electronic supplementary material, figure S4).

Breaking up reciprocity into exchanges in kind and trade of food for other commodities yielded overall weighted effect sizes of r = 0.19 each, similar to the ones reported above (see electronic supplementary material, results for details). Even though tests for subgroup differences were not significant (all p > 0.2), only human foragers showed significant effects of in-kind reciprocity (r = 0.24, 95% CI = 0.14–0.33, k = 16), whereas the CIs for monkeys (r = 0.14, 95% CI = −0.02–0.29, k = 8) and apes (r = 0.11, 95% CI = −0.02–0.23, k = 6) overlapped with zero. Several moderators significantly decreased effect sizes and after reducing the sample to studies controlling for them, these species differences became even more pronounced (see electronic supplementary material, results). Conversely, only apes showed a significant effect of trade (r = 0.17, 95% CI = 0.05–0.30, k = 8), although trade was only examined in a small number of study populations in other subgroups (monkeys: r = 0.07, 95% CI = −0.20–0.33, k = 4; humans: r = 0.32, 95% CI = −0.06–0.59, k = 2).

Finally, there was some indication that study period length influenced correlations between giving and receiving as the number of observation hours but not study duration in weeks significantly increased overall effect size of reciprocity (see electronic supplementary material, results for details). The average number of transfers per dyad influenced effect size in a way similar to the number of dyads, with effect sizes being noisier for small sample sizes (see electronic supplementary material, figures S2 and S3). Using number of transfers per dyad rather than number of dyads to weight effect sizes yielded an identical overall weighted effect size of r = 0.20 (see the electronic supplementary material, results for details).
Relative effects of reciprocity, kinship and tolerated scrounging

Relative effect sizes for reciprocity, kinship and tolerated scrounging were calculated for studies that included all three factors in the same model (eight studies with 10 independent study populations: three apes and seven humans; electronic supplementary material, table S1). As shown in figure 2, all effects were significantly greater than zero but not significantly different from each other (reciprocity: \( r = 0.20 \), 95% CI = 0.12–0.28, \( z = 4.57, p < 0.001 \); kinship: \( r = 0.14 \), 95% CI = 0.06–0.22, \( z = 3.50, p < 0.001 \); tolerated scrounging: \( r = 0.22 \), 95% CI = 0.08–0.35, \( z = 2.98, p = 0.002 \); all \( k = 10 \)). Differences between taxonomic subgroups were not explored due to the small number of studies.

4. Discussion

We performed meta-analyses of 23 studies on 32 independent study populations testing for the effect of receiving food and/or other commodities on sharing food. With a total sample size of almost 10,000 dyads, we obtained a significant overall effect of reciprocity that did not differ between humans, apes and monkeys (figure 1) and was robust across methodological differences (table 1). The conservative lower estimate of the overall weighted effect size was \( r = 0.20 \) but it ranged up to \( r = 0.48 \) for the highest possible effect obtained from human studies with multiple sharing measures (see electronic supplementary material, figure S4). The relative effect of reciprocity in explaining sharing was not significantly different from those of kinship and tolerated scrounging in the subset of studies including all three factors in the same analysis (figure 2). This study is the first to use modern meta-analytical techniques to quantify independent effects of reciprocity and alternative hypotheses in food sharing, a form of helping inseparably linked to the evolved human life-history and pro-social psychology [16–18].

The conservative effect size of \( r = 0.20 \) for reciprocity (figure 1) and the independent effects of reciprocity, kinship and tolerated scrounging (figure 2) were similar to the ones...
Figure 2. The relative effects of reciprocity (RA), kin selection (KS) and tolerated scrounging (TS) on food sharing in a subset of studies including all three factors in the same analysis (three on apes and seven on humans). Bars represent overall weighted effect sizes with 95% CIs. Species differences were not investigated due to the small number of studies.

reported for the partial effects of grooming on coalitional support ($r = 0.16$ [10]) and kinship on grooming in primates ($r = 0.18$ [11]) but considerably lower than in-kind reciprocity in grooming ($r = 0.47$ [9]). The combined effects of reciprocity, kin selection and tolerated scrounging in studies including all three factors was $r = 0.49$. Overall, our results thus corroborate previous work showing significant effects of reciprocity on primate cooperation independent of other factors [9–11], but also highlight the need to consider different explanations of sharing simultaneously [26,31,32,40,58–60].

The highest effect of reciprocity obtained for human studies with multiple sharing measures was $r = 0.48$, almost twice as high as the most conservative estimate for this subsample, and independently explained 23% of the variance in food given. Even though no single sharing measure consistently yielded the highest estimates of reciprocity (see electronic supplementary material, table S1), different measures may add ecological and psychological realism [38]. For instance, proportions shared may be a better indicator of the relative value attributed to social partners than total quantities shared, emphasizing the long-term aspects of social relationships over absolute short-term gains [38]. Furthermore, applying discounting functions to quantities shared better approximates the diminishing marginal returns to consumption, especially for large items [23]. Which particular component of sharing is maximized may depend on local production economy, as argued for differences between Ache and Hiwi foragers [38,67].

Finally, the proximate mechanisms underlying reciprocity warrant more investigation, especially given that sharing among human foragers and primates alike is often initiated by recipients [16–18]. Thus, preferential sharing may already start with selective association, restricting the range of potential recipients and reducing transaction costs of monitoring and exchange [46]. Long-term allies may be more likely to approach food possessors, sometimes excluding other potential recipients through the formation of begging clusters [28]. Finally, possessors can direct sharing through differential responses to approaches, tolerating some but not others, as well as active forms of sharing [31,32]. Evidence presented here suggests that the responses of possessors may be affected by some form of score-keeping such as long-term representations of partner value [4,15,16] independent of positive assortment operationalized as kinship, spatial proximity or rank distance (cf. table 1). Ultimately, no matter what proximate mechanisms brought about the significant correlations between giving and receiving observed here, they yield benefits to both individuals and should thus be under positive selection.

Acknowledgements. We thank Ray Hames, Paul Hooper, Lisa Rapaport and John Ziker for kindly providing unpublished material, Gabriele Schino for generous methodological advice, James Coxworth, Pavel Duda, Drew Gerkey and Kristen Hawkes for discussions, and Christoph Boesch, Bobbi Low, Shane Macfarlan, Chris von Rueden, Ben Trumble and one anonymous referee for helpful comments on earlier versions of this manuscript.

Funding statement. This work was supported by a postdoc grant from the Swiss National Science Foundation (PBZHP3-133443) and a Sage Junior Research Fellowship to A.V.J.


52. Bates D, Maechler M, Bolker BM. 2011 lme4: Linear mixed-effects models using S4 classes. See http://lme4.r-forge.r-project.org/.


60. Hooper PL. 2011 The structure of energy production and redistribution among Tsimane’ forager–horticulturalists. PhD, Anthropology, University of New Mexico: Albuquerque New Mexico, xv, 131.


