ANIMAL BEHAVIOUR, 2006, **71**, 585–592
doi:10.1016/j.anbehav.2005.05.015

Available online at www.sciencedirect.com

SCIENCE DIRECT

Measuring and testing the steepness of dominance hierarchies

HAN DE VRIES*, JEROEN M. G. STEVENS†‡ & HILDE VERVAECKE†‡

*Department of Behavioural Biology, Utrecht University

†Department of Biology, University of Antwerp

‡Centre for Research and Conservation, Royal Zoological Society of Antwerp

(Received 10 February 2005; initial acceptance 25 March 2005; final acceptance 24 May 2005; published online 10 February 2006; MS. number: 8464)

In the analysis of social dominance in groups of animals, linearity has been used by many researchers as the main structural characteristic of a dominance hierarchy. In this paper we propose, alongside linearity, a quantitative measure for another property of a dominance hierarchy, namely its steepness. Steepness of a hierarchy is defined here as the absolute slope of the straight line fitted to the normalized David's scores (calculated on the basis of a dyadic dominance index corrected for chance) plotted against the subjects' ranks. This correction for chance is an improvement of an earlier proposal by de Vries (appendix 2 in de Vries, *Animal Behaviour*, 1998, **55**, 827–843). In addition, we present a randomization procedure for determining the statistical significance of a hierarchy's steepness, which can be used to test the observed steepness against the steepness expected under the null hypothesis of random win chances for all pairs of individuals. Whereas linearity depends on the number of established binary dominance relationships and the degree of transitivity in these relationships, steepness measures the degree to which individuals differ from each other in winning dominance encounters. Linearity and steepness are complementary measures to characterize a dominance hierarchy.

Dominance hierarchies can be characterized in terms of two properties: linearity and steepness. Although a measure of linearity, along with a statistical test procedure, is available (de Vries 1995), a quantitative operational measure of the steepness of a dominance hierarchy does not exist. In an often-quoted paper about primate socioecology, van Schaik (1989, page 206) used the terms 'egalitarian' and 'despotic' (see also Vehrencamp 1983) to describe dominance hierarchies that are 'weakly linear and shallow' and 'steep and linear', respectively. Although the term 'steepness' was thus introduced conceptually, an operational measure, along with a statistical test procedure, to be used in empirical studies has not been provided. Nevertheless, the concept has since been used (sometimes referred to as 'dominance gradient') in several behavioural studies such as biological market models (e.g. Barrett et al. 1999; Henzi & Barrett 1999; Leinfelder et al. 2001) and theoretical modelling studies by Hemelrijk (1999) and Hemelrijk & Gygax (2004), who used the coefficient of variation as a measure of rank differentiation. The concept is further pivotal in the realms of social power and dominance styles (Flack & de Waal 2004) and reconciliation (e.g. Thierry 2000; Demaria & Thierry 2001).

Linearity in a set of binary dominance relationships depends on the number of established relationships and on the degree to which these relationships are transitive (Landau 1951; Kendall 1962; Appleby 1983; de Vries 1995). The steepness of a dominance hierarchy refers to the size of the absolute differences between adjacently ranked individuals in their overall success in winning dominance encounters (i.e. dominance success). When these differences are large the hierarchy is steep; when they are small the hierarchy is shallow. Whereas linearity is based on the binary dyadic dominance relationships, steepness requires a cardinal rank measure (Flack & de Waal 2004).

Two broad types of methods can be used to produce a linear hierarchy (reviewed in de Vries 1998; also Jameson et al. 1999; de Vries & Appleby 2000; Albers & de Vries 2001; Gammell et al. 2003). In the first type the dominance matrix is reorganized such that a numerical criterion, calculated for the matrix as a whole, is minimized or maximized. This yields an ordinal rank order. The second type provides a suitable measure of individual overall

Correspondence: H. de Vries, Department of Behavioural Biology, Padualaan 14, Postbus 80.086, 3508 TB Utrecht, The Netherlands (email: J.deVries@bio.uu.nl). J. M. G. Stevens and H. Vervaecke are at the Department of Biology, University of Antwerp, Universiteitsplein 1 B 2610 Wilrijk, Belgium and at the Centre for Research and Conservation, Royal Zoological Society of Antwerp, Koningin Astridplein 26, B-2018 Antwerp, Belgium.

0003–3472/05/$30.00/0

success from which a cardinal rank order can be directly derived. Such overall success measures can be used to calculate a steepness measure of a hierarchy.

Recently, Gammell et al. (2003) showed that David's score (David 1987, 1988) appears to be the most suitable of the proposed measures of individual overall success. David's score is based on an unweighted and a weighted sum of the individual's dyadic proportions of wins combined with an unweighted and a weighted sum of its dyadic proportions of losses. The crucial advantage of the David's score is that the overall success of an individual is determined by weighting each dyadic success measure by the unweighted estimate of the interactant's overall success, so that relative strengths of the other individuals are taken into account. Thus, defeating a high-ranking animal is weighted heavier than defeating a low-ranking one.

In the current paper we develop a measure of steepness of a hierarchy based on David's scores. More specifically, we present a formula for normalizing the David's scores, which can be used to define an initial steepness measure. Next, we present a dyadic dominance index corrected for chance to be used in the calculation of the David's scores. The final steepness measure is based on these adjusted, normalized David's scores. Finally, we describe a randomization test procedure for calculating the statistical significance of the steepness of a hierarchy.

NORMALIZING DAVID'S SCORES

The procedure for calculating David's score (DS) for each individual in a group of N individuals on the basis of the observed numbers of dyadic wins and losses is as follows. First, the dyadic proportions of wins are calculated. The proportion of wins by individual i in its interactions with another individual j (P_{ij}) is the number of times that i defeats j (s_{ij}) divided by the total number of interactions between i and j (n_{ij}), i.e. $P_{ij} = s_{ij}/n_{ij}$. The proportion of losses by i in its interactions with j (P_{ji}) equals $1 - P_{ij}$. If $n_{ij} = 0$ then $P_{ij} = 0$ and $P_{ji} = 0$ (David 1988). DS for each member, i, of a group is calculated with the formula:

$$DS = w + w_2 - l - l_2$$

Here w represents the sum of i's P_{ij} values, i.e. $w = \sum P_{ij}$ ($j = 1 \ldots N$; $j \neq i$); w_2 represents a weighted sum of i's P_{ij} values (weighted by the w values of its interactants), i.e. $w_2 = \sum w_j P_{ij}$ ($j = 1 \ldots N$; $j \neq i$). Similarly, l represents the sum of i's P_{ji} values, i.e. $l = \sum P_{ji}$ ($j = 1 \ldots N$; $j \neq i$); l_2 represents a weighted sum of i's P_{ji} values (weighted by the l values of its interactants), i.e. $l_2 = \sum l_j P_{ji}$ ($j = 1 \ldots N$; $j \neq i$) (David 1988, page 108; de Vries 1998). Tables 1 and 2 give a worked example showing these calculations. Table 1 presents the numbers of wins and losses of dyadic dominance encounters in a group of seven bonobos, *Pan paniscus* (J. M. G. Stevens, unpublished data). Table 2 presents the proportions of wins (P_{ij}) and the calculated w, w_2, l and l_2 values and the resulting DS values. For instance, for the bonobo Dz $w(Dz) = 0.0 + 1.0 + 1.0 + 0.0 + 0.99 + 1.0 = 3.99$ and $w_2(Dz) = 5.0 \times 0.0 + 3.94 \times 1.0 + 2.0 \times 1.0 + 1.29 \times 0.0 + 1.53 \times 0.99 + 0.25 \times 1.0 = 7.71$. Similarly, $l(Dz) = 0.01$ and $l_2(Dz) = 0.05$ and thus $DS(Dz) = w + w_2 - l - l_2 = 3.99 + 7.71 - 0.01 - 0.05 = 11.64$.

Table 1. Dominance interaction matrix with numbers of wins and losses among seven bonobos

	He	Dz	Ho	De	Ko	Re	Ki
He	*	0	1	2	10	63	8
Dz	0	*	2	3	0	88	4
Ho	0	0	*	4	65	84	3
De	0	0	0	*	0	80	10
Ko	0	0	0	0	*	4	1
Re	0	1	5	0	10	*	6
Ki	0	0	0	0	0	2	*

Data from J. M. G. Stevens, unpublished.

To obtain a steepness measure that varies between 0 and 1, it is necessary to convert DS into a normalized DS (NormDS) as follows:

$$NormDS = \{DS + MaxDS(N)\}/N = \{DS + N(N-1)/2\}/N,$$

where $MaxDS(N)$ is the highest potential David's score that can be obtained by an individual in a group of size N. The structure of this formula can be reasoned as follows. Consider a group of individuals in which a perfect linear dominance hierarchy exists and for which the wins in every dyad are completely unidirectionally distributed (i.e. for each dyad (i,j) either i or j has won all dominance encounters between i and j). The wins and losses in such a group can be presented in a matrix in which the animals have been ordered from highest to lowest rank. This matrix contains all zeroes in the left lower triangular half and positive counts in the upper triangular half. It can easily be seen that for such a matrix the DS of the top-ranking animal equals $N(N-1)/2$ and the DS of the animal at the bottom equals $-N(N-1)/2$. By adding $N(N-1)/2$ (i.e. the maximum DS value) to the DS of each animal, the scores will lie between 0 and $2 \times N(N-1)/2$. By dividing this score by N we arrive at a normalized David's score that varies between 0 and $N-1$. The last column in Table 2 presents the normalized David's scores for the bonobo example.

STEEPNESS MEASURE

When the animals, ranked from the highest rank 1 to the lowest rank N in the rank order found by NormDS, are put on the X axis, and are given the normalized DS value on the Y axis, ordinary least-squares linear regression can be used to find the best-fitting straight line. We propose to use the absolute value of the slope of this line as a measure of steepness of the dominance hierarchy. Figure 1 shows this regression line for the bonobo data. The equation of this line is $Y = -0.74X + 5.94$, so the steepness of this hierarchy is 0.74.

In general, the steepness can vary between 0 and 1 when the normalized DS is used. When there is perfect linearity in the set of dominance relationships and when all proportions of wins P_{ij} are 1, the slope equals -1, and steepness is thus at its maximum 1 (Fig. 2). When the

Table 2. Matrix of proportions of wins (P_{ij}), matrix of dyadic dominance indices corrected for chance (D_{ij}) and the values for w, w_2, l and l_2 used to calculate David's score (DS) and the normalized DS (NormDS)

	He	Dz	Ho	De	Ko	Re	Ki	w	w_2	DS	NormDS
Win proportions											
He	*	0.0	1.0	1.0	1.0	1.0	1.0	5.00	9.01	14.01	5.00
Dz	0.0	*	1.0	1.0	0.0	0.99	1.0	3.99	7.71	11.64	4.66
Ho	0.0	0.0	*	1.0	1.0	0.94	1.0	3.94	4.98	6.61	3.94
De	0.0	0.0	0.0	*	0.0	1.0	1.0	2.00	1.78	−1.29	2.82
Ko	0.0	0.0	0.0	0.0	*	0.29	1.0	1.29	0.69	−5.99	2.14
Re	0.0	0.01	0.06	0.0	0.71	*	0.75	1.53	1.37	−8.73	1.75
Ki	0.0	0.0	0.0	0.0	0.0	0.25	*	0.25	0.38	−16.25	0.68
l	0.00	0.01	2.06	3.00	2.71	4.47	5.75				
l_2	0.00	0.05	0.26	2.07	5.25	7.16	11.13				
Dyadic dominance indices											
He	*	0.00	0.75	0.83	0.95	0.99	0.94	4.47	8.40	10.99	4.57
Dz	0.00	*	0.83	0.88	0.00	0.98	0.90	3.59	7.74	9.69	4.38
Ho	0.25	0.17	*	0.90	0.99	0.94	0.88	4.12	7.09	7.39	4.06
De	0.17	0.13	0.10	*	0.00	0.99	0.95	2.34	3.93	0.86	3.12
Ko	0.05	0.00	0.01	0.00	*	0.30	0.75	1.10	1.33	−7.26	1.96
Re	0.01	0.02	0.06	0.01	0.70	*	0.72	1.51	1.75	−8.85	1.74
Ki	0.06	0.10	0.13	0.05	0.25	0.28	*	0.85	1.93	−12.82	1.17
l	0.53	0.41	1.88	2.66	2.90	4.49	5.15				
l_2	1.37	1.23	1.94	2.75	6.79	7.63	10.46				

non-normalized DS is used the slope of the fitted straight line can vary between 0 and N. With a normalized DS the maximum value of the steepness measure has become independent of the number of subjects in the group. Furthermore, the steepness measure is generally independent of the number of subjects, which is obviously a desirable feature when slopes of different-sized groups are to be compared (Appendix 1).

So, normalizing DS is a necessary first step to arrive at a suitable steepness measure.

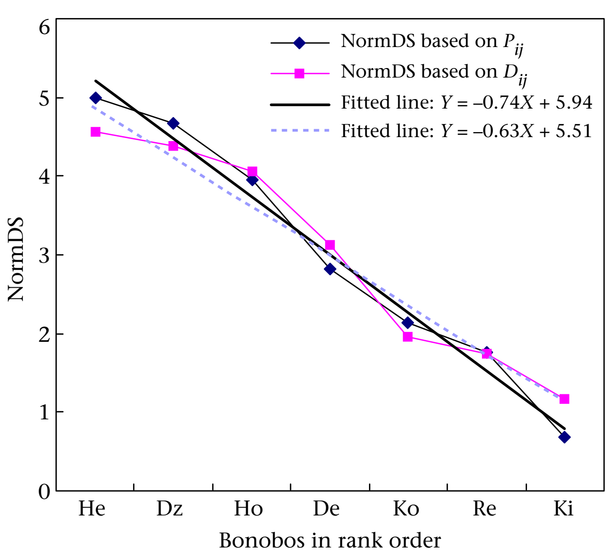


Figure 1. The normalized David's scores (NormDS) plotted against the rank of seven bonobos, He–Ki, ranked from 1 (highest) to 7 (lowest). The straight line fitted through the normalized David's scores based on the proportions of wins (P_{ij}) has slope −0.74, so the steepness of the hierarchy based on these scores is 0.74. The steepness hierarchy based on the normalized David's scores based on the dyadic dominance index (D_{ij}) is 0.63.

THE DYADIC DOMINANCE INDEX CORRECTED FOR CHANCE

So far we have used the dyadic proportions of wins, P_{ij}, in the calculation of the normalized David's scores. However, as noted by David (1988, page 108), these P_{ij} values are not

	a	b	c	d	e
a	*	5	1	1	3
b	0	*	1	2	1
c	0	0	*	4	2
d	0	0	0	*	1
e	0	0	0	0	*

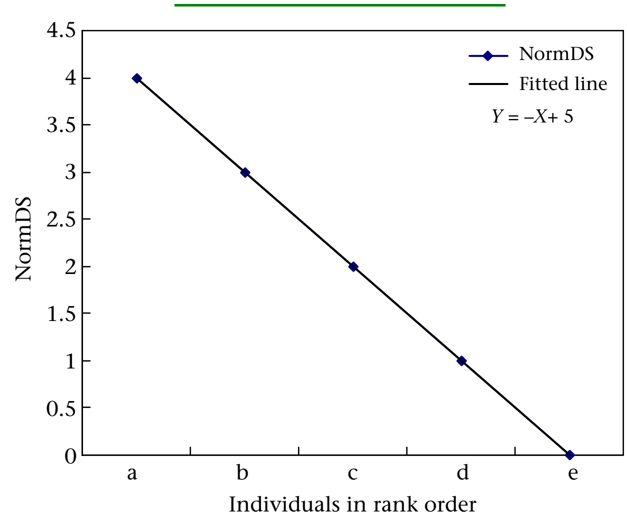


Figure 2. Fictive interaction matrix of wins and losses to show the necessity of normalizing DS. In the graph, the normalized David's scores (NormDS) are plotted against the rank of five animals, a–e, ranked from 1 (highest) to 5 (lowest). The (absolute value of the) slope of the straight line fitted through these values equals 1, indicating maximum steepness of the hierarchy. With non-normalized David's scores the slope would be −5 (not shown in the graph).

wholly satisfactory when the interaction numbers differ greatly between dyads. For instance, whether A defeats B in two of two interactions or A defeats B in 10 of 10 interactions, in both situations P_{AB} equals 1. When estimating A's chances of defeating B we have to take the number of interactions into account. To this purpose de Vries (appendix 2 in de Vries 1998) proposed a dyadic dominance index d_{ij} in which the observed proportion of wins, P_{ij}, is corrected for the chance occurrence of this observed outcome. de Vries proposed calculating the chance occurrence of the observed outcome on the basis of a binomial distribution with each animal having an equal chance of winning or losing in every dominance encounter. In the present paper we propose calculating the chance occurrence of the observed outcome on the basis of a uniform distribution, that is, given a certain number of observed dominance encounters, n_{ij}, then by chance every possible division of these encounters in wins and losses among the two animals is equally likely. For two reasons this is better than the former proposal: (1) this new dominance index corrected for chance, D_{ij}, turns out to be equal to the well-known Bayesian estimator under Jeffreys' prior distribution used for estimating the parameter p in a sequence of Bernoulli trials, and (2) a Monte Carlo simulation study (Appendix 2) shows that D_{ij} performs better as an estimator of the win probability p than the dyadic dominance index d_{ij} that was proposed by de Vries (appendix 2 in de Vries 1998).

To be specific, the dyadic dominance index D_{ij} of i over j, corrected for chance under the assumption that every outcome is equally likely, is defined as:

$$D_{ij} = \text{observed proportion} - \{(\text{observed proportion} - \text{expected proportion}) \times \text{Prob}[\text{observed proportion}]\};$$

that is,

$$D_{ij} = P_{ij} - \{(P_{ij} - 0.5) \times \text{Prob}[P_{ij}]\},$$

where Prob[P_{ij}] is the probability that the observed proportion will occur by chance. This probability is equal to $1/(n_{ij}+1)$.

An example illustrates the calculation of D_{ij}. Suppose $s_{ij} = 4$ and $s_{ji} = 1$, and hence $n_{ij} = 5$, then

$$P_{ij} = 4/5 = 0.8, \text{ while}$$

$$D_{ij} = 4/5 - \{(4/5 - 0.5) \times \text{Prob}[s_{ij} = 4 | n_{ij} = 5 \text{ and each outcome is equally likely}]\} = 0.8 - 0.3(1/6) = 0.75.$$

For $n_{ij} = 5$ interactions there are six possible outcomes of s_{ij}: 0, 1, 2, 3, 4 and 5, corresponding to six possible outcomes of P_{ij}: 0, 0.2, 0.4, 0.6, 0.8 and 1. Under the assumption that each of these outcomes is equally likely to occur, the chance of a particular outcome is 1/6. In general, when individuals i and j have had n_{ij} interactions there are $n_{ij}+1$ equally likely outcomes of P_{ij}. So, the formula for D_{ij} can be rewritten as:

$$D_{ij} = s_{ij}/n_{ij} - \{(s_{ij}/n_{ij} - 0.5)/(n_{ij}+1)\}.$$

It can easily be seen that $D_{ij} = 1 - D_{ji}$, and that for $s_{ij} = s_{ji}$ (not zero) the value of D_{ij} equals 0.5. If $n_{ij} = 0$, D_{ij} and D_{ji} are zero (cf. David 1988). The value of D_{ij} approaches 1 if the difference between s_{ij} and s_{ji} approaches infinity. For example, $s_{ij} = 1$ and $s_{ji} = 0$ gives a D_{ij} of 0.75, while $s_{ij} = 5$ and $s_{ji} = 0$ gives a D_{ij} of 0.917.

By some algebraic manipulation it can be seen that D_{ij} is identically equal to $(s_{ij}+0.5)/(n_{ij}+1)$. In Appendix 2 we explain why D_{ij} is preferable to the earlier proposed dominance index d_{ij} as well as to the simple proportion of wins P_{ij}, the main reason being that D_{ij} is a better estimator of the win probability p than d_{ij} and P_{ij} (Fig. 3).

We now show why the dyadic dominance index D_{ij} should be used rather than the P_{ij} in the calculation of the normalized DS to arrive at a suitable steepness measure. In the bonobo example, Fig. 1 shows the values of NormDS based on D_{ij} plotted against the ranks of the seven bonobos. The steepness of the straight line fitted to the normalized David's scores based on D_{ij} is 0.63, whereas the steepness based on the P_{ij} is 0.74. In this example, differences between NormDS based on P_{ij} and NormDS based on D_{ij} are rather small. However, in the following example comparing two artificial interaction matrices, the benefits of using D_{ij} instead of P_{ij} become

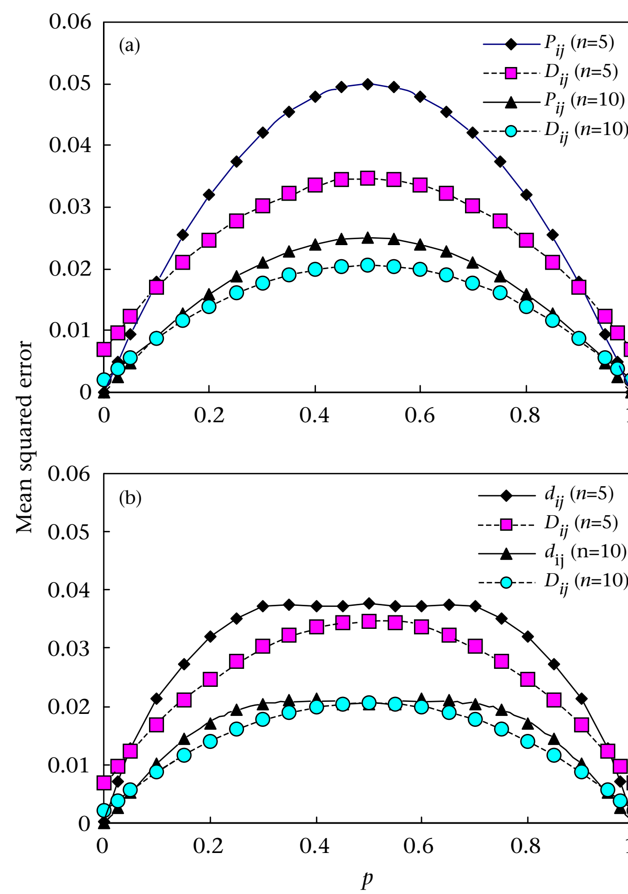
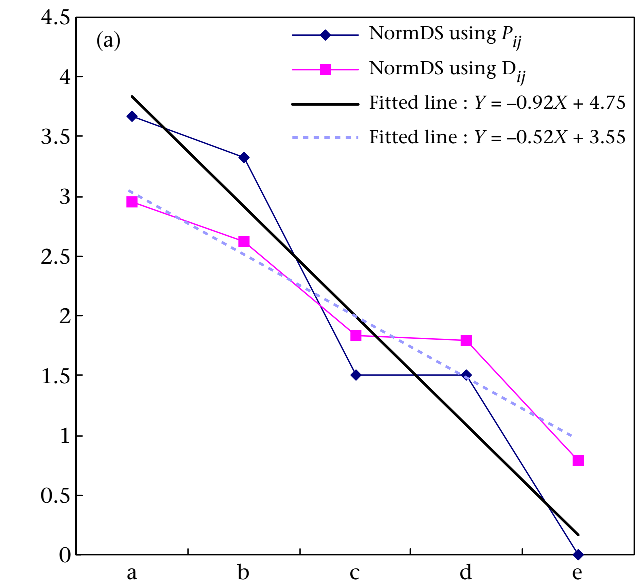


Figure 3. Mean squared error (MSE) of the Bayesian estimator D_{ij} compared to (a) the MSE of the maximum likelihood estimator P_{ij} and (b) the MSE of the estimator d_{ij}. D_{ij} performs better (has smaller MSE) across a large range of win probabilities p. n = number of interactions per dyad.

	a	b	c	d	e
a	*	2	1	2	1
b	1	*	1	1	1
c	0	0	*	1	2
d	0	0	1	*	3
e	0	0	0	0	*



	a	b	c	d	e
a	*	10	5	10	5
b	5	*	5	5	5
c	0	0	*	5	10
d	0	0	5	*	15
e	0	0	0	0	*

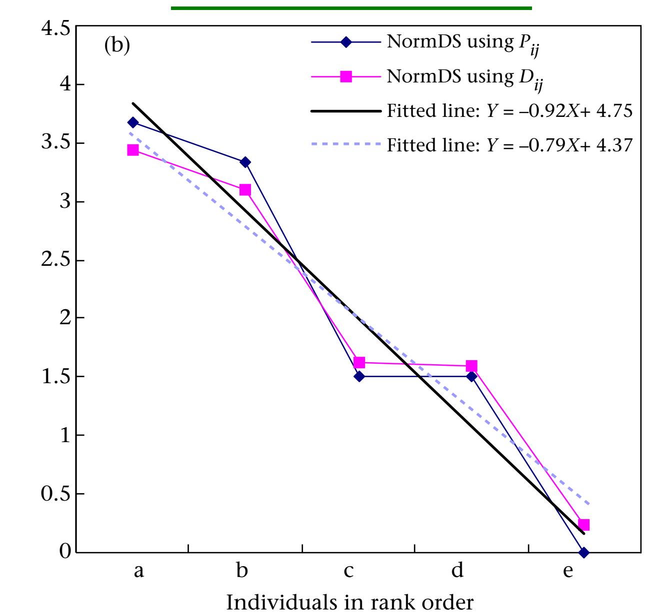


Figure 4. Fictive interaction matrices of wins and losses. In (a) there are five times fewer interactions than in (b). The normalized David's scores (NormDS) using the proportions of wins P_{ij} are plotted against the rank of five animals, a–e, ranked from 1 (highest) to 5 (lowest). The (absolute value of the) slope of the straight line fitted through these points equals 0.92 in (a) and (b). The slope of the straight line fitted through the normalized David's scores using the dyadic dominance index D_{ij} equals 0.52 in (a) and 0.79 in (b).

obvious. The first matrix (Fig. 4a) includes for each dyad five times fewer interactions than the second (Fig. 4b). Figure 4a shows the normalized DS values based on P_{ij} and the normalized DS values based on D_{ij} (called 'NormDS using P_{ij}' and 'NormDS using *Dij*', respectively). The steepness of the fitted line is 0.92 for NormDS using P_{ij} and 0.52 for NormDS using *Dij* (Fig. 4a). So, with NormDS using P_{ij} the steepness would be considered to be high, whereas with NormDS using D_{ij} it is much lower. When the numbers of wins and losses are increased by a factor of five (Fig. 4b), the P_{ij}'s remain unchanged. This property of P_{ij} is undesirable, since it means that winning one of one interaction gives the same P_{ij} value as winning two of two interactions or 10 of 10 interactions (see for instance dyads (He, Ho), (He, De) and (He, Ko) in Table 1); in each case P_{ij} equals 1 (Table 2). In contrast, D_{ij} differentiates between these three cases: D_{ij} for the (1,0) dyad (He, Ho) equals 0.75, for the (2,0) dyad (He, De) it equals 0.83, and for the (10,0) dyad (He, Ko) it equals 0.95 (Table 2). So, with increasing numbers of interactions, the D_{ij} values will approach the P_{ij} values, and thus also the steepness of the dominance order based on D_{ij} will approach the one based on P_{ij}: the steepness based on NormDS using P_{ij} is 0.92, whereas that based on NormDS using D_{ij} is 0.79 (Fig. 4b). From this example it is clear that when the dyadic dominance index D_{ij} is used instead of the simple win proportion P_{ij}, the individual overall success as measured by the normalized David's score takes the number of encounters for each dyad into account. Thus, the use of D_{ij} allows the comparison of matrices containing different interaction frequencies. Because of this desirable feature of D_{ij} compared to P_{ij} the resulting steepness found using the normalized DS based on D_{ij} is more suitable than the one based on P_{ij}.

SIGNIFICANCE TEST

To test whether the observed steepness differs significantly from the steepness to be expected under the null hypothesis of random win chances for all pairs of individuals we can use the following randomization test procedure. Generate for each and every dyad (i,j) a random number of wins r for individual i by randomly drawing a number from the integers 0, 1, 2 … n_{ij}. Then $n_{ij} - r$ will be the number of losses by i from j. Calculate the steepness for the resulting random win–loss matrix.

If this procedure is repeated 2000 times (or more if a more precise estimate of the P value is needed), this creates a null frequency distribution of steepness values. The significance (the right-tailed P value) of the observed steepness can be obtained by calculating the proportion of times that a randomly generated steepness under the null hypothesis is greater than or equal to the actually observed steepness. We applied the test procedure to the bonobo example matrix presented in Table 1. None of the randomly generated steepness values are greater than or equal to the observed steepness of 0.63, so $P < 0.001$, indicating that the degree of steepness of the bonobo dominance hierarchy is significant. In addition, application of the linearity test (de Vries 1995) shows that the degree

Table 3. Dominance interaction matrix with fictive numbers of wins and losses among seven individuals

	a	b	c	d	e	f	g	NormDS
a	*	1	1	4	2	6	10	4.19
b	0	*	4	5	0	10	4	3.49
c	0	2	*	4	65	8	3	3.41
d	2	3	2	*	0	80	10	3.21
e	1	0	0	0	*	6	7	2.60
f	1	8	5	0	2	*	6	2.10
g	4	0	1	8	5	3	*	2.01

The degree of linearity in the set of dominance relationships is strong ($h' = 0.946$) and differs significantly (right-tailed $P = 0.008$) from the expected h' value of 0.375. Yet, the observed steepness of the hierarchy is 0.362 and not significantly different (right-tailed $P = 0.17$) from the steepness value (0.283) to be expected under the null hypothesis that for every dyad in the matrix each division of wins and losses over the two opponents is equally likely.

of linearity in this set of dominance relationships is also significant: $h' = 0.86$ (right-tailed $P = 0.028$).

As a further illustration, Table 3 presents a fictive dominance matrix with a strong and highly significant degree of linearity, while at the same time the degree of steepness is rather low and nonsignificant. This example goes to show that linearity and steepness measure two different characteristics of a dominance hierarchy.

DISCUSSION

Several animal behaviour researchers have suggested that it is important to characterize the dominance hierarchy in terms of being more or less 'egalitarian' or 'despotic' (e.g. Vehrencamp 1983; van Schaik 1989). van Schaik (1989) proposed that dominance hierarchies that are both weakly linear and shallow can be called 'egalitarian', whereas steep, linear hierarchies can be called 'despotic'. Flack & de Waal (2004) stated that attempts to measure steepness were hereto incomplete because of the absence of a quantitative assessment. Here we have presented a methodology to quantify steepness of dominance hierarchies, based upon the same sociometric matrices that can be used to quantify linearity of those hierarchies. The question can now be addressed whether more egalitarian or more despotic dominance hierarchies, in the sense of van Schaik (1989), also lead to a more balanced or more unequal distribution of resources/reproduction (Vehrencamp 1983 and others). Future research will reveal whether the measure that is proposed here fills this gap.

In their model studies, Hemelrijk (1999) and Hemelrijk & Gygax (2004) used the coefficient of variation (standard deviation divided by the mean) of the individual dominance values as a measure of the hierarchical gradient, and they used this term interchangeably with hierarchical differentiation or rank differentiation. Although the coefficient of variation (CV) is indeed a measure of rank differentiation, it is not fully suitable as a measure of the gradient or steepness of a hierarchy. This can easily be seen by comparing the following two sets of dominance values: set A: 14, 9, 7, 2; set B: 23, 17, 14, 5 (these values lie within the range of possible DOM values presented in Figure 2a of Hemelrijk 1999). The CV of set A equals 0.62 and the CV of set B equals 0.51; yet, it is clear that the four animals in set B form a steeper hierarchy than the four animals in set A. When the slope of the fitted linear regression line is taken as a steepness measure no such illogical values are found: for set A a slope of 3.8 and for set B a slope of 5.7.

As in studies of linearity, one should be cautious in interpreting the results with respect to observational zeroes, which arise when two individuals are never observed to interact because of biased sampling or inadequate observation effort (de Vries 1995; Galimberti et al. 2003). Thus, it is important that adequate sampling efforts and observation methodologies are used. When sufficient time has been spent observing all animals of the study group but some dyads have still not had any dominance interactions, one can use circumstantial observations to interpret the relationships of these dyads. First, it might be that during the observation period itself a subordinate animal keeps a safe distance from a dominant one, without having any dominance interactions, because they had established their dominance–subordination relationship well before the observation period started, by means of aggressive dominance interactions (Hemelrijk 1999). When it is possible reliably to observe dyadic 'avoiding at a distance' interactions, the number of dyads with missing values can be reduced by including 'avoiding at a distance' into the class of dominance interactions. Second, when two animals are regularly seen in each other's (close) neighbourhood, without having any dominance interactions with each other, these animals show by their behaviour that a clear dominance–subordination relationship between them is absent. In this case, the zeroes in the respective cells of this dyad are appropriate. Finally, when two animals stay at a relatively large distance from each other (but still clearly belong to the same group), and neither one is clearly responsible for this (i.e. neither one shows 'avoiding at a distance' behaviour), it is likely that their dominance relationship is unresolved, and therefore in this case the zeroes in the respective cells of the dyad are also appropriate.

Within-species comparison of different dominance matrices requires not only similar observational efforts, but also analysis on the same sociospatial level, the latter to minimize the number of structural zeroes. For example, Galimberti et al. (2003) pointed out that population structure affects the likelihood of interaction among elephant seals, *Mirounga leonine*: seals of the same harem have a higher interaction rate than seals belonging to the same population but different harems. Between species, linear and steep hierarchies may be easier to detect among close-knit groups of social animals, but more difficult to find among individuals living in loose societies.

Meanwhile, the concept of steepness, as defined here, has shown its value in comparing characteristics of dominance hierarchies within and between captive groups of bonobos (Stevens et al., in press). Sampling effort and social interaction frequencies between all animals were relatively high, so that observational zeroes were minimal. We found that hierarchies among male bonobos were generally steeper than those among females. We also used the steepness concept to test predictions of biological market models. We found that in bonobo groups with a relatively shallow hierarchy grooming was exchanged reciprocally, whereas grooming was more unidirectional in groups with a steep hierarchy (Stevens et al. 2005). So, steepness is a useful additional measure of dominance hierarchies in animal societies.

Acknowledgments

We thank John Hutchinson and four other referees for their useful comments on the manuscript.

References

Albers, P. C. H. & de Vries, H. 2001. Elo-rating as a tool in the sequential estimation of dominance strengths. *Animal Behaviour*, **61**, 489–495.

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

Barrett, L., Henzi, S. P., Weingrill, T., Lycett, J. E. & Hill, R. A. 1999. Market forces predict grooming reciprocity in female baboons. *Proceedings of the Royal Society of London, Series B*, **266**, 665–670.

Casella, G. & Berger, R. L. 2001. Estimation: point and interval. In: *International Encyclopedia of the Social and Behavioral Sciences* (Ed. by N. J. Smelser & P. B. Baltes), pp. 4744–4749. New York: Elsevier.

David, H. A. 1987. Ranking from unbalanced paired-comparison data. *Biometrika*, **74**, 432–436.

David, H. A. 1988. *The Method of Paired Comparisons*. London: C. Griffin.

Demaria, C. & Thierry, B. 2001. A comparative study of reconciliation in rhesus and Tonkean macaques. *Behaviour*, **138**, 397–410.

Flack, J. C. & de Waal, F. B. M. 2004. Dominance style, social power, and conflict management: a conceptual framework. In: *Macaque Societies: A Model for the Study of Social Organization* (Ed. by B. Thierry, M. Singh & W. Kaumanns), pp. 157–182. New York: Cambridge University Press.

Galimberti, F., Fabiani, A. & Boitani, L. 2003. Socio-spatial levels in linearity analysis of dominance hierarchies: a case study on elephant seals. *Journal of Ethology*, **21**, 131–136.

Gammell, M. P., de Vries, H., Jennings, D. J., Carlin, C. M. & Hayden, T. J. 2003. David's score: a more appropriate dominance ranking method than Clutton-Brock et al.'s index. *Animal Behaviour*, **66**, 601–605.

Gelman, A., Carlin, J. B., Stern, H. S. & Rubin, D. B. 2004. *Bayesian Data Analysis*. 2nd edn. London: Chapman & Hall.

Hemelrijk, C. K. 1999. An individual-oriented model of the emergence of despotic and egalitarian societies. *Proceedings of the Royal Society of London, Series B*, **266**, 361–369.

Hemelrijk, C. K. & Gygax, L. 2004. Dominance style, differences between the sexes and species. *Interaction Studies*, **5**, 131–146.

Henzi, S. P. & Barrett, L. 1999. The value of grooming to female primates. *Primates*, **40**, 47–59.

Jameson, K. A., Appleby, M. C. & Freeman, L. C. 1999. Finding an appropriate order for a hierarchy based on probabilistic dominance. *Animal Behaviour*, **57**, 991–998.

Kendall, M. G. 1962. *Rank Correlation Methods*. 3rd edn. London: C. Griffin.

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.

Leinfelder, I., de Vries, H., Deleu, R. & Nelissen, M. 2001. Rank and grooming reciprocity among females in a mixed-sex group of captive hamadryas baboons. *American Journal of Primatology*, **55**, 25–42.

van Schaik, C. P. 1989. The ecology of social relationships amongst female primates. In: *Comparative Socioecology of Mammals and Man* (Ed. by V. Standen & R. Foley), pp. 195–218. Oxford: Blackwell Scientific.

Stevens, J. M. G., Vervaecke, H., de Vries, H. & van Elsacker, L. 2005. The influence of the steepness of dominance hierarchies on reciprocity and interchange in captive groups of bonobos (*Pan paniscus*). *Behaviour*, **142**, 941–960.

Stevens, J. M. G., Vervaecke, H., de Vries, H. & van Elsacker, L. In press. Sex differences in the steepness of dominance hierarchies in captive bonobo groups. *International Journal of Primatology*.

Thierry, B. 2000. Conflict management patterns across macaque species. In: *Natural Conflict Resolution* (Ed. by F. Aureli & F. B. M. de Waal), pp. 106–128. Berkeley, California: University of California Press.

Vehrencamp, S. L. 1983. A model for the evolution of despotic versus egalitarian societies. *Animal Behaviour*, **31**, 667–682.

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.

de Vries, H. & Appleby, M. C. 2000. Finding an appropriate order for a hierarchy: a comparison of the I&SI and the BBS methods. *Animal Behaviour*, **59**, 239–245.

Appendix 1

To show that the steepness measure is generally independent of the number of subjects, consider matrices of size N ($N = 2, 3, 4, \ldots$) with, in the upper right triangle, the win proportions P_{ij} ($j > i$) equal to some value p (≥ 0.5). It can easily be seen that the normalized DS for individual i equals $(N - i)p + (i - 1)(1 - p)$, and that the steepness is equal to $\mathrm{DS}_i - \mathrm{DS}_{i+1} = (N - i)p + (i - 1)(1 - p) - (N - i - 1)p + (i + 1 - 1)(1 - p) = 2p - 1$. So, indeed the steepness does not depend on N.

Appendix 2

By some algebraic manipulation it can be seen that D_{ij} is identically equal to $(s_{ij} + 0.5)/(n_{ij} + 1)$. This shows that D_{ij} is in fact equal to the Bayesian estimator of the binomial parameter p of the binomial distribution $\mathrm{Bin}(n_{ij};p)$ under the noninformative Jeffreys' prior distribution (Gelman et al. 2004). When a Bayesian estimator under a noninformative prior distribution is required, Jeffreys' prior distribution is usually recommended on the ground that the Bayesian estimator under such a prior distribution is invariant under reparameterization (Jeffreys' invariance principle), a property that is not shared by other prior distributions (Gelman et al. 2004). To compare the performance of the Bayesian estimator D_{ij} in estimating the true win probability p with the performance of the maximum likelihood estimator P_{ij} and the performance of the formerly proposed estimator d_{ij}, the mean squared error

(MSE) can be used (Casella & Berger 2001). Low MSE values indicate better performance of the respective estimator. The MSE of D_{ij} is defined by $E\{(D_{ij} - p)^2\}$, where E is the expected value. The MSE of P_{ij} and d_{ij} are defined similarly. The MSE of D_{ij} and P_{ij} are calculated via equation (13) in Casella & Berger (2001); the MSE of d_{ij} has been obtained by means of a Monte Carlo procedure. Figure 3a shows that the MSE of D_{ij} is smaller than the MSE of P_{ij} when the win probability p lies between about 0.1 and 0.9. Similarly, Fig. 3b shows that the MSE of D_{ij} is smaller than the MSE of d_{ij} when the win probability p lies between about 0.05 and 0.95. For all these reasons, the Bayesian estimator $(s_{ij} + 0.5)/(n_{ij} + 1)$, i.e. the dyadic dominance index D_{ij}, is to be preferred to both the earlier proposed dominance index d_{ij} and the maximum likelihood estimator P_{ij}.