ANIMAL BEHAVIOUR, 2006, **72**, 835–842
doi:10.1016/j.anbehav.2006.01.024

Assessing the reliability of dominance scores for assigning individual ranks in a hierarchy

MAUD POISBLEAU, STEPHANIE JENOUVRIER & HERVE FRITZ
Centre d'Etudes Biologiques de Chizé, CNRS UPR 1934

(Received 28 January 2005; initial acceptance 18 April 2005;
final acceptance 11 January 2006; published online 7 September 2006; MS. number: 8446R)

The dominance score (number of wins divided by the total number of interactions) is the most widely used procedure in field studies to rank individuals. Its reliability depends on the number of interactions on which it is calculated. However, most authors use it without any estimate of the associated error. We describe the precision associated with a dominance score estimate as a function of the number of interactions on which it is based, and hence provide a tool to plan field protocols and effort. The precision error decreases according to a power function with increasing number of interactions, but with more precision for extreme scores for any given number of interactions. We discuss the fact that the minimum number of interactions should be based on the precision associated with the 50% score, the least precise of all scores. We also emphasize the trade-off between recording effort and precision of the estimator, and give an example of our choice of 26 interactions for fieldwork on ducks and geese. When comparing individual ranks based on dominance scores with ranks given by the dominance matrix, we found a good correlation, with more mismatches around the middle of the hierarchy. This was consistent with the precision calculated with our model. We conclude that dominance score is a reliable tool, but conclusions must take into account the number of interactions on which the calculations are done. We also discuss the importance of initial assumptions and sources of bias in field studies.

Dominance relationships have been observed and figured prominently in a variety of studies on social species, ranging from bumblebees to baboons (Gauthreaux 1978; Huntingford & Turner 1987; Piper 1997; Koivula 1999). A dominance hierarchy decreases fighting in a group because individuals learn to evaluate their chances of winning conflicts. Therefore, stable dominant–subordinate relationships may benefit both dominants and subordinates by diminishing the incidence of physically dangerous conflicts, but also by reducing the time devoted to social interactions, thus increasing the time that can be allocated to other key behaviours (Rowell 1974; Bernstein 1981; Archer 1988; Belthoff et al. 1994; Guillemin et al. 2000; Macdonald et al. 2002). Thus, dominance is central to most behavioural studies in highly social or gregarious animals, and has contributed greatly to our understanding of social structure (Drews 1993), and variation in individual fitness (Robinson 1986; Zahavi 1989).

The first step in studying the correlates of social dominance, or including the effect of dominance in any behavioural or ecological study, is to derive a reliable score of dominance status by which individuals can be ranked (Boyd & Silk 1983; de Vries 1998; Jameson et al. 1999; de Vries & Appleby 2000; Gammell et al. 2003). The most common method used to quantify dominance status is to derive dominance scores. The dominance score for a given individual is calculated as the number of interactions won, divided by the total number of interactions in which that individual participated (for details see Ens & Goss-Custard 1984; Lamprecht 1986; Caldow & Goss-Custard 1996). This method involves three important assumptions: (1) encounters are always won by the higher-ranking individual, regardless of the magnitude of the difference in ranks, (2) animals interact with those that are above and below them in rank in proportion to their numbers and (3) a linear dominance hierarchy occurs among animals in the group (Goss-Custard et al. 1995a). Goss-Custard et al. (1995a) discussed how to test these assumptions but, to our knowledge, no study has determined how many interactions are required to obtain

Correspondence: M. Poisbleau, Max Planck Institute for Ornithology, Vogelwarte Radolfzell, Schlossallee 2, 78315 Radolfzell, Germany (email: poisbleau@orn.mpg.de). S. Jenouvrier and H. Fritz are at the Centre d'Etudes Biologiques de Chizé, CNRS UPR 1934, 79360 Beauvoir-sur-Niort, France.

0003–3472/06/$30.00/0

a good estimate of the score in all situations. This is essential if we are to have any confidence in such scores because dominance scores may be imprecise if the number of interactions on which they are calculated is too low, or biased if the sample of individuals is not representative of the population studied.

There is little risk of error when estimating the dominance of highly dominant and very subordinate individuals with few interactions. They always win or lose, respectively. The problem lies with the middle-ranking individuals, which require more interactions to stabilize the calculation of their dominance score. The number of observations available to assess the rank of these individuals will obviously be a trade-off between the desired performance of the estimator and the field constraints. However, several researchers have assigned and then analysed dominance scores without a fixed minimum sample size, and when a minimum was set it varied greatly (e.g. three, Stahl et al. 2001; five, Kikkawa 1980; Arcese & Smith 1985; $\geq$10, Goss-Custard et al. 1995a, b; Caldow & Goss-Custard 1996; $\geq$42, Ens & Goss-Custard 1984).

To assess the reliability of the calculated score, and to take a decision about the minimum number of interactions necessary to calculate it, given the field constraints, one must know the error associated with the score as a function of the sample size on which it is based. Furthermore, to test the validity of the dominance score as an index of true dominance status it is valuable to compare the dominance score with the results of a more accurate method of assigning dominance status, for instance, the matrix method (de Vries et al. 1993). The dominance matrix seems to be one of the most efficient way to classify individuals because the classification takes into account the identity of each opponent and all the interactions are considered. The matrix of dominance is built in such a way as to minimize inconsistencies; hence, it is considered to provide the closest estimate of the true hierarchy (de Vries et al. 1993; de Vries 1995, 1998; de Vries & Appleby 2000). Although it is easy to use this method in captivity where most individuals can be marked, it is much more difficult to use in natural conditions where in most cases relatively few individuals are recognizable. The study of oystercatchers, *Haematopus ostralegus*, by Ens & Goss-Custard (1984) provides an exception to this.

In this study, we investigated for the first time the reliability of the dominance score by quantifying the way in which the error associated with the dominance score changes as a function of the number of interactions on which it is based. We also tested the validity of conventionally derived dominance scores through two comparisons with ranks determined by a dominance matrix, one in captive mallards, *Anas platyrhynchos*, and one in wild dark-bellied brent geese, *Branta bernicla bernicla*.

METHODS

Simulation Approach

To determine the performance of the dominance score estimator in relation to the number of interactions available, we simulated a virtual data set of interactions for a range of individuals of known true dominance score. The model was constructed under the hypothesis of independence of interactions for an individual. We assumed therefore, that individuals move in a large group of conspecifics with which they randomly interact and that there is no memory of the outcome of the previous interaction.

We considered that an individual's dominance score corresponds to the probability (p) of winning an interaction against other individuals. We simulated the pattern of win/lose at each interaction using a Bernoulli drawing of probability p, and estimated the dominance score ($\hat{p}$) for a given number of interactions (n). The output of the procedure gives the mean dominance score ($\bar{\hat{p}}$), which we called the 'apparent score' for 10 000 simulations, and the standard deviation ($\mathrm{SD}(\hat{p})$). We studied a dominance score range of 0.1–50%, and simulated up to 3000 interactions per score. We restricted our results to values of the dominance score equal to and below 50%, because the results for the scores above 50% are symmetrical to those below. The simulation procedure was realized by a FORTRAN program and is available on request from the first author in a DOS format.

The performance of an estimator is evaluated by both its bias and its precision. Under our hypothesis of independence of interactions there was no bias in the estimator. We focused on the precision of the estimator, measured by its standard deviation ($\mathrm{SD}(\hat{p})$). For each true score, p, we plotted $\mathrm{SD}(\hat{p})$ against the number of interactions and fitted a power function to the data. The equation of the power function provides an estimated value of the standard deviation associated with the apparent score, and allows the definition of the 95% confidence interval in which the corresponding true score lies.

Recording Data on Captive Mallards

We carried out the experiments at the Centre d'Etudes Biologiques de Chizé (CEBC), in western France, using adult mallards descended from individuals caught in the wild. The birds had been kept at the CEBC for at least 3 years before the experiments, and were therefore accustomed to their aviary environment. During the day, they were free to move in the field station's yard (approximately 1000 m^2) with food and water ad libitum. At night, they were caged in 200-m^2 aviaries (100 m^2 grass, 100 m^2 concrete), equipped with a 25-m^2 pool. Food in aviaries (ad libitum) was composed of a mixture of crushed corn, wheat and commercial duck food.

For the experiments, birds were kept apart in three separate 10 × 10-m grass areas during the day and in 5 × 10-m concrete aviaries during the night, all equipped with a 10-m^2 pool. We selected individuals randomly from a larger flock of 30 individuals to obtain an even sex ratio ($N = 8$ males, 8 females). Food was still provided ad libitum. While they were in these more restricted areas, we monitored birds three times a day to check for any hyperaggressive behaviour or any injuries. None were observed.

The birds were marked individually with plastic badges glued on their backs with nontoxic glue (3 × 6 cm, with black and white codes). The tags were not strongly glued, and hence fell from the plumage in the week after the trials. No damage to the plumage was apparent, and we checked for any symptoms associated with this marking each time birds were manipulated. The ducks were observed for 45 min in the morning and for 45 min in the afternoon for 5 consecutive days (18–22 December 2001). During each 45-min session, the group of birds was placed in a 2 × 4-m arena and videotaped with a digital video camera recorder (Sony, Digital handycam) positioned 2 m from the enclosure and 2 m from the ground to facilitate reading of the badge codes. No observer was visible to the ducks after the camera was positioned. In the morning session, the camera field (2 × 2 m) included a bowl of 400 g of wheat and a bowl with an area of 700 cm^2 of water. In the afternoon a 1-m^2 pool of clear water replaced the bowl of water and two cameras covered the whole area of the enclosure. All interactions on the videotapes were noted, close to the resources or elsewhere in the enclosure. For each aggressive encounter recorded, we determined the winner and the loser, the latter being the individual moving away from or avoiding the former. The birds were deprived of food for 10 h during the night before each day of observation. This served to increase competition for resources and ensured a sufficient number of interactions during each observation session.

Recording Data on Wild Brent Geese

During the winter 2003–2004, we observed the wintering population of brent geese around Oléron Island (45°56′N, 1°21′W, France). Between 2001 and 2004, we caught and marked more than 100 brent geese. They were caught by cannon-net and a team with long experience in the long-term European brent ringing programme (Coordinator B. Ebbinge, Alterra, Netherlands). The capture was carried out on the beach, close to a watering place where brent geese came in small flocks. Care was taken not to fire the net when geese or other animals could be injured. About 2% of the birds in the wintering population were marked with coded Darvic rings that could be read from a distance of up to 250 m with a ×20–60 telescope. The Darvic rings are those used for the last 14 years in the European brent network, and no significant adverse effects have been reported.

We determined both the dominance scores of the marked individuals and their position within the hierarchy of marked birds by observing interactions between marked individuals and both marked and unmarked individuals. We defined an interaction as a direct confrontation between two birds, ranging from threats with lowered head and neck to active chases with flapping wings (Stahl et al. 2001). We observed the flock from a car and noted any marked participant in an interaction as well as the outcome, while the geese were on the intertidal zone. We considered an agonistic interaction as being won by an individual when the opponent turned and walked or ran away (Stahl et al. 2001).

Comparing Dominance Score and Matrix

We used the dominance score defined above for individuals in the mallard and brent goose data sets calculated with a minimum total number of 26 interactions, defined from our simulation model (see Results). We then compared ranks calculated with this dominance score and ranks determined by a dominance matrix.

We organized interactions between ringed birds into sociometric matrices from which we calculated Kendall's coefficient of linearity K, Landau's index h and the index of linearity h' (de Vries 1995), using MatMan 1.0 (Noldus Information Technology, Wageningen, Netherlands; and see de Vries et al. 1993). Each index varies from 0 (absence of linearity) to 1 (complete linearity). The index h' is based on h and takes into account the existence of unknown relationships, when two members of a dyad have not been observed to interact aggressively with each other. Statistical significance of K is provided by a chi-square test. For the h' index, a resampling process using 10 000 randomizations is performed (de Vries 1995). When the dominance hierarchy was significantly linear, individuals were reordered by a two-step iterative procedure (10 000 sequential trials), finding the rank order most consistent with a linear hierarchy by minimizing the number of inconsistencies and then minimizing the total strength of the inconsistencies (de Vries 1998). Each bird was then assigned a rank from 1 (most dominant) to n, (most subordinate $n > 1$).

We used a linear regression to study the correlation between dominance score rank and matrix rank. If the ranks assigned through the scores reflect the ranks defined by the matrix, then the slope of the regression should not differ from 1 and the Y intercept should not be significantly different from zero. We plotted the residuals from the regression of the score rank on the matrix rank against the dominance scores to see whether medium-ranking individuals (score around 50%) correspond to extreme values in the residuals, that is, if they correspond to the individuals the furthest away from the predicted regression line. This would mean that medium-ranking individuals were more difficult to rank correctly with the dominance score. For our captive mallards, as all individuals were known, we expected that the residuals would truly reflect the error associated with ranking in relation to dominance score. For brent geese, few birds were known compared to those unknown, hence we expected more discrepancies between the score rank and the matrix rank. For all statistical analyses, we used SYSTAT 7.0 (release 9 for Windows; SPSS Inc., Chicago, IL, U.S.A.; Wilkinson 1997).

RESULTS

Dominance Score Reliability

As expected, when the number of interactions increased, the standard deviation of the apparent dominance score decreased for all true scores (Fig. 1). The precision of the estimator first increased rapidly, then more slowly for larger numbers of interactions, eventually reaching an asymptotic value as suggested by the fit of

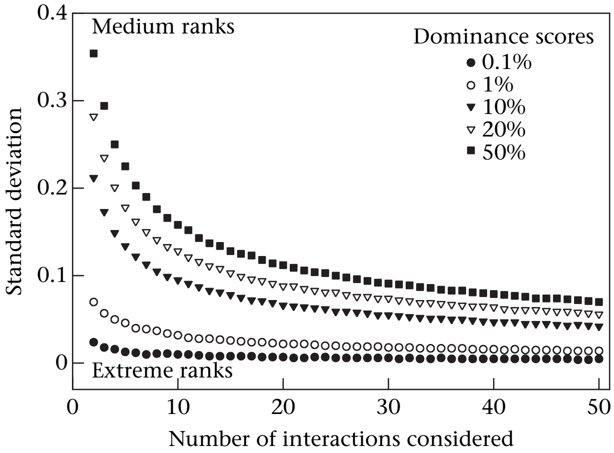

Figure 1. Standard deviation of the score (number of wins divided by the total number of interactions) according to the number of simulated interactions between individuals, for a set of true dominance scores. Standard deviation decreases with interaction number according to the different dominance scores. The asymptote can be used to assess the number of interactions above which increasing the sampling effort does not significantly improve the precision of the estimated score.

a power function to the data (Fig. 1, Table 1). The precision tended to stabilize around a certain number of interactions that varied between scores. For example, for a true score of 50%, the standard deviation was around 7% for 50 interactions, whereas with the same number of interactions the standard deviation was around 1% for a true score of 1% (or 99%; Fig. 1). The true dominance score of 50% was logically the least precise of all. The equations of the power functions in Table 1 allow the value of the standard deviation of the apparent score to be calculated for any number of interactions and provide an indicator of the possible error in attributing the scores.

Because the least precise estimator is that of the 50% score, the minimum number of interactions to record in the field or in an experiment should be determined according to this score. Examination of the simulation results shows that for a bird with a true dominance score of 50%, 26 interactions yield an error around ±10% for the estimated score. To reduce this error to ±5%, 101 interactions would be necessary (Table 1). However, because this number is too high for our logistic constraints, we chose to record at least 26 interactions per individual in our field study on brent geese and in our experiments on mallard ducks that correspond to a precision of ±10% in our dominance score estimation.

Captive Mallards

The mallards' sociometric matrix comprised 1160 interactions. The values of the Kendall's linearity index ($K = 0.51$, $X^2_{23} = 62.7$, $P < 0.0001$), the Landau's index and the corrected index were high ($h = 0.52$, $h' = 0.53$, improved linearity test using h', $P < 0.0001$) indicating that the ranking was neither circular nor random. Therefore, the analysis of this matrix revealed a significant linear hierarchy allowing the use of a linear order to rank individuals (Fig. 2a).

Matrix ranks (X) and dominance score ranks (Y) were significantly related ($Y = 0.912X + 0.75$; $R^2 = 0.831$, $F_{1,14} = 68.995$, $P < 0.001$; Table 2, Fig. 3a). The slope was not significantly different from 1 ($t_{15} = -0.80$, $P = 0.438$) and the Y intercept was not significantly different from 0 ($t_{14} = 0.707$, $P = 0.491$). As predicted, the correlation between ranks was very high for the extreme ranks, but there were more discrepancies for intermediate ranks, around the 50% score (between 35 and 72%), as shown by the plot of residuals of score–matrix rank regression on dominance scores.

Table 1. Results (equations and R^2) of the logarithmic simulations of standard deviation of the score (number of wins divided by the total number of interactions) according to the number of interactions considered to calculate this score, for a set of true dominance scores (see Fig. 1)

			Minimum number of interactions			
Dominance score (%)	Equation	R^2	SD = 0.01	SD = 0.05	SD = 0.1	SD = 0.2
0.1 (or 99.9)	$SD = 0.0118X - 0.3322$	0.7642	2	1	1	1
1 (or 99)	$SD = 0.1009X - 0.5036$	0.9667	99	5	2	1
5 (or 95)	$SD = 0.2250X - 0.5054$	0.9900	474	17	5	2
10 (or 90)	$SD = 0.2937X - 0.4964$	0.9939	906	36	9	3
15 (or 85)	$SD = 0.3509X - 0.4971$	0.9959	1284	51	13	4
20 (or 80)	$SD = 0.4006X - 0.5002$	0.9962	1601	65	17	5
25 (or 75)	$SD = 0.4359X - 0.5011$	0.9972	1869	76	19	5
30 (or 70)	$SD = 0.4574X - 0.4997$	0.9971	2102	84	21	6
35 (or 65)	$SD = 0.4759X - 0.4997$	0.9977	2276	91	23	6
40 (or 60)	$SD = 0.4937X - 0.5012$	0.9978	2393	97	25	7
45 (or 55)	$SD = 0.4993X - 0.5007$	0.9976	2466	100	25	7
50	$SD = 0.5018X - 0.5007$	0.9977	2491	101	26	7

From these equations, we obtain the number of interactions (X) required to obtain a fixed standard deviation ($Y = 0.01$, 0.05, 0.1 or 0.2) corresponding to a given score (in %) ±1, 5, 10 or 20%. As standard deviation largely depends on the dominance score (extreme versus medium), we present results for different ranges of dominance score. In the field, when dominance scores are unknown, we recommend adopting the more pessimistic hypothesis when the subject has a medium score and to choose a minimum number of interactions corresponding to a true dominance of 50%.

(a)	C9	C16	C13	C12	C5	C1	C8	C7	C14	C11	C10	C15	C3	C6	C4	C2
C9	■	5/0	11/1	4/0		4/0	4/0	1/0	7/0	15/0	15/0	12/0	4/0	1/0		16/0
C16		■	12/0	8/0	2/0	4/0		5/3	9/0	49/0	29/0	31/0	16/0	14/0	6/0	26/1
C13			■	30/0	1/0		2/0		11/0	35/0	30/0	15/0	43/0	9/0	2/0	58/0
C12				■		2/0	1/0			5/0	4/0		6/0	9/0	8/0	7/0
C5	3/1				■	5/1		2/0		6/1	5/3	1/1		1/0	1/1	
C1			6/0			■	1/0	1/0		7/2				4/0	3/0	4/2
C8		6/2					■	2/0	1/0	2/0	7/0	3/0	4/0	6/0	3/0	
C7			3/2	34/0				■	4/1	9/1			9/2	4/1	2/2	14/1
C14				1/0					■	4/0	6/2	2/0	2/0	2/0	1/0	1/0
C11										■	50/0	23/9	10/0	17/0	24/0	8/5
C10						2/0		7/5			■	15/5	9/0	12/0	14/1	18/4
C15								7/0				■	5/0	6/0	4/0	3/3
C3													■	5/1	16/1	18/1
C6														■	12/1	13/2
C4	1/0														■	14/0
C2					5/0		3/0									■

(b)	-7	1T	=5	K8	K-	1Y	DT	A9	DD	=F	=H	-3	K3	=J	A-	=K	=T	DK	D2
-7	■	2/1		1/0			2/0							1/0					7/0
1T		■		1/0				5/0	2/0	1/0	1/0		4/0		1/0	1/0	1/0		4/0
=5			■	2/0			1/0	4/0	5/0							1/0	1/0	1/0	2/0
K8				■		1/0				5/0	5/0	1/0	2/0						
K-					■	1/0		1/1	1/1	1/0	1/0		1/0					3/0	
1Y						■				1/0	1/0			1/0				3/0	1/0
DT							■	1/0	2/0										
A9								■		2/0	4/0			5/4	1/0	3/0	3/0	1/0	1/0
DD									■	1/0	1/0					1/0	1/0		1/0
=F										■		5/0	4/0	2/1				9/2	
=H											■	2/0	2/0	2/1				2/2	
-3					2/0							■		2/1	1/0	1/0	1/0	2/0	3/0
K3													■	1/0	2/0	1/0	2/0	1/0	1/0
=J							1/0							■		2/2	2/0		1/0
A-															■	1/0	1/0		
=K										1/0	1/0					■	1/0		
=T																	■	1/0	
DK							2/0											■	
D2																			■

Figure 2. Dominance matrices for (a) mallards ($N = 1160$ interactions) and (b) brent geese ($N = 187$ interactions). The birds are ordered according to dominance rank, decreasing from left to right and from top to bottom. In each filled cell, the number on the left indicates the number of encounters won by the bird in that row over the bird in that column and the number on the right shows the number of encounters won by the 'column' bird over the 'row' bird. As this kind of matrix is symmetrical about the diagonal, only cells in which the 'row' bird won a majority or equal number of the encounters are completed. Pair specific inconsistencies are below the diagonal (underlined); they correspond to pairs of interacting birds for which the direct outcomes are victorious scores opposite to those expected from the overall classification.

Wild Brent Geese

To construct the matrix, we considered only brent geese that had interacted with more than five other marked individuals to avoid a bias of dominance scores caused by a large number of interactions with a few particular opponents. This left 19 marked brent geese with which to construct a dominance matrix. This sociometric matrix comprised 187 interactions. The values of the Kendall's linearity index ($K = 0.22$, $X^2_{26} = 36.9$, $P < 0.05$), the Landau's index and the corrected index were high ($h = 0.22$, $h' = 0.29$, improved linearity test using h', $P < 0.05$) indicating that the ranking was neither circular nor random. Therefore, the analysis of this matrix revealed a significant linear hierarchy and indicated that we could use a linear order to rank individuals, although linearity was less pronounced than in mallards (Fig. 2b).

Matrix ranks (X) and dominance score ranks (Y) were significantly related ($Y = 0.784X + 2.158$; $R^2 = 0.615$, $F_{1,17} = 27.154$, $P < 0.001$; Table 2, Fig. 3b). The slope was not significantly different from 1 ($t_{18} = -1.44$, $P = 0.173$) and the Y-intercept was not significantly different from 0 ($t_{17} = 1.258$, $P = 0.226$). The plot of the residuals of the score−matrix rank regression on dominance scores shows that the rank correspondence was very high for high ranks, but that there were discrepancies among the middle-ranking and subordinate birds (scores below 55%).

DISCUSSION

The results of our modelling approach provide a framework to evaluate the reliability of observed dominance scores by giving an idea of the precision of this measure of individual status and hence a tool to discuss uncertainties in the observed classification. Our results can, therefore, assist in planning the amount of effort necessary to record dominance reliably in both experimental and field studies.

Our model was constructed under several simple hypotheses with the aim of providing a general application to determine the number of interactions needed to obtain a reliable estimation of the score. Therefore, (1) our simulation does not include a specific number of individuals and simulates individuals interacting with an

Table 2. Mallard and brent goose classification by two methods: dominance score and matrix

Code	Number of interactions	Dominance score (%)	Score rank	Matrix rank
Mallards				
C9	105	95	1	1
C16	228	93	2	2
C13	271	88	3	3
C8	47	72	4	7
C7	122	71	5	8
C5	40	60	6	5
C1	48	56	7	6
C11	282	48	8	10
C14	54	37	9	9
C12	119	35	10	4
C10	243	34	11	11
C15	145	28	12	12
C3	152	27	13	13
C6	120	23	14	14
C4	117	18	15	15
C2	227	12	16	16
Brent geese				
-7	424	99	1	1
1T	177	91	2	2
=5	292	89	3	3
K8	122	72	4	4
K-	68	72	5	5
1Y	212	62	6	6
=F	211	55	7	10
=H	134	47	8	11
=K	142	46	9	16
=T	116	42	10	17
DT	114	41	11	7
A9	141	40	12	8
-3	160	35	13	12
DK	259	32	14	18
DD	75	31	15	9
K3	140	24	16	13
D2	218	23	17	19
=J	160	23	18	14
A-	60	15	19	15

Codes represent the individual identities of different birds within their groups.

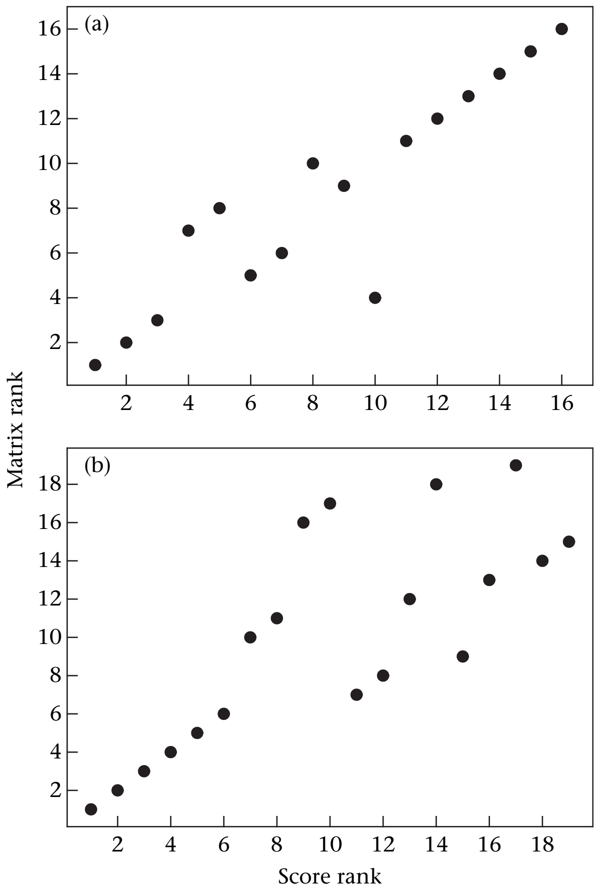

Figure 3. Relations between ranks obtained from two different ranking methods for (a) mallards and (b) brent geese. On the abscissa, birds are ranked according to their relative dominance score (number of wins divided by the total number of interactions). On the ordinate, birds are ranked according to their relative position within the matrix. Birds are classified from most dominant (rank = 1 for both) to most subordinate (rank = 16 for mallards and rank = 19 for brent geese).

infinite number of individuals, (2) we assumed that the individual interacted randomly with different individuals and (3) each new encounter is independent of preceding ones. We fully acknowledge that integrating group structure (Mesterton-Gibbons & Dugatkin 1995; D'Eath & Keeling 2003), individual recognition bias (Grasso et al. 1996; Senar 1999) and the effects of prior winning and losing (Chase et al. 1994; Brotons 1998; Kim & Zuk 2000; Otten et al. 2002; Dugatkin & Earley 2004) would be more realistic. However, with so many possible elaborations existing for some species and not for others, we decided to focus on a set of restrictive assumptions, allowing a wider application, without involving specific bias such as those highlighted above. Indeed, under these constraining hypotheses, increasing the number of recorded interactions should not affect the bias but, rather, increases the precision around the estimated value. In brief, this simple theoretical approach allows the estimation of the precision around an unbiased estimator.

As a result of our model simulations, we selected 26 interactions as the minimum number of interactions for each individual in our studies of captive mallard and wild brent geese. This value was chosen because in the worst-case scenario of a bird with a true dominance of 50% it yields a precision of ±10% that cannot be substantially improved upon without an unrealistic amount of effort. Of course this was a trade-off between the desired precision of the estimator and the feasibility of our studies. Each study will have to draw its own compromise between precision and practicality. The number of individuals studied could also influence the threshold of precision required. As the range of dominance scores is bounded between 0 and 100%, the precision of the estimates required to rank two individuals correctly is lower than the precision of the estimates required to rank 20 or 50 individuals, as the number of possible misclassifications is lower. The type of analyses performed on dominance scores also affects the precision requirements. For instance, Stahl et al. (2001) used only two dominance groups for one analysis, one with animals with a dominance score of 50% or more (referred to as dominant) and one with birds with a dominance score of less than

50% (referred to as subordinate). For this analysis, precision requirements were lower, although according to our simulations about 20% of the individuals could have been wrongly classified in the worst-case scenario.

Dominance Score Validity

We compared dominance score ranks and matrix ranks in two different situations: in captivity where all animals were identified and in the field where only a small proportion of the population was marked. For captive ducks, matrix ranks were precise because of the small number of unknown relationships (eight, corresponding to 6.67% of the possible dyads) and a significant linear hierarchy. We can therefore assume that in this case the matrix ranks reflected the true hierarchy among the birds. Thus, the lack of perfect correlation between dominance score ranks and matrix ranks around the 50% score probably reflects the greater error associated with the estimated dominance of the middle-ranked birds. For brent geese, the matrix determination of the hierarchy order was more difficult. Because there were few ringed birds, with few interactions between them, there were many unknown relationships in the matrix (84, corresponding to 49.12% of the possible dyads). This translated into a significant but weak linear hierarchy. As a consequence, the relation between dominance score ranks and matrix ranks was poorer than in ducks, although still significant, as errors in the matrix ranks combined with those associated with the dominance score ranks. In fact, as the number of different opponents per marked individual was low in the matrix (between six and 16), it is possible that the matrix does not represent the true social context in which birds live. In this typical field study case, with a lot of interactions per marked individual but only a few between marked birds, the simple dominance score gives a fairly precise ranking and is thus probably better to use, and easier, than a very incomplete matrix.

In this paper, we present the first calibration of the dominance score method against the dominance matrix, which is usually considered more reliable. Where a well-estimated matrix was available, simple dominance score ranks proved to be highly correlated with the ranks derived from the matrix. Thus, given a sufficient number of interactions, simple dominance scores may provide a very reliable indicator of dominance status. Indeed, in field studies where attaining a complete matrix is often difficult, simple dominance scores of marked individuals may provide a rank estimate with greater precision than that derived from the few interactions that occur between them. We call for caution, however, when using dominance scores based on few interactions to rank individuals. Researchers must consider the error associated with their estimated dominance scores and the way in which this changes with sample size in planning their behavioural sampling protocols.

As expected, we found that an increased sample size increases the precision of the estimator, but the magnitude of the improvement is unequal between scores, and hence between ranks. The intermediate scores are the least precise; hence the intermediate-ranking individuals are those most likely to be misclassified. There is also a biological reason for this number, not just a statistical one. Middle-ranked individuals probably have similar fighting capacities and from a biological point of view it is difficult to discern the hierarchy between these individuals. Indeed, the uncertainty over positions in the middle of the hierarchy has implications for the probability of escalation of conflicts. The problem that lies with the middle ranked individuals has key implications for behavioural ecology beyond the scope of this paper.

The required precision increases with the number of individuals monitored, but the number of individuals must be sufficient to be representative of the population studied. Field work effort is therefore likely to be a compromise, dictated by time and money. The table we provide in this paper can help in setting threshold sample sizes on which to base field studies for a desired precision, and consequently ease planning decisions, at least in situations when group size is large enough partly to satisfy the initial assumptions. In addition, this table may also be used to define the range of possible error associated with ranks attributed from dominance score analysis.

Acknowledgments

We are grateful to David Pinaud for statistical help. We also highly appreciate the skilful assistance with duck and brent goose manipulations and field work of Nadine and Noël Guillon, Matthieu Guillemain, Pierre-Yves Perroi, Sébastien Dalloyau, Léopold Denonfoux, Sandra Blais, Daphné Durant and particularly Barwolt S. Ebbinge, Gerard Müskens and Janien van der Greft. The manuscript benefited greatly from critical comments by Richard Caldow, Matthieu Guillemain, David Carslake, Marcel M. Lambrechts, Graeme Ruxton, Rob Slotow and two anonymous referees. The programme was funded by the CNRS, the Office National de la Chasse et de la Faune Sauvage and the EU ‘Coast Bird Diversity’ programme EVK2-2000-00612 for the brent geese component.

References

Arcese, P. & Smith, J. N. M. 1985. Phenotypic correlates and ecological consequences of dominance in song sparrows. *Journal of Animal Ecology*, **54**, 817–830.

Archer, J. 1988. *The Behavioural Biology of Aggression.* Cambridge: Cambridge University Press.

Belthoff, J. R., Dufty, A. M., Jr. & Gauthreaux, S. A., Jr. 1994. Plumage variation, plasma steroids and social dominance in male house finches. *Condor*, **96**, 614–625.

Bernstein, I. S. 1981. Dominance: the baby and the bathwater. *Behavioral and Brain Sciences*, **4**, 419–458.

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

Brotons, L. 1998. Status signalling in the coal tit (*Parus ater*): the role of previous knowledge of individuals. *Etologia*, **6**, 49–52.

Caldow, R. W. G. & Goss-Custard, J. D. 1996. Temporal variation in the social rank of adult oystercatchers *Haematopus ostralegus*. *Ardea*, **84**, 389–399.

Chase, I. D., Bartolomeo, C. & Dugatkin, L. A. 1994. Aggressive interactions and inter-contest interval: how long do winners keep winning? *Animal Behaviour*, **48**, 393–400.

D'Eath, R. B. & Keeling, L. J. 2003. Social discrimination and aggression by laying hens in large groups: from peck orders to social tolerance. *Applied Animal Behaviour Science*, **84**, 197–212.

Drews, C. 1993. The concept and definition of dominance in animal behaviour. *Behaviour*, **125**, 285–313.

Dugatkin, L. A. & Earley, R. L. 2004. Individual recognition, dominance hierarchies and winner and loser effects. *Proceedings of the Royal Society of London, Series B*, **271**, 1537–1540.

Ens, B. J. & Goss-Custard, J. D. 1984. Interference among oystercatchers, *Haematopus ostralegus*, feeding on mussels, *Mytilus edulis*, on the Exe Estuary. *Journal of Animal Ecology*, **53**, 217–231.

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.

Gauthreaux, S. A., Jr. 1978. The ecological significance of behavioral dominance. *Perspectives in Ethology*, **3**, 17–54.

Goss-Custard, J. D., Caldow, R. W. G., Clarke, R. T., Le V dit Durell, S. E. A. & Sutherland, W. J. 1995a. Deriving population parameters from individual variations in foraging behaviour. I. Empirical game theory distribution model of oystercatchers *Haematopus ostralegus* feeding on mussels *Mytilus edulis*. *Journal of Animal Ecology*, **64**, 265–276.

Goss-Custard, J. D., Caldow, R. W. G., Clarke, R. T. & West, A. D. 1995b. Deriving population parameters from individual variations in foraging behaviour. II. Model tests and population parameters. *Journal of Animal Ecology*, **64**, 277–289.

Grasso, M. J., Savalli, U. D. & Mumme, R. L. 1996. Status signalling in dark-eyed juncos: perceived status of other birds affects dominance interactions. *Condor*, **98**, 636–639.

Guillemin, M.-L., Atramentowicz, M. & Charles-Dominique, P. 2000. Dominance relationships in captive male bare-tailed woolly opossum (*Caluromys philander*, Marsupialia: Didelphidae). *Revue d'Ecologie (Terre et Vie)*, **55**, 337–349.

Huntingford, F. & Turner, A. 1987. *Animal Conflict*. London: Chapman & Hall.

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.

Kikkawa, J. 1980. Winter survival in relation to dominance classes among silvereyes *Zosterops lateralis chlorocephala* of Heron Island, Great Barrier Reef. *Ibis*, **122**, 437–446.

Kim, T. & Zuk, M. 2000. The effects of age and previous experience on social rank in female red junglefowl, *Gallus gallus spadiceus*. *Animal Behaviour*, **60**, 239–244.

Koivula, K. 1999. Ecological consequences of social dominance in birds. In: *Proceedings of the 22nd International Ornithological Congress* (Ed. by N. J. Adams & R. H. Slotow), pp. 1580–1591. Johannesburg: Bird Life South Africa.

Lamprecht, J. 1986. Structure and causation of the dominance hierarchy in a flock of bar-headed geese (*Anser indicus*). *Behaviour*, **96**, 28–48.

Macdonald, D. W., Stewart, P. D., Johnson, P. J., Porkert, J. & Buesching, C. 2002. No evidence of social hierarchy amongst feeding badgers, *Meles meles*. *Ethology*, **108**, 613–628.

Mesterton-Gibbons, M. & Dugatkin, L. A. 1995. Toward a theory of dominance hierarchies: effects of assessment, group size and variation in fighting ability. *Behavioral Ecology*, **6**, 416–423.

Otten, W., Puppe, B., Kanitz, E., Schön, P. C. & Stabenow, B. 2002. Physiological and behavioral effects of different success during social confrontation in pigs with prior dominance experience. *Physiology and Behavior*, **75**, 127–133.

Piper, W. H. 1997. Social dominance in birds. Early findings and new horizons. In: *Current Ornithology* (Ed. by V. Nolan Jr, E. D. Ketterson & C. F. Thompson), pp. 125–187. New York: Plenum.

Robinson, S. K. 1986. Benefits, costs and determinants of dominance in a polygynous oriole. *Animal Behaviour*, **34**, 241–255.

Rowell, T. 1974. The concept of social dominance. *Behavioral Biology*, **11**, 131–154.

Senar, J. C. 1999. Plumage colouration as a signal of social status. In: *Proceedings of the 22nd International. Ornithological. Congress, Durban* (Ed. by N. J. Adams & R. H. Slotow), pp. 1669–1686. Johannesburg: BirdLife South Africa.

Stahl, J., Tolsma, P. H., Loonen, M. J. J. E. & Drent, R. H. 2001. Subordinates explore but dominants profit: resource competition in high Arctic barnacle goose flocks. *Animal Behaviour*, **61**, 257–264.

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

de Vries, H., Netto, W. J. & Hanegraaf, P. L. H. 1993. MatMan: a program for the analysis of sociometric matrices and behavioural transition matrices. *Behaviour*, **125**, 157–175.

Wilkinson, L. 1997. *Systat 7.0, New Statistics*. Chicago: SPSS Inc.

Zahavi, A. 1989. Arabian babbler. In: *Lifetime Reproduction in Birds* (Ed. by I. Newton), pp. 253–275. London: Academic Press.