Behavior Research Methods
2008, 40 (2), 626-634
doi: 10.3758/BRM.40.2.626

Testing reciprocity in social interactions: A comparison between the directional consistency and skew-symmetry statistics

David Leiva, Antonio Solanas, and Lluís Salafranca
University of Barcelona, Barcelona, Spain

In the present article, we focus on two indices that quantify directionality and skew-symmetrical patterns in social interactions as measures of social reciprocity: the directional consistency (DC) and skew-symmetry indices. Although both indices enable researchers to describe social groups, most studies require statistical inferential tests. The main aims of the present study are first, to propose an overall statistical technique for testing null hypotheses regarding social reciprocity in behavioral studies, using the DC and skew-symmetry statistics (Φ) at group level; and second, to compare both statistics in order to allow researchers to choose the optimal measure depending on the conditions. In order to allow researchers to make statistical decisions, statistical significance for both statistics has been estimated by means of a Monte Carlo simulation. Furthermore, this study will enable researchers to choose the optimal observational conditions for carrying out their research, since the power of the statistical tests has been estimated.

Statistical analysis of social interactions should consider some particularities that social researchers have increasingly taken into account in their studies. Specifically, most researchers are interested in estimating actor, partner, and relationship effects and even in describing groups as a whole. These characteristics often make some classical statistical tests unsuitable for analyzing social interaction data. For analyzing social phenomena, therefore, dominance and social reciprocity statistical tests have been put to use, along with well-known statistical tests (Appleby, 1983; de Vries, 1995; Hemelrijk, 1990a, 1990b; Kenny & La Voie, 1984; Landau, 1951; Rapoport, 1949; Warner, Kenny, & Stoto, 1979).

A correlational approach has been proposed for testing reciprocity and interchange at group level (Hemelrijk, 1990a, 1990b). In this approach, statistical significance is obtained by means of a kind of permutation test. This procedure quantifies social reciprocity as a whole, because the association coefficient is an overall measure of global reciprocity or interchange in the social group. According to Hemelrijk (1990b), there are three types of reciprocity: relative, absolute, and qualitative. The present study is concerned with absolute reciprocity. A group presents absolute reciprocity when there is exact matching between agents' amount (or duration) of behavior. Relative reciprocity requires data to be ranked within each individual, whereas qualitative reciprocity implies that the comparison is done on a binary scale. Hemelrijk (1990b) also described two models that can be applied in ethological studies of reciprocity: the actor–reactor and the actor–receiver models. According to the first model, individuals give most often to those who more frequently give them something in return. The actor–receiver model involves the comparison between what each individual gives and receives in return. We here focus on absolute reciprocity, since we are interested in testing the symmetry of a sociomatrix, defined as the number of behaviors given and received among individuals. Moreover, we use an actor–receiver model because we assume that actors in dyads compare only what is given and received from their partners, without taking into account what is given and received from the others. The actor–reactor model requires more complex cognitive abilities (i.e., each individual must be able to trace the acts of the other individuals) and does not allow construction of a complete reciprocation sociomatrix for odd group sizes; as a consequence, it is not possible to test for social reciprocity (Hemelrijk, 1990b). For these reasons, the present research is concerned with the more parsimonious and unrestricted actor–receiver model.

The social relations model (SRM; Kenny & La Voie, 1984) is useful to analyze data from round-robin designs, since it uses dyadic relations for the study of social phenomena. Although the SRM has been commonly used in interpersonal perception studies (Kenny, 1994), it can be also applied to analyze interaction behaviors in groups. In addition to the mean level, the SRM decomposes each dyadic observation of sociomatrices into the actor effect, partner effect, and relationship effect. The SRM uses a random effects two-way ANOVA, which allows estimat-

D. Leiva, dleivaur@ub.edu

ing the actor variance, partner variance, and relationship variance, to make statistical decisions. This model also enables social researchers to assess different kinds of social relations: dyadic and generalized reciprocity. Whereas dyadic reciprocity refers to interdependence in dyads, generalized reciprocity measures dependencies at the individual level. The SRM quantifies dyadic and generalized reciprocity in groups by means of the product–moment correlation coefficient; therefore, this procedure is also founded on a correlational approach to measure social reciprocity. It should be noted that the SRM does not enable social researchers to measure social reciprocity at the global level and does not take into account the absolute dyadic differences to quantify social reciprocity.

The directional consistency index (DC; van Hooff & Wensing, 1987) has widely been used by biologists in order to quantify the directionality of behavior in social interactions (Côté, 2000; Koenig, Larney, Lu, & Borries, 2004; Pelletier & Festa-Bianchet, 2006; Stevens, Vervaecke, de Vries, & van Elsacker, 2005; Vervaecke, de Vries, & van Elsacker, 1999; Vogel, 2005). The DC is obtained by dividing the number of the total interactions in the most frequent direction (H) minus the number of interactions in the less frequent direction (L) by the total of interactions performed by all individuals in the group. It should be noted that this is the same as the sum of absolute dyadic differences divided by the total number of interactions:

$$\mathrm{DC} = \frac{(H-L)}{(H+L)} = \frac{\sum_{i=1}^{n}\sum_{j=i+1}^{n}\left|x_{ij}-x_{ji}\right|}{N};$$

$$N = \sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} x_{ij};\ 0 \le \mathrm{DC} \le 1,$$

where x_{ij} is the number of interactions that individual i addresses to individual j, x_{ji} is the number of interactions that individual i receives from individual j, N is the total number of interactions in the group, and n is the number of individuals. The index ranges from 0 to 1. When the DC index takes a value close to 0, social reciprocity is near its maximum. On the other hand, when the index is close to 1, most dyadic interactions are unidirectional and social reciprocity is near its minimum value.

Another quantification of social reciprocity has recently been proposed (Solanas, Salafranca, Riba, Sierra, & Leiva, 2006). This method is based on partitioning a sociomatrix (**X**) into its symmetrical and skew-symmetrical parts (Constantine & Gower, 1978):

$$\mathbf{X} = \frac{\mathbf{X}+\mathbf{X}'}{2} + \frac{\mathbf{X}-\mathbf{X}'}{2} = \mathbf{S}+\mathbf{K},$$

where $\mathbf{X}'$ denotes the transpose of the sociomatrix **X**, the elements in matrix **S** are the average of the amount of behavior addressed and received by each individual, and the elements in matrix **K** represent the average of differences between the number of behaviors emitted and received by each individual in the group. This method enables researchers to describe groups at individual, dyadic, and group levels, assuming that global phenomena depend on dyadic interactions. It takes into account the absolute differences among agents' dyadic behaviors in order to compute a measure of reciprocity. The method also allows researchers to quantify generalized and dyadic reciprocity by means of discrepancy measures. Furthermore, a proximity matrix can be obtained and multidimensional scaling can be applied to determine underlying dimensions in groups and to represent individuals in a Euclidean space.

Solanas et al. (2006) proposed quantifying overall reciprocity in groups by means of the skew-symmetry index or, if preferred, the symmetry index. The skew-symmetry index is computed as follows:

$$\Phi = \frac{tr(\mathbf{K}'\mathbf{K})}{tr(\mathbf{X}'\mathbf{X})} = \frac{\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} k_{ij}^2}{\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} x_{ij}^2},\ tr(\mathbf{X}'\mathbf{X}) > 0;\ 0 \le \Phi \le .5,$$

where **X** and **K** denote any sociomatrix and its corresponding skew-symmetrical matrix, respectively. The symmetry index, denoted by Ψ, is equal to $1 - \Phi$. Note that the larger the skew symmetry is in groups, the closer the skew-symmetry value is to .5.

Now, we can write the DC index in the following re-expression:

$$\mathrm{DC} = \frac{2\sum_{i=1}^{n}\sum_{j=i+1}^{n}\left|k_{ij}\right|}{\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} x_{ij}}.$$

Looking at the mathematical expressions of the DC and skew-symmetry indices, it should be noted that these indices will be monotonically correlated.

Note that the DC and skew-symmetry indices have some advantages over other techniques that have been proposed for analyzing groups. Most importantly, these methods enable social reciprocity to be analyzed without any loss of information. That is, the DC and skew-symmetry statistics take into account the differences in the ways in which people behave with one another. In other words, there is no lost information, as occurs when the linear index of hierarchy is computed (Landau, 1951; Rapoport, 1949). Moreover, the DC and skew-symmetry statistics can be useful in studies in which researchers are interested in analyzing absolute differences in behavior instead of calculating any association coefficient for measuring social reciprocity, as occurs when other procedures are used (Hemelrijk, 1990a, 1990b; Kenny, 1994).

Researchers can study patterns of reciprocity in groups using the DC or skew-symmetry statistics, as we have shown above. After describing the group by means of these statistical indices, researchers may be interested in making statistical decisions regarding the null hypothesis. In this case, the null hypothesis often corresponds to complete reciprocation among individuals. For this reason, we propose an overall statistical technique for testing symmetry in any group, although the procedure can be used to test other null hypotheses. Thus, researchers will be able to make decisions about whether the group under analysis presents a statistically significant unidirectional or skew-symmetrical pattern.

In the present article, we pursue several aims. First, we present statistical tests for testing null hypotheses regarding reciprocity in social interactions. Thus, social researchers will be able to associate statistical significance to the DC and skew-symmetry values obtained in their studies. Second, we are interested in comparing two statistics by means of a simulation study. We estimate several sampling distributions for the DC and skew-symmetry statistics, and a power analysis is carried out to allow social researchers to make an optimal choice of statistic depending on the observational conditions. The statistical tests and the simulation study are carried out by using a Monte Carlo method.

A Procedure for Testing Social Reciprocity

We propose a Monte Carlo sampling procedure to test statistical hypotheses concerning social reciprocity, since this method has been recommended for use in studies when the exact distributions are unknown (Noreen, 1989; Peres-Neto & Olden, 2001). Given that the sampling distributions of the DC and skew-symmetry statistics are currently unknown, a Monte Carlo test can be used to estimate them. We emphasize that this statistical method enables social researchers to obtain statistical significance for any sociomatrix, independently of the number of individuals and the amount of behavior for each dyad. It will be needed only to specify the parameter values to be tested, the number of individuals in the group, and the number of behaviors for each dyad.

We denote the number of times that the behavior of interest is registered between individuals i and j by N_{ij}. X_{ij} represents the number of times the individual i addresses behavior to j. We assume that the probability of the event "i addresses behavior to j" (p_{ij}) is a constant value for every trial during the observation period. Note that this assumption is needed if repeated interactions among individuals is gathered and aggregated in a unique sociomatrix (Adams, 2005; Boyd & Silk, 1983; Tufto, Solberg, & Ringsby, 1998), since it is made in round-robin designs. In addition, we assume that the outcomes of successive encounters are independent during the observation period (Appleby, 1983; Boyd & Silk, 1983). This assumption, for example, is also made in the SRM (Warner et al., 1979), since interaction behaviors between the individuals of each pair are counted or aggregated, which means that this dependency cannot be estimated from data at hand. Furthermore, the SRM does not include a term in which dependency between successive interactions is taken into account, and, unfortunately, no general strategy is known for controlling these kinds of order effects (Kenny, Kashy, & Cook, 2006, p. 217). As a consequence of this second assumption, the number of times that i addresses behavior to j, X_{ij}, is binomially distributed with parameters N_{ij} and p_{ij}. This probabilistic approach has previously been used to model social interactions (Tufto et al., 1998). Note that if $p_{ij} = p_{ji}$ for all dyads, all relationships are reciprocal. Thus, $E(X_{ij}) = E(X_{ji})$, and the DC and skew-symmetry statistics computed in samples are expected to be close to 0; otherwise, the value of the DC statistic will be near 1 and that of the skew-symmetry statistic close to .5, as a function of the lack of reciprocity among the dyads. Finally, we also assume that there are no dependency effects between dyads, an assumption also made in the SRM (Kenny et al., 2006, p. 216; Warner et al., 1979).

We propose to use Monte Carlo sampling to generate a specified number of simulated sociomatrices. The p_{ij} parameter values should be established according to the particular null hypothesis to be tested. Additionally, the exact N_{ij} value for each dyad needs to be specified, and the number of individuals in the group, n, has to be established. Given the values for p_{ij}, N_{ij}, and n, sampling distributions for the DC and skew-symmetry statistics can be estimated by Monte Carlo sampling. Therefore, the values of the two statistics obtained from the original sociomatrix can be located at their corresponding sampling distributions, thus obtaining statistical significance.

An asymptotic test could have been proposed to test the complete reciprocation hypothesis, the assumptions mentioned above being also required. According to the reproductive or additive property of the χ^2 distribution, the following statistic is χ^2 distributed with $n(n-1)/2$ degrees of freedom:

$$Z^2 = \sum_{i=1}^{n} \sum_{j=i+1}^{n} \left(\frac{\dfrac{x_{ij}}{N_{ij}} - .5}{\sqrt{\dfrac{.5 \cdot .5}{N_{ij}}}} \right)^2 = \sum_{i=1}^{n} \sum_{j=i+1}^{n} \frac{N_{ij} \left(\dfrac{x_{ij}}{N_{ij}} - .5 \right)^2}{.5^2}.$$

The main problem of using this statistical test is that the statistic does not follow a χ^2 distribution if p_{ij} values clearly differ from .5 and the number of observations is not large enough. A more general solution is to take into account the additive property of the binomial distribution. This property states that if $X_1, X_2, \ldots, X_q$ are binomially distributed with parameters $n_1, n_2, \ldots, n_q$ and p, the random variable $Y = X_1 + X_2 + \ldots + X_q$ follows a binomial distribution with parameters $n_1 + n_2 + \ldots + n_q$ and p. The main drawback of the latest procedure is that the parameter p has to be equal for all dyads. The procedure described in this article is intended to allow social researchers to test any null hypothesis—for instance, $p_{12} = .7$, $p_{13} = .5$, and $p_{23} = .6$.

With respect to the application of other computer-intensive methods, bootstrapping would allow estimating sampling distributions (Manly, 2007; Noreen, 1989). The main problem would be how to carry out the resampling procedure. It seems clear that the elements of a sociomatrix should not be randomly selected to draw bootstrap samples, since dyads are the unit of analysis and the number of behaviors for each dyad in each resampling sociomatrix needs to be equal to those values of the original sociomatrix. The statistical problem would appear again if dyads were randomly chosen for drawing resampling sociomatrices. Note that it is not proper to estimate sampling distributions if there is not a concordance between the amount of behavior in the original and resampling sociomatrices.

For the reasons mentioned above, we propose to use Monte Carlo sampling for testing social reciprocity hypotheses. We used the SAS/IML procedure to develop an SAS program, in order to compute both statistics and obtain statistical significance. This program was also developed in R code. A Monte Carlo test is used to estimate sampling distributions, then to estimate Type I error rates. The programs analyze sociomatrices and provide measures of social reciprocity at group level. Moreover, social researchers can choose the number of simulations of the Monte Carlo sampling to test several null hypotheses. These codes can be useful for social psychologists and ethologists, because they include indices that allow them to measure social processes and make statistical decisions about dyadic interactions in groups. We emphasize that researchers can specify any group size and any number of interactions within dyads.

Several null hypotheses can be tested by means of the proposed Monte Carlo procedure. For instance, social researchers could be interested in testing the null hypothesis of complete reciprocation among individuals. This null hypothesis can be expressed as follows:

$$H_0\colon p_{ij} = p_{ij};\ i, j = 1, 2, 3, \ldots, n;\ i \neq j \text{ and } i < j.$$

Note that the null hypothesis of complete reciprocation states that probabilities of occurrence of behavior are the same for all individuals—for instance, an equal proportion of wins during play interactions (Bauer & Smuts, 2007). However, other null hypotheses can be tested, as has been mentioned.

The simulation steps in Monte Carlo sampling are as follows: (1) Group size is defined according to the size of the original matrix; (2) a random number a is generated from a binomial distribution with parameters N_{ij} and p_{ij}; (3) the random number is assigned to the element on the upper triangular matrix (x_{ij}), and the value on the lower triangular matrix is obtained by the formula $x_{ji} = N_{ij} - x_{ij}$; (4) if the element belongs to the principal diagonal, a 0 value is assigned; (5) Steps 2 and 3 are repeated for each element in the matrix; (6) once the simulated sociomatrix has been generated, the programs compute both the DC and skew-symmetry statistics associated to this simulated sociomatrix; (7) Steps 2–6 are repeated according to the number of iterations that has been previously specified. Statistical significance is computed as (NOS + 1)/(rep + 1), where rep equals the number of the generated sociomatrices and NOS is the number of significant cases. The number of significant cases for both statistics is obtained as the number of simulated statistics that is greater than or equal to the original statistic. This is a valid statistical test, because it ensures that the original statistic is among the set of simulated statistics, so statistical significance can never be smaller than 1/rep (Noreen, 1989; Onghena & May, 1995). Finally, the programs provide some summary statistics related to the simulated sociomatrices, such as mean, standard deviation, and several percentiles.

The following sociomatrix shows grooming interactions in a group of six captive spider monkeys (*Ateles belzebuth hybridus*):

$$\mathbf{X} = \begin{array}{cc} & \begin{array}{cccccc} 1 & 2 & 3 & 4 & 5 & 6 \end{array} \\ \begin{array}{c} 1\\2\\3\\4\\5\\6 \end{array} & \begin{pmatrix} 0 & 9 & 3 & 6 & 12 & 18 \\ 13 & 0 & 7 & 17 & 45 & 6 \\ 7 & 2 & 0 & 5 & 8 & 0 \\ 11 & 26 & 12 & 0 & 8 & 6 \\ 4 & 3 & 0 & 1 & 0 & 2 \\ 2 & 5 & 0 & 1 & 0 & 0 \end{pmatrix} \end{array}.$$

Table 1 shows a brief description of the group. Data contained in the sociomatrix were collected from January to April of 2006 at Barcelona Zoo as a part of a wider study. In short, behavior frequencies were collected, and, in order to obtain a quantification of social reciprocity, individuals were considered as actors and receivers.

When analyzing the matrix **X**, we obtained the following original indices: DC $\approx$ .5146 and $\Phi \approx$ .2768. Using a Monte Carlo test with 99,999 simulated matrices, we found that both statistics were statistically significant (p = .00001). Thus, a significant unidirectional and skew-symmetrical pattern in grooming interactions exists in the studied group; that is, there is a clear lack of complete reciprocation in that group. Table 2 shows several results of the simulation.

Simulation Study

The amount of behavior per dyad was established as constant for each dyad in the simulation (i.e., $N = N_{ij} = N_{ji}$). This constraint allowed us to control an important factor in the simulation, and therefore to study the effect of increasing the number of individuals in group on Type II error rates for the DC and skew-symmetry statistics. Note that the quantity of possible sociomatrices is infinite and we were unable to study all possibilities. Additionally, we were also interested in making a comparison between the two statistics in order to know which is less biased under the null hypothesis of complete reciprocation, which states that dyadic relations are reciprocal among all individuals (p_{ij} = .5). We focused on studying the null hypothesis of complete reciprocation, since it seems the most significant test for social researchers; although, as mentioned above, the proposed statistical procedure enables them to test other null hypotheses.

Table 1
Spider Monkey Descriptions

Individual	Description
1	Adult female
2	Adult female
3	Adult female
4	Juvenile female
5	Adult male
6	Infant male

Table 2
Some Results of the Monte Carlo Test for Grooming Interaction Data in a Captive Group of Spider Monkeys Under Null Hypothesis $p_{ij} = .5$

	DC	Φ
Original statistical value	.5146	.2768
p value	.00001	.00001
N simulations	99,999	99,999
Mean	.1811215	.0359286
Variance	.0015176	.0002522
MS_e	.0343226	.0061722
Maximum	.3807531	.1525672
Minimum	.0627615	.0035971
25th percentile	.1548117	.0243492
50th percentile	.1799163	.0332322
75th percentile	.2050209	.0442544

Note—Both statistics were significant (DC $= .5146$, $p = .00001$; $\Phi = .2768$, $p = .00001$). Mean square error (MS_e) values correspond to a 0–1 scale.

Method

In total, we studied 300 experimental conditions as a result of varying three factors: group size (n), amount of behavior for each dyad (N), and the probability associated with the event "individual i addresses behavior to individual j" (p_{ij}). Specifically, 12 values were established for group size ($n = 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25$, and 30); 5 values for the total amount of behavior in each dyad ($N = 5, 10, 20, 30$, and 60); and 5 values for reciprocity levels ($p_{ij} = .5, .6, .7, .8$, and .9).

This intensive computer-simulation experiment allowed us to estimate sampling distributions for the DC and skew-symmetry statistics. We established the following statistical significance levels for studying empirical Type II error rates: .05 and .01. Thus, we investigated the power of the statistical test under the null hypothesis of complete reciprocation. In order to estimate statistical power ($1-\beta$), we obtained the DC and Φ cutoff points for specific α, n, N, and $p_{ij} = .5$. We compared these cutoff points with the values of the two statistics obtained in sociomatrices randomly drawn from populations in which $p_{ij} = .6, .7, .8$, and .9, keeping the values of α, n, and N constant. Statistical power was estimated as the proportion of values as large as, or larger than, the DC and Φ cutoff points under the null hypothesis and for specific α. That is to say, Type II error rates (β) were estimated as the proportion of simulated values lower than the cutoff points. Once we had estimated statistical power for the DC and skew-symmetry statistical tests, we were able to compare the two statistical tests in order to choose the most optimal one. In addition, some of the simulated data were used to compare the two statistics regarding their bias.

A FORTRAN 90 program was developed to carry out the simulations, using the Salford FTN90 v2.19.1 compiler for Windows. The NAG Release 3 libraries for Windows program was used to generate sociomatrices under different conditions, specifically the nag_rand_discrete and nag_rand_contin modules. The simulation steps were as follows: (1) n, N, and p_{ij} were specified; (2) the nag_rand_discrete module was used to generate a random number vector, and each variable followed a binomial distribution with parameters N and p_{ij}; and (3) each value of the random number vector was assigned to one location in the sociomatrix. The value was assigned to x_{ij} if the random number generated by the nag_rand_uniform module was greater than .5 or to x_{ji} if it was less than or equal to .5; (4) Steps 2 and 3 were iterated 100,000 times; (5) the DC and skew-symmetry statistics were calculated for each sociomatrix, and their empirical distributions were estimated; (6) Steps 1 to 5 were iterated for each experimental condition.

Results

Regarding sampling distributions under the null hypothesis, Table 3 shows the averages for the DC and skew-symmetry statistics and their variances for all experimental conditions. The mean value of both statistics depends on the number of individuals, whereas their variances decrease as a function of the amount of behavior in dyads and group sizes. Table 3 also shows the estimated mean square error (MS_e). Given that both indices are biased and their variances are not equal, the MS_e criterion is needed to make comparisons to choose the most appropriate estimator. The formulas to compute the MS_e are as follows:

$$MS_e(\mathrm{DC}) = E^2(\mathrm{DC}) + Var(\mathrm{DC})$$
$$MS_e(\Phi) = E^2(\Phi) + Var(\Phi)$$

In order to make suitable comparisons, the variance of the skew-symmetry statistic was multiplied by 4 and its mathematical expectancy was multiplied by 2. It should be noted that the DC ranges between 0 and 1 and that the skew symmetry takes values between 0 and .5. Thus, to make comparisons possible, these statistics should be expressed on the same scale, so we have turned the skew-symmetry statistics into a 0–1 scale. That permits a correct comparison between the DC and (a transformation of) the skew-symmetry statistics regarding their estimation properties. It means that this comparison refers to the transformed skew-symmetry statistic—that is, 2Φ.

We provide MS_e values for the DC and the transformed skew-symmetry statistics in all experimental conditions (Figure 1). MS_e values decrease as a function of the group size for both statistics. For instance, Experimental Conditions 1, 6, 11, and 16—where there are five behaviors per dyad and $n = 3, 4, 5$, and 6—show a slow fall in MS_e values in these conditions. It should be noted that MS_e values decrease more quickly as the number of behaviors per dyad increases. For example, in Experimental Conditions 1–5—where $n = 3$ and $N = 5, 10, 20, 30$, and 60—MS_e values fall sharply. MS_e values for the transformed skew-symmetry statistic are lower than those for the DC statistic in all experimental conditions, suggesting that the former should be used as a better estimate of social reciprocity. In addition, as noted above, both statistics must be monotonically correlated. In the 60 experimental conditions studied under the null hypothesis, we found Spearman's coefficients to be greater than .9 (Table 3) and all were statistically significant ($p < .0001$).

Regarding the results of the statistical power analyses for different values of n, N, and α, we found that both statistical tests are powerful enough, since they show acceptable empirical Type II error rates. We show some results corresponding to $\alpha = .05$ and $\alpha = .01$, since both values of α represent the best balance between Type I and Type II error rates (Figures 2 and 3) for $p_{ij} = .7$ and those values of n and N that we included in the simulation study. In fact, the power of the statistical tests for the aforementioned values of α is almost equal to .8 for $n = 6$, $N = 10$, and $p_{ij} = .7$. Statistical power increases if larger values of n and N are considered. Obviously, statistical power is also better if the effect sizes under analysis are more evident. In general, similar results of statistical power are obtained for the DC and skew-symmetry statistics.

Table 3
Results of the Simulation Study for the DC and Skew-Symmetry Statistics Under the Null Hypothesis

Condition (ID)	E(DC)	Var(DC)	$E(\Phi)$	Var(Φ)	MS_e(DC)	$MS_e(\Phi)$	r_s
$n = 3$ $N = 5$ (1)	.3748520	.0196687	.3106900	.0349655	.1601827	.1314938	.9918
$n = 3$ $N = 10$ (2)	.2457940	.0130647	.1730924	.0140623	.0734794	.0440232	.9591
$n = 3$ $N = 20$ (3)	.1759780	.0063293	.0923980	.0047250	.0372976	.0132624	.9620
$n = 3$ $N = 30$ (4)	.1445860	.0041571	.0633764	.0023482	.0250623	.0063647	.9642
$n = 3$ $N = 60$ (5)	.1025900	.0020494	.0324714	.0006583	.0125741	.0017127	.9663
$n = 4$ $N = 5$ (6)	.3752280	.0099300	.3217440	.0190348	.1507261	.1225540	.9774
$n = 4$ $N = 10$ (7)	.2462090	.0065527	.1774866	.0075791	.0671716	.0390806	.9338
$n = 4$ $N = 20$ (8)	.1766000	.0031571	.0942662	.0024928	.0343447	.0113789	.9413
$n = 4$ $N = 30$ (9)	.1445420	.0020760	.0639600	.0012172	.0229684	.0053081	.9448
$n = 4$ $N = 60$ (10)	.1026970	.0010205	.0326652	.0003323	.0115672	.0013993	.9480
$n = 5$ $N = 5$ (11)	.3749710	.0059509	.3261200	.0118074	.1465541	.1181616	.9734
$n = 5$ $N = 10$ (12)	.2463440	.0039535	.1794642	.0047033	.0646388	.0369107	.9259
$n = 5$ $N = 20$ (13)	.1759440	.0019066	.0942292	.0015253	.0328629	.0104044	.9341
$n = 5$ $N = 30$ (14)	.1443890	.0012466	.0641064	.0007373	.0220948	.0048469	.9367
$n = 5$ $N = 60$ (15)	.1025510	.0006163	.0326800	.0002025	.0111330	.0012705	.9396
$n = 6$ $N = 5$ (16)	.3748650	.0039666	.3283700	.0080101	.1444904	.1158370	.9723
$n = 6$ $N = 10$ (17)	.2459210	.0026177	.1797976	.0031574	.0630949	.0354846	.9220
$n = 6$ $N = 20$ (18)	.1762970	.0012583	.0946694	.0010165	.0323389	.0099788	.9298
$n = 6$ $N = 30$ (19)	.1442700	.0008293	.0641312	.0004937	.0216431	.0046065	.9324
$n = 6$ $N = 60$ (20)	.1026150	.0004118	.0327508	.0001361	.0109417	.0012088	.9363
$n = 7$ $N = 5$ (21)	.3752930	.0028397	.3302880	.0057794	.1436845	.1148696	.9715
$n = 7$ $N = 10$ (22)	.2459880	.0018796	.1804346	.0022956	.0623896	.0348522	.9209
$n = 7$ $N = 20$ (23)	.1762020	.0008982	.0948466	.0007320	.0319453	.0097278	.9276
$n = 7$ $N = 30$ (24)	.1444720	.0005925	.0643910	.0003561	.0214646	.0045023	.9307
$n = 7$ $N = 60$ (25)	.1024890	.0002930	.0326816	.0000973	.0107970	.0011654	.9329
$n = 8$ $N = 5$ (26)	.3750660	.0021339	.3307940	.0043764	.1428084	.1138011	.9713
$n = 8$ $N = 10$ (27)	.2460130	.0014099	.1807674	.0017323	.0619323	.0344091	.9194
$n = 8$ $N = 20$ (28)	.1761650	.0006772	.0948898	.0005513	.0317114	.0095553	.9261
$n = 8$ $N = 30$ (29)	.1444560	.0004427	.0643634	.0002651	.0213103	.0044077	.9282
$n = 8$ $N = 60$ (30)	.1025390	.0002185	.0327360	.0000727	.0107328	.0011444	.9320
$n = 9$ $N = 5$ (31)	.3750940	.0016347	.3314420	.0033696	.1423302	.1132234	.9712
$n = 9$ $N = 10$ (32)	.2459750	.0010913	.1809092	.0013463	.0615950	.0340745	.9193
$n = 9$ $N = 20$ (33)	.1761670	.0005294	.0949976	.0004330	.0315642	.0094576	.9254
$n = 9$ $N = 30$ (34)	.1445480	.0003427	.0644402	.0002060	.0212368	.0043586	.9268
$n = 9$ $N = 60$ (35)	.1024900	.0001702	.0327038	.0000565	.0106744	.0011260	.9311
$n = 10$ $N = 5$ (36)	.3750270	.0013200	.3317480	.0027207	.1419652	.1127775	.9711
$n = 10$ $N = 10$ (37)	.2461400	.0008728	.1813210	.0010785	.0614577	.0339558	.9178
$n = 10$ $N = 20$ (38)	.1762770	.0004214	.0951084	.0003465	.0314949	.0093921	.9252
$n = 10$ $N = 30$ (39)	.1444230	.0002769	.0643976	.0001667	.0211349	.0043137	.9274
$n = 10$ $N = 60$ (40)	.1025660	.0001377	.0327614	.0000455	.0106575	.0011188	.9310
$n = 15$ $N = 5$ (41)	.3749280	.0005633	.3325260	.0011660	.1411343	.1117396	.9708
$n = 15$ $N = 10$ (42)	.2459770	.0003730	.1814628	.0004626	.0608777	.0333913	.9173
$n = 15$ $N = 20$ (43)	.1762080	.0001808	.0951576	.0001485	.0312300	.0092034	.9230
$n = 15$ $N = 30$ (44)	.1444130	.0001193	.0644478	.0000721	.0209745	.0042256	.9259
$n = 15$ $N = 60$ (45)	.1025910	.0000587	.0327906	.0000195	.0105836	.0010947	.9275
$n = 20$ $N = 5$ (46)	.3750080	.0003121	.3329520	.0006484	.1409431	.1115055	.9710
$n = 20$ $N = 10$ (47)	.2460500	.0002088	.1816338	.0002595	.0607494	.0332504	.9176
$n = 20$ $N = 20$ (48)	.1762190	.0000991	.0951984	.0000816	.0311523	.0091444	.9230
$n = 20$ $N = 30$ (49)	.1445230	.0000657	.0645374	.0000397	.0209526	.0042047	.9260
$n = 20$ $N = 60$ (50)	.1025960	.0000322	.0327908	.0000107	.0105581	.0010859	.9279
$n = 25$ $N = 5$ (51)	.3750010	.0001992	.3330900	.0004131	.1408249	.1113621	.9710
$n = 25$ $N = 10$ (52)	.2461450	.0001315	.1817696	.0001633	.0607188	.0332035	.9164
$n = 25$ $N = 20$ (53)	.1762110	.0000629	.0952168	.0000519	.0311132	.0091182	.9219
$n = 25$ $N = 30$ (54)	.1444570	.0000415	.0644936	.0000251	.0209093	.0041845	.9253
$n = 25$ $N = 60$ (55)	.1025650	.0000206	.0327766	.0000069	.0105402	.0010812	.9292
$n = 30$ $N = 5$ (56)	.3749860	.0001363	.3331720	.0002830	.1407508	.1112866	.9706
$n = 30$ $N = 10$ (57)	.2461420	.0000904	.1817822	.0001123	.0606763	.0331571	.9161
$n = 30$ $N = 20$ (58)	.1761840	.0000439	.0951948	.0000362	.0310847	.0090982	.9226
$n = 30$ $N = 30$ (59)	.1445120	.0000289	.0645456	.0000175	.0209126	.0041836	.9254
$n = 30$ $N = 30$ (60)	.1025720	.0000141	.0327800	.0000047	.0105351	.0010792	.9267

Note—Table shows mathematical expectancy (E), variance (Var), and mean square error (MS_e) for both statistics. Spearman's correlation coefficients between the DC and Φ statistics (r_s) are shown in the last column for all experimental conditions. $E(\Phi)$ and Var(Φ) are expressed on a 0–1 scale to make possible comparisons with the DC statistic.

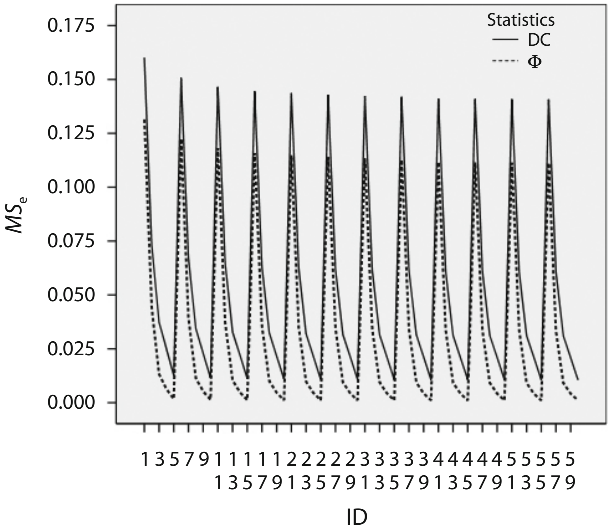

Figure 1. Mean square error (MS_e) of the DC and the skew-symmetry statistics for the 60 conditions (ID) under the null hypothesis.

Discussion

In the present study, we have focused on two overall indices: the DC index for quantifying directionality (van Hooff & Wensing, 1987), and the recently proposed skew-symmetry index for describing asymmetrical social systems (Solanas et al., 2006). These two indices are measures of social reciprocity at group level. However, most researchers require a procedure not only for describing social systems as a whole, but also for allowing researchers to make statistical decisions.

We propose a statistical method founded on Monte Carlo sampling to test null hypotheses. Although most social researchers will be interested in testing the null hypothesis that assumes complete reciprocation, the proposed procedure also enables them to test other null hypotheses (for instance, $p_{12} = .7$, $p_{13} = .8$, and $p_{23} = .5$). Hence, one advantage of this procedure is its flexibility and adaptability regarding both the different number of null hypotheses that can be tested and the observed number of behaviors in empirical studies. Regarding the latter point, it is improbable that the same number of encounters occurs for all dyads in a group in natural settings (de Vries, 1998). Given that most social research is carried out in natural settings, a statistical method for testing social reciprocity in a wider set of conditions is needed. Specifically, the amount of behavior per dyad should not be restricted to an equal amount for each dyad. The proposed statistical procedure does not involve any constraint regarding the number of behaviors per dyad.

The proposed procedure's main drawback is that it requires three assumptions to be met in order for statistical decisions to be made. In fact, these assumptions have been commonly made when developing indices and statistical tests for quantifying and making decisions about social relations (Appleby, 1983; Boyd & Silk, 1983; de Vries, 1995; Hemelrijk, 1990a; Kenny et al., 2006; Landau, 1951; Rapoport, 1949; Warner et al., 1979). Specifically, we have assumed that p_{ij} values are constant for every trial during the observation period, that outcomes of successive interactions are independent, and that dyad behaviors are not influenced by extradyadic effects. The first assumption suggests that social researchers would apply the proposed procedure if data were gathered for periods of time that were as short as possible. The validity of this assumption could be called into question if sociomatrix data were obtained for long periods of time, although it has been often assumed in social interaction analysis (Adams, 2005; Boyd & Silk, 1983; Tufto et al., 1998). The second and third assumptions cannot be realistically made in most social research; but it is not possible to estimate dependency effects, if available data correspond to aggregated sociomatrices. Many social studies analyze aggregated data, since a large number of observation periods is needed to gather

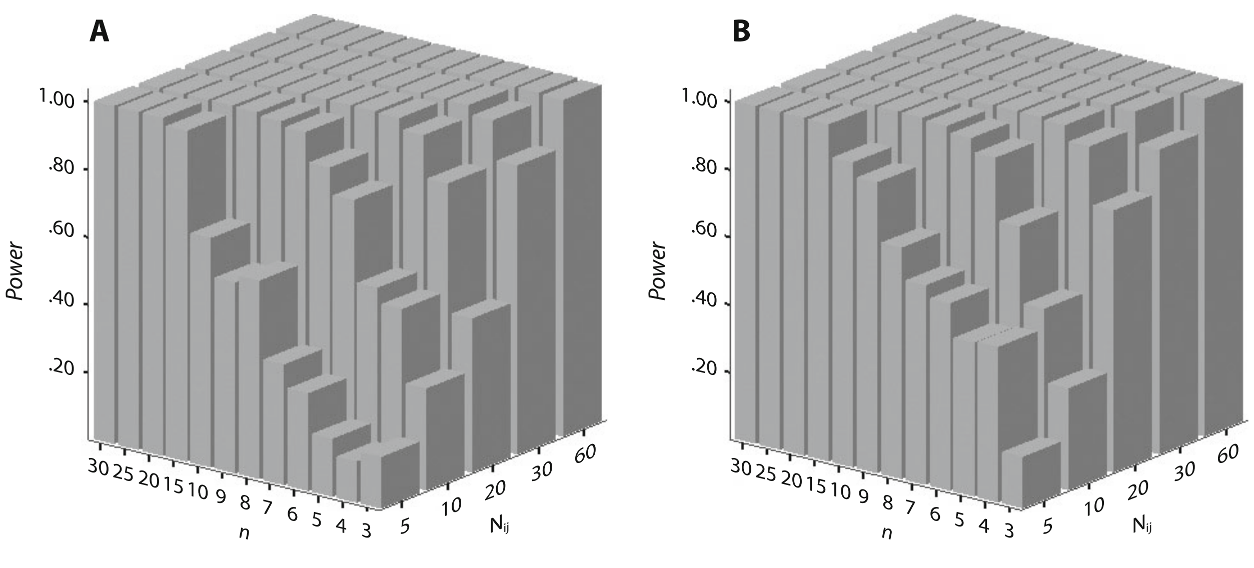

Figure 2. Statistical power for the DC statistic for several values of n, N_{ij}, and $p_{ij} = .7$. (A) Results for $\alpha = .01$. (B) Results for $\alpha = .05$.

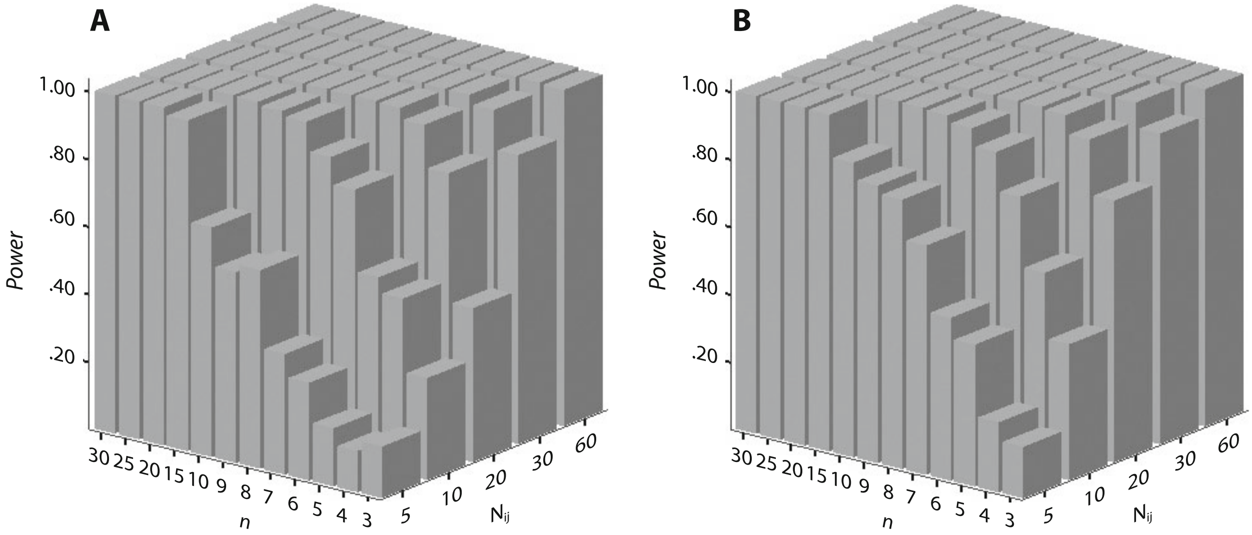

Figure 3. Statistical power for the skew-symmetry statistic for several values of n, N_{ij}, and $p_{ij} = .7$. (A) Results for $\alpha = .01$. (B) Results for $\alpha = .05$.

a significant amount of behavior for all dyads. As long as data are aggregated in a unique sociomatrix, dependency between successive interactions and pairs of dyads cannot be estimated. Unfortunately, both assumptions may not be often met in many social studies. Regarding the assumption of independence between successive interactions, the nonexistence of a general strategy for controlling this sort of order effect in round-robin designs has been pointed out in a recent work (Kenny et al., 2006, p. 217). In relation to the assumption of independence between pairs of dyads, it has also been stated that this assumption is needed in the SRM (Warner et al., 1979). These three constraints mean that the proposed procedure should be used in those natural or experimental settings in which the three assumptions could be supposed to approximately represent empirical phenomena. In any case, future research should be carried out to develop techniques for dealing with dependency between successive interaction results and between pairs of dyads.

Regarding the comparison between the DC and skew-symmetry statistics, the results of the simulations show that the statistical tests are powerful enough for the studied conditions. For instance, the power of the tests is approximately .8 for $n = 6$, $N = 10$, and $p_{ij} = .7$. This means that both tests are sensitive to moderate discrepancies from overall reciprocity, since $p_{ij} = .7$ represents groups that are relatively close to reciprocation—or, as it might be better to say, close to an egalitarian social interaction pattern. Furthermore, it should be noted that 6 individuals, and 10 trials per dyad, are not extremely large conditions in social studies. A transformation of the skew-symmetry statistical test shows better results than does the DC test, if MS_e is considered. Thus, it seems that this transformation of the skew-symmetry statistic is the best choice in order to obtain more accurate estimates of social reciprocity, if complete reciprocation is assumed. Moreover, the skew-symmetry statistic allows researchers to obtain quantifications of individual and dyadic effects that could be of interest to social researchers for studying social reciprocity at its different levels (see Solanas et al., 2006).

In our simulation study, we have studied the DC and the skew-symmetry statistics for a set of particular conditions. Regarding the amount of behavior per dyad, because the number of possible sociomatrices is infinite, we established a constant number of encounters. Thus, establishing an equal number of N_{ij} for each dyad in sociomatrices allowed us to study Type II error rates in a systematic manner and to learn how this kind of statistical error decreases as the number of individuals in a group increases for the DC and skew-symmetry statistics.

To sum up: This article presents a statistical procedure for testing null hypotheses concerning the global social reciprocity for a set of individuals in a group. Therefore, our work enables social researchers to make statistical decisions about directionality and skew symmetry in groups. The results of the simulation study enable researchers to make decisions about the optimal observational conditions for the null hypothesis that assumes complete reciprocation, since empirical Type II error rates for both statistical tests have been estimated. Finally, we emphasize that this statistical method can be applied in natural settings in which the number of behaviors is not the same for every dyad.

AUTHOR NOTE

This research was partially supported by the Spanish Ministry of Education and Science Grant SEJ2005-07310-C02-01/PSIC and by the Generalitat of Catalonia's Ministry of Universities, Research and the Information Society Grants 2005SGR00098 and 2007FIC00747. We thank Carles Riba and two anonymous reviewers for their comments on the manuscript. We also thank Laia Palou for allowing us partial use of data recorded for her doctoral research. Additional figures and the SAS and R programs are available on request. Correspondence concerning this article should be addressed to D. Leiva, Departament de Metodologia de les Ciències del Comportament, Facultat de Psicologia, Universitat de Barcelona, Passeig de la Vall d'Hebron 171, 08035 Barcelona, Spain (e-mail: dleivaur@ub.edu).

REFERENCES

Adams, E. S. (2005). Bayesian analysis of linear dominance hierarchies. *Animal Behaviour*, **69**, 1191-1201.

Appleby, M. C. (1983). The probability of linearity in hierarchies. *Animal Behaviour*, **31**, 600-608.

Bauer, E. B., & Smuts, B. B. (2007). Cooperation and competition during dyadic play in domestic dogs, *Canis familiaris*. *Animal Behaviour*, **73**, 489-499.

Boyd, R., & Silk, J. B. (1983). A method for assigning cardinal dominance ranks. *Animal Behaviour*, **31**, 45-58.

Constantine, A. G., & Gower, J. C. (1978). Graphical representation of asymmetric matrices. *Applied Statistics*, **27**, 297-304.

Côté, S. D. (2000). Determining social rank in ungulates: A comparison of aggressive interactions recorded at a bait site and under natural conditions. *Ethology*, **106**, 945-955.

de Vries, H. (1995). An improved test of linearity in dominance hierarchies containing unknown or tied relationships. *Animal Behaviour*, **50**, 1375-1389.

de Vries, H. (1998). Finding a dominance order most consistent with a linear hierarchy: A new procedure and review. *Animal Behaviour*, **55**, 827-843.

Hemelrijk, C. K. (1990a). A matrix partial correlation test used in investigations of reciprocity and other social interaction patterns at group level. *Journal of Theoretical Biology*, **143**, 405-420.

Hemelrijk, C. K. (1990b). Models of, and tests for, reciprocity, unidirectionality and other social interaction patterns at a group level. *Animal Behaviour*, **39**, 1013-1029.

Kenny, D. A. (1994). *Interpersonal perception: A social relations analysis*. New York: Guilford.

Kenny, D. A., Kashy, D. A., & Cook, W. L. (2006). *Dyadic data analysis*. New York: Guilford.

Kenny, D. A., & La Voie, L. J. (1984). The social relations model. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 18, pp. 142-182). San Diego: Academic Press.

Koenig, A., Larney, E., Lu, A., & Borries, C. (2004). Agonistic behavior and dominance relationships in female Phayre's leaf monkeys—preliminary results. *American Journal of Primatology*, **64**, 351-357.

Landau, H. G. (1951). On dominance relations and the structure of animal societies: I. Effect of inherent characteristics. *Bulletin of Mathematical Biophysics*, **13**, 1-19.

Manly, B. F. J. (2007). *Randomization, bootstrap and Monte Carlo methods in biology* (3rd ed.). Boca Raton, FL: Chapman & Hall/CRC.

Noreen, E. W. (1989). *Computer-intensive methods for testing hypotheses: An introduction*. New York: Wiley.

Onghena, P., & May, R. B. (1995). Pitfalls in computing and interpreting randomization test *p* values: A commentary on Chen and Dunlap. *Behavior Research Methods, Instruments, & Computers*, **27**, 408-411.

Pelletier, F., & Festa-Bianchet, M. (2006). Sexual selection and social rank in bighorn rams. *Animal Behaviour*, **71**, 649-655.

Peres-Neto, P. R., & Olden, J. D. (2001). Assessing the robustness of randomization tests: Examples from behavioural studies. *Animal Behaviour*, **61**, 79-86.

Rapoport, A. (1949). Outline of a probabilistic approach to animal sociology: I. *Bulletin of Mathematical Biophysics*, **11**, 183-191.

Solanas, A., Salafranca, L., Riba, C., Sierra, V., & Leiva, D. (2006). Quantifying social asymmetric structures. *Behavior Research Methods*, **38**, 390-399.

Stevens, J. M. G., Vervaecke, H., de Vries, H., & van Elsacker, L. (2005). Peering is not a formal indicator of subordination in bonobos (*Pan paniscus*). *American Journal of Primatology*, **65**, 255-267.

Tufto, J., Solberg, E. J., & Ringsby, T.-H. (1998). Statistical models of transitive and intransitive dominance structures. *Animal Behaviour*, **55**, 1489-1498.

van Hooff, J. A. R. A. M., & Wensing, J. A. B. (1987). Dominance and its behavioral measures in a captive wolf pack. In H. W. Frank (Ed.), *Man and wolf* (pp. 219-252). Dordrecht, Netherlands: Junk Publishers.

Vervaecke, H., de Vries, H., & van Elsacker, L. (1999). An experimental evaluation of the consistency of competitive ability and agonistic dominance in different social contexts in captive bonobos. *Behaviour*, **136**, 423-442.

Vogel, E. R. (2005). Rank differences in energy intake rates in white-faced capuchin monkeys, *Cebus capucinus*: The effects of contest competition. *Behavioral Ecology & Sociobiology*, **58**, 333-344.

Warner, R. M., Kenny, D. A., & Stoto, M. (1979). A new round robin analysis of variance for social interaction data. *Journal of Personality & Social Psychology*, **37**, 1742-1757.

(Manuscript received June 27, 2007;
revision accepted for publication November 2, 2007.)