



## ARTICLES

# Finding an appropriate order for a hierarchy based on probabilistic dominance

KIMBERLY A. JAMESON\*, MICHAEL C. APPLEBY† & LINTON C. FREEMAN‡

\*University of California, San Diego

†University of Edinburgh

‡University of California, Irvine

(Received 3 July 1998; initial acceptance 18 July 1998;  
final acceptance 21 November 1998; MS. number: A8096R)

Methods of ranking individuals in a dominance hierarchy that use transitivity of relationships may obscure irregularities. Furthermore, these methods use only a small proportion of the information available from dominance encounters. This paper presents an intuitively appealing and easily implemented alternative to existing methods for ordering dominance data, developed from the work of Batchelder et al. (1992, *Journal of Mathematical Psychology*, **36**, 185–212). The procedure presented here is based on a mathematical model of paired comparisons and it involves only simple estimation procedures. We illustrate its use with data on dominance among red deer, *Cervus elaphus*, stags. The results indicate that dominance relationships are well characterized by the scale values that the model provides, and, because the method provides predictions for all pairings of animals, dominance predictions also exist for pairs of animals that have yet to be observed. Moreover, the dominance outcomes predicted by the model using the order scale are highly correlated with actual dominance observations at all levels. Overall, the procedure described provides a solution to the problem of identifying an appropriate order for a near-linear dominance hierarchy.

© 1999 The Association for the Study of Animal Behaviour

Dominance hierarchies in groups of animals are widely studied, yet methods for the description of hierarchies are still under discussion. In particular, many hierarchies appear to be linear or near-linear. In a linear hierarchy, one individual dominates all the others, a second dominates all but the first and so on. Linear hierarchies are rare except in small groups of animals (Wood-Gush 1955; Drews 1993), and in near-linear hierarchies the question arises of how best to order the group members. de Vries (1998) has recently reviewed methods that have been used for such ranking. The majority are based not on dominance interactions or encounters, but on dominance relationships within pairs of animals (dyads). In

many dyads, one member wins most or all interactions: indeed, that is the definition of dominance (Drews 1993). A matrix of encounters can therefore be simplified into a matrix of relationships with each entry being either 1 or 0: usually a 1 indicates that the 'row individual' is dominant to the 'column individual', a 0 that it is subordinate (missing and tied relationships are also possible). The criterion of dominance used for this purpose may be simply that one individual has won the majority of interactions, or that the proportion of interactions it has won is significantly greater than half by binomial test. de Vries (1998) argues in favour of methods using 1/0 dominance relationships, principally because alternatives may assume that individual dominance encounters are statistically independent, an assumption that is unlikely to hold. This point will be discussed below.

A common approach to ranking using relationships is to arrange the individuals in an order that minimizes the instances of individuals dominating others higher in rank (Schein & Fohrman 1955; Brown 1975). This approach assumes that dominance is transitive: thus, among animals A, B and C (a triad), if A dominates B and B

Correspondence: K. A. Jameson, Department of Psychology, University of California at San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0109, U.S.A. ([kjameson@ucsd.edu](mailto:kjameson@ucsd.edu)). L. C. Freeman is at the Institute for Mathematical Behavioural Sciences, School of Social Sciences SSPA 2143, University of California, Irvine, CA 92697-5100, U.S.A. M. C. Appleby is at the University of Edinburgh, School of Agriculture Building, West Mains Road, Edinburgh EH9 3JG, Scotland.

dominates C, then A will also dominate C. Before using such a method, it is therefore appropriate to test statistically whether dominance is more transitive than chance (Appleby 1983). de Vries (1998) describes a number of methods that formalize Brown's (1975) approach, and presents a new method of his own.

Two problems remain with such methods. First, as pointed out by Beilharz & Mylrea (1963) and emphasized by Appleby (1983), 'tidying up' hierarchies in this way may be misleading. Just because transitivity holds statistically does not mean that every triad not known to be intransitive can be assumed to be transitive.

For example, a weak dependence of dominance on the absolute characteristics of the individuals involved would be sufficient to produce significant linearity in a large group, and assuming dominance to be transitive throughout could obscure various irregularities in the hierarchy (Appleby 1983, pp. 605–606).

The main reason for ranking group members must be to consider the correlates of rank, including causes and effects. If there is no intrinsic reason for using transitivity in ranking, then it may be inappropriate. Alternative methods of ranking using relationships that do not involve transitivity of dominance include simple indices such as the proportion of opponents encountered that are beaten (Wagnon et al. 1966) and more complex indices that take into account success of opponents (Wei 1952; Kendall 1955; Clutton-Brock et al. 1979).

Second, information on dominance in groups of animals is often incomplete. This problem can be made worse if only some of the information collected is used, as occurs when multiple encounters in a dyad are represented simply in a 1/0 form. For example, Appleby (1983) found that the pattern of relationships among seven red deer, *Cervus elaphus*, stags could not be distinguished from random, suggesting that no hierarchy was present. Iverson & Sade (1990), however, showed that when the individual interactions were taken into account, the pattern of wins and losses was clearly non-random. Freeman et al. (1992) further explored analyses of encounters and asked not only 'who beats whom?' but also 'who fights with whom?' and 'who fights at all?' (page 239), although the detail of some of their conclusions was later called into question (Appleby 1993).

Both these problems are addressed by a method offered by Boyd & Silk (1983). Their index of dominance does not assume transitivity, and uses all the information available from individual dominance encounters rather than just a simplified matrix of relationships. In addition, their method has the attractive feature of an iterative procedure that uses the success of opponents to yield rankings even in cases where many dyads in the group have never been seen to interact. Unfortunately the method has a number of restrictive requirements and assumptions, in addition to the assumption that individual dominance encounters are statistically independent. de Vries (1998) recognized both the procedure's disadvantages and advantages in describing it as 'the method that is based on the most restrictive model, but that also yields the most informative results' (page 833).

Boyd & Silk (1983) assume that dominance is not deterministic. They see it as essentially stochastic, such that depending on fluctuations in health, fatigue and the like, either of a given pair of animals might win any particular encounter. In the present paper we also take the probabilistic approach, and introduce an alternative probability-based method that is more appropriate for application to dominance interactions than that proposed by Boyd & Silk (1983). In common with all the methods discussed above, the method presented here is appropriate for ranking near-linear hierarchies. It has yet to be proven appropriate for more complex social structures, such as those in mixed groups of male and female animals with partially independent hierarchies, or groups with frequent alliances or maternally inherited dominance. Such complexities are particularly common in primates (Sade 1992). However, in studies of complex human behaviours like chess playing (Elo 1978), our method is ideal as a procedure for assigning performance rankings of individual chess players involved in tournament play. Thus it is plausible that our method may be amenable to more complex structures of animal dominance, but this question is not addressed in the present paper. However, for near-linear hierarchies, this paper provides an intuitively appealing and easily implemented alternative to existing methods for ordering dominance data. We illustrate the method by applying it to data from observations of 68 red deer stags on the Isle of Rhum (Appleby 1982).

## THE BOYD & SILK MODEL

Both the model proposed by Boyd & Silk (1983) and the one that we introduce here are based on Thurstone's (1927) method of paired comparisons. Thurstone's method involves observing a series of events in which pairs of objects are formed in such a way that one object in each pair is selected. The win/loss matrix defined above is generated by just such observations. The events that the ethologist observes are agonistic encounters; the object selected is the winner.

The model described by Boyd & Silk was originally introduced by Bradley & Terry (1952). Consider a particular pair of animals  $a_i$  and  $a_j$ . Let  $\pi_{ij}$  be the probability that in any encounter  $a_i$  dominates  $a_j$ . Then the Bradley–Terry procedure uses probabilities to assign a dominance scale value  $s(a_i)$  to each animal  $a_i$  in the community. The procedure assigns dominance scale values to  $a_i$  and  $a_j$  in such a way that the probability  $\pi_{ij}$  is a monotone increasing function  $F$  of the difference between their scale values:

$$\pi_{ij} = F[s(a_i) - s(a_j)]$$

The values of  $\pi_{ij}$  are assumed to be continuous and are estimated from the observed proportions  $p_{ij}$  defined above. The magnitudes of the  $p_{ij}$ s are used to estimate values of  $s(a_i)$  by using a maximum likelihood procedure.

The model itself requires three assumptions: (1) the underlying distribution of dominance is continuous; (2) the probability that a particular animal will defeat a

particular other,  $\pi_{ij}$ , is constant; and (3) the outcome of a particular encounter between any pair of animals is independent of the outcomes of their previous encounters. These assumptions may be questionable in behavioural terms, but they are essential in any stochastic approach. As mentioned above, the third assumption has been particularly questioned (de Vries 1998). However, it should be noted that neither the method described by Boyd & Silk nor the alternative that we suggest below are statistical tests, requiring independent data. Although the theory that produced the algorithm makes assumptions of independence of individual dominance encounters, the results of the algorithm can be tested directly without that assumption, as explained below (also see Discussion).

Greater difficulties are raised by three additional assumptions that must be met in order to estimate the parameters of the Bradley–Terry model. First, estimating values of  $s(a_i)$  requires the use of an iterative computational procedure to solve a set of nonlinear equations. Unfortunately, the procedure converges on a solution only when the data meet a rather stringent and, for this application, probably inappropriate, additional condition that there should be no extreme individuals. As Boyd & Silk (1983, page 49) describe it, ‘... convergence is not achieved in groups that include individuals who are never dominated by other individuals ...’, but we can also add that convergence is not achieved when an individual is dominated by all the others. In a linear order, however, there is a top-ranking animal that is never defeated and a bottom-ranking animal that is always defeated, and because we are seeking an approximation of such an order, the Bradley–Terry model is of questionable utility for this application.

Second, an estimation of  $s(a_i)$  entails the assumption of constant interaction rates. Boyd & Silk describe it as assuming ‘... that the rate of interaction between individuals is independent of their relative dominance ranks ...’, but it has been established, at least for some species (Cole 1981; Clutton-Brock et al. 1982; Freeman et al. 1992), that interaction rates are strongly associated with relative dominance positions of the individuals involved. So, again, we see that the Bradley–Terry model is not really appropriate for applications to dominance.

The third and perhaps most telling deficiency of the Bradley–Terry model in this context is that to get stable estimates of its parameters, the data must include repeated observations on each pair of individuals. Thus, the data must provide information on a kind of extended round-robin series of encounters (Roberts 1990). However, ethologists cannot usually produce this kind of information. In natural settings, it is almost universally true that a good many pairs of individuals are simply never observed interacting. Although the Bradley–Terry model has many strengths, overall, it is probably not the model we want for applications to dominance.

#### THE BATCHELDER–BERSHAD–SIMPSON (BBS) SCALING METHOD

Batchelder & Bershad (1979) introduced an alternative model also based on Thurstone’s paired comparisons.

Their model was later elaborated by Batchelder & Simpson (1989) and Batchelder et al. (1992). The Batchelder–Bershad–Simpson (BBS) method is a formal statement of an idea that was originally developed informally by Elo (1978), a statistician and chess enthusiast. Its original aim was to provide a systematic procedure for scaling the performance of a relatively large number of chess players on the basis of outcomes from a relatively small number of pairwise games. Because the data from animal communities that concern ethologists also tend to display this kind of sparseness, the BBS model has some special properties that make it appealing for application to dominance data.

The BBS model shares the same assumptions about continuity, stability and independence of outcomes with the model discussed above, but, in addition, Batchelder & Bershad add a fourth assumption that permits each animal to be assigned a scale score, even under conditions like those we are faced with here, where the entries in the dominance matrix are relatively sparse. Their additional assumption is that the distribution of dominance that underlies the observed behaviour is normal, an assumption which is discussed below.

The BBS model uses a simple equation to provide initial estimates of scale values  $s(a_i)$  for each animal:

$$s(a_i) = [\alpha(2W_i - N_i)/2N_i] \quad (1)$$

where  $\alpha = \sqrt{2\pi} = 2.50663$  is a constant, taken from the Taylor expansion of the normal distribution;  $W_i$  is the number of encounters in which animal  $a_i$  was observed to have won; and  $N_i$  is the number of encounters in which  $a_i$  was involved. Using this equation, each animal is assigned an initial scale score as a simple function of the proportion of agonistic encounters it won. Then, given the initial scale scores provided by this first equation, a second equation is used recursively to rescale the animals repeatedly until their scale scores become invariant:

$$s(a_i) = [2(W_i - L_i)/N_i] + Q_i \quad (2)$$

where  $L_i$  is the number of encounters in which animal  $a_i$  lost; and  $Q_i$  is the mean scale score of those animals that  $a_i$  met in agonistic encounters.

Thus, an animal’s scale position depends on (1) the proportion of wins in its encounters with others, (2) its proportion of losses, and (3) the scale scores of the others that it has met in agonistic encounters. This measure is similar to the one proposed by Clutton-Brock et al. (1982), which is based on the numbers of opponents beaten and lost to, and the success of those opponents. The measure proposed here takes similar features of the data into account.

This equation is applied repeatedly to observed data. Initial values of  $Q_i$  are calculated from the initial scale scores. Then, as a result of that first calculation, each animal is assigned a new scale score that depends, in part, on the scale scores of its opponents. These new scores are then used to calculate new values of  $Q_i$  for each animal. Using successive score updates, this process is repeated until the scores converge and continuing calculations no

**Table 1.** A portion of the Appleby (1983) data consisting of the raw dominance encounter matrix for 68 red deer stags

	BUDY	CLEC	MAXI	SBBW	HAMI	SOOC	RGRC	YHOC	PETE	SLIP	.	.	.
BUDY	0	0	0	0	4	0	0	0	0	0	.	.	.
CLEC	0	0	0	0	0	1	0	0	0	0	.	.	.
MAXI	3	1	0	7	4	2	3	0	0	6	.	.	.
SBBW	0	11	0	0	0	3	4	0	0	0	.	.	.
HAMI	0	5	0	13	0	7	3	5	0	1	.	.	.
SOOC	0	4	0	0	0	0	1	0	0	0	.	.	.
RGRC	0	2	0	0	0	1	0	0	0	0	.	.	.
YHOC	0	0	0	8	0	3	5	0	0	0	.	.	.
PETE	8	0	0	1	1	0	3	0	0	1	.	.	.
SLIP	5	3	0	3	7	3	1	3	0	0	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.

longer induce changes. These final scale scores can then be used to order the animals. The scale scores for a given pair of animals  $a_i$  and  $a_j$  are  $s(a_i)$  and  $s(a_j)$ . If  $s(a_i) > s(a_j)$ , then the model concludes that  $a_i$  is dominant to  $a_j$ . And if  $s(a_i) < s(a_j)$  then  $a_j$  is judged to be dominant.

**Predicting Dominance Outcomes from a Derived Dominance Ordering Scale**

An additional important feature of the BBS method rests in its ability to predict unobserved dominance patterns using the formal model and the obtained dominance scale. This feature is valuable, has been found to be robust under empirical test, and is not a feature that is presented by the existing methods in the literature.

Briefly, if  $A = \{a_1, a_2, \dots, a_n\}$  denotes the set of  $n$  animals; and  $S = \{s(a_1), s(a_2), \dots, s(a_n)\}$  denotes the convergent dominance scale, then the model's predictions are obtained by employing each scale value  $s(a_i)$  in a piecewise linear equation given by the model to predict outcome probabilities for all possible encounters between scaled animals (see Jameson 1996 for details). The essential feature of this Thurstonian-based procedure is that the model's predictions are a monotonic function of the difference between the true dominance rankings of the animals being considered. This procedure derives the complete  $n \times n$  matrix of conditional probabilities for all pairings of animals, including estimating outcomes for agonistic encounters between animals that have never previously met.

**An Application to Dominance Data**

To show how the BBS model works, we used Appleby's (1982) data from a field study of 68 red deer stags during a 3.5-month period, from January to April 1978 on the Isle of Rhum. This data set records all of the encounters among the  $(68)(67)/2 = 2278$  pairs of stags that were observed during this period (Table 1). Of these 2278 pairs, 1632, or 72%, were never observed together in any encounter. Thus, the dominance matrix for these red deer stags is, as is typically the case, relatively sparse.

The data exhibit many of the problems typically found in studies of dominance. There were 24 pairs of stags that did not display a completely stable pattern of dominance. When confrontations occurred between the two animals in these pairs, sometimes one of them won and sometimes the other; in four of these cases, each animal in the pair won exactly half of their encounters with each other. There were 12 triads of stags that displayed cyclical patterns, where one stag dominated a second, the second dominated a third and the third dominated the first. Stags varied widely in their aggressiveness, from individuals that initiated over 100 interactions to others that initiated none, and they distributed their aggression unevenly, attacking some opponents more than others (cf. Freeman et al. 1992).

Initial scale values were assigned to individual animals according to equation (1) above (Table 2).

We then used the number of encounters that each stag won and lost, and the initial scale scores of the

**Table 2.** Initial estimated dominance scale ordering for 68 red deer stags

FERD	1.2533	TRCC	0.2121	MYRC	-0.5371
CORK	1.2533	BLTC	0.2040	HECT	-0.5371
ORPH	1.2533	YHOC	0.1119	COC4	-0.5396
BL45	1.2533	FROD	0.0964	RGCC	-0.5483
PETE	1.2533	SOOC	0.0869	SPRI	-0.6133
MAXI	1.2533	CRTC	0.0836	UPCC	-0.6267
JUNC	1.0743	CR14	0.0783	TR34	-0.6267
LICC	0.8355	BDLT	0.0000	TAL5	-0.6650
BREA	0.8322	STUL	0.0000	ELSI	-0.6963
BLT2	0.7976	CLEC	-0.0404	BST4	-0.7110
SLIP	0.7365	SOLO	-0.0418	RG14	-0.7127
CONA	0.7050	SBBW	-0.0508	COCC	-0.7520
HAMI	0.6611	FEAT	-0.0836	BROC	-0.9984
SX95	0.6483	UPT3	-0.0917	GRE4	-1.1059
BOSS	0.5994	REDC	-0.1139	BST5	-1.1320
FECC	0.5106	RGRC	-0.1253	CHE5	-1.1725
CLYD	0.4896	LTSP	-0.2027	M4	-1.2533
TORM	0.4700	TALC	-0.2243	DOR5	-1.2533
DICK	0.4424	CHOC	-0.2938	RECC	-1.2533
YEST	0.4178	TA24	-0.3374	RGR4	-1.2533
CLCC	0.3531	GIPS	-0.3418	ALTO	-1.2533
BUDY	0.2984	SC01	-0.3581	SC05	-1.2533
GILL	0.2350	WHIC	-0.4178		

**Table 3.** Dominance scale ordering of the final scale scores convergent at iteration 17, for 68 red deer stags

PETE	3.0073	YHOC	0.3218	SPRI	-1.1691
MAXI	2.3447	GIPS	0.2379	ELSI	-1.2256
BREA	2.0432	SBBW	0.2376	COCC	-1.2651
TORM	2.0244	BDLT	0.2073	RGCC	-1.3209
BLT2	1.9716	BLTC	0.1388	BROC	-1.3432
CONA	1.9591	FROD	0.1352	TR34	-1.3584
JUNC	1.8876	CRTC	0.1253	TA24	-1.3591
SLIP	1.7518	RGRC	0.1042	UPCC	-1.4714
CORK	1.7356	TALC	0.0687	COC4	-1.5565
SX95	1.7292	LICC	0.0676	BST4	-1.7059
ORPH	1.6363	SOOC	-0.0114	CR14	-1.7671
HAMI	1.3478	TRCC	-0.1676	RECC	-2.0114
BOSS	1.2056	CLEC	-0.1778	RG14	-2.1087
YEST	1.0609	HECT	-0.2142	M4	-2.1676
CLYD	0.9318	BL45	-0.2521	TAL5	-2.2521
BUDY	0.8802	FECC	-0.2640	GRE4	-2.3151
LTSP	0.7758	CLCC	-0.3056	RGR4	-2.4543
GILL	0.7316	FEAT	-0.3953	DORS	-2.4929
DICK	0.7042	CHOC	-0.4213	ALTO	-2.5023
FERD	0.6239	REDC	-0.4707	CHE5	-3.0757
UPT3	0.4474	SC01	-0.4929	BST5	-3.0960
STUL	0.4284	WHIC	-0.5259	SC05	-3.3591
SOLO	0.3225	MYRC	-0.7918		

individuals he had encountered to estimate new scale scores. We repeated this process until the resulting scale converged on the 17th recursion (Table 3).

The resulting scale, in which no two stags had the same score, allowed us to rank all 68 individuals. In general, scale scores need not be uniquely assigned to animals. For example, if two animals are observed with identical interaction patterns (i.e. encountering exactly the same animals, and sharing identical patterns of wins and losses), both will earn the same scale score.

Note that the range of initial scores presented in Table 2 is -1.25-1.25, whereas that presented in Table 3 is -3.35-3.0. The difference between these ranges reflects the influence of scaling factor  $Q_i$  in equation (2), which incorporates scale information about the ratings of a given animal's opponents into Table 3. In contrast, the scale in Table 2, derived using equation (1), simply reflects a scaling factor equal to zero, an initial consequence of the normal distribution assumption.

The model uses the overall pattern of wins and losses to generate an order for all the pairs of animals, those that were observed and those that were not. We can see how well this order captures the observed pairwise data by using the model results to estimate which animal in each pair was dominant and compare this with actual observations. We can then calculate a proportional reduction of error measure to determine whether the results of the model are better than chance.

The predictions made by this model are obtained by using each animal's scale value in a piecewise linear equation that produces predicted probabilities for all possible encounters between the animals scaled (Jameson 1996). Because most of the pairs studied here were never observed together, we cannot evaluate the performance of most of these predictions. We can, however, evaluate those predictions that were generated for all the pairs in which the two animals were actually observed together.

**Table 4.** Comparison between dominance relationships predicted by the BBS model and those observed

Probability range excluded	Pairs excluded	Pairs retained	Kendall's tau	
			100% Data	90% Data
None	0	642	0.853	0.847
0.45-0.55	56	586	0.925	0.912
0.40-0.60	111	531	0.958	0.941
0.30-0.70	195	447	0.991	0.982

Values of Kendall's tau are presented both for the complete data set (100% Data) and with 10% of the predictors removed from the computations (90% Data). Successive rows were calculated as pairs of animals with relatively poor predictability were excluded.

Because the predictions of the model are based only on each individual animal's overall pattern of wins and losses and on the average win/loss pattern of those he encountered, this is a rigorous test.

A corollary of the BBS model suggests that animals whose dominance scale scores are close together will give rise to more error in the estimation process. As a consequence of this consideration, we calculated several indices of agreement (Table 4). One index was based on all the pairs for which an order was observed. One index excluded 56 pairs that yielded predictions near 50/50, that is, those falling in the range from 0.45 to 0.55. Another index further excluded 55 pairs that fell in the range from 0.40 to 0.60, and a third excluded another 84 pairs that fell between 0.30 and 0.70.

The penultimate column of Table 4 shows that the dominance outcomes predicted by the model using the order scale were highly correlated with actual dominance observations at all levels. Not too surprisingly, the results also show that the model was more effective as a predictor for those pairs of animals with large differences in scale values. Thus, the assumption of the Thurstonian model is correct, in that the error in the estimation process is systematically distributed in a way that is consistent with differences in dominance estimates. In contrast to a uniformly distributed estimation error, the model's estimation error decreases as the difference between two dominance scale estimates increases. This is what one would expect to observe if the Thurstone model were appropriate for dominance data.

Table 4 demonstrates that actual dominance interactions can be predicted by the scale values the model provides, and, because the model provides predictions for all pairings of animals, dominance predictions also exist for pairings of animals that have yet to be observed. We can get a sense of this potential by eliminating some of our observations and repeating the scaling process. We randomly selected 65 (10%) pairs among those for which a dominance order had been observed, and dropped those observations. We substituted zeros in their matrix cells as if no encounter had been recorded. Then we repeated our calculations to produce a new initial scaling (like Table 2), and a new final order (like Table 3), and a correlation table (final column of Table 4). The final correlations demonstrate

that there was very little loss of predictive power even with the loss of input data.

So, even when the BBS model is confronted with sparse data and circular triads, it can obtain an order with very little distortion of the observed patterning of pairwise dominance. This order can then be used to predict dominance interaction outcomes for any possible encounter between scaled animals.

## DISCUSSION

The BBS model provides a workable approach to the problem of identifying an appropriate order from dominance data that are not quite linear. It uses only minimal information about individual animals (their overall proportions of winning and losing encounters and an index of the dominance level of their opponents) to estimate the overall dominance patterning of all the pairs, including those pairs not observed together in any encounter. Perhaps the most important feature of this model is that its estimates are based on assumptions not automatically violated when applied to dominance data. While the Bradley–Terry model can provide estimates only by assuming that there are no extreme individuals and that there are constant interaction rates, the present model involves no such restrictions. Because the computational algorithm we used assumes that dominance is normally distributed, the present model requires only simple and direct estimation procedures that entail no additional counterintuitive assumptions. Two assumptions made by the BBS model need discussion. The first assumption is that the outcome of a given encounter between two animals is probabilistically independent of the outcomes of previous encounters. This assumption, which is common to both the Thurstone Case V model (Thurstone 1927) underlying the BBS algorithm and to the Bradley–Terry model (Bradley & Terry 1952) employed by Boyd & Silk (1983), gives ease of derivation of the underlying mathematical form and permits a characterization of dominance interactions using just a few parameters. Without this assumption, these relations involve more parameters than data points, making a solution indeterminate. As mentioned above, the assumption of the independence of individual dominance encounters is a theoretical assumption that allows for one method of justifying the algorithm. The procedure and results of the algorithm and the methods of testing the dominance rankings do not assume independence of dominance encounters. Thus, the assumption of independence is not tantamount to acceptance of an unrealistic expectation for behaviour, rather it is a simplification made at the level of the underlying model which imposes no serious constraints or expectations at the level of the observed behaviour.

However, we may also note that the outcomes of encounters observed in a study are not wholly dependent either. This is because observations record only a short period in the longer ongoing life of the group, except in studies of newly formed groups, and thus, the encounters observed will be only a subset of a longer sequence. So the probability of  $a_i$  beating  $a_j$  in the second observed

encounter may not be completely independent of the outcome of the first, but the influence of the first encounter is diluted by the influence of all previous unobserved encounters. Indeed, in many dyads, the probability will in fact be nearly 1 or 0 and will not change over time.

In addition, results from psychological research suggest that the BBS method is likely to be of use in studies of newly mixed groups, in which the encounters observed really are the first, second and so on. The special case of hierarchy formation, as opposed to hierarchy maintenance, is addressed elsewhere (Chase 1974; Chase et al. 1994), and future applications of the BBS method will clarify the utility of the procedure in the case of hierarchy formation.

The second assumption is that the distribution underlying dominance interactions is normal. This is the only new assumption made by the BBS method. Indeed, it can be questioned whether the notion of a distribution underlying dominance interactions is a valid construct. On this point Iverson (1987) demonstrates, in a direct generalization of the model, that the binary choice probability of animal  $x$  dominating animal  $y$  (denoted  $P_{x,y}$ ) is a function of a random variable representation of the underlying (dominance) process, denoted  $U_x$  and  $U_y$ , respectively. Thus, the probability that  $x$  will dominate  $y$  is:

$$P_{x,y} = P(U_x > U_y)$$

According to Iverson (1987, page 221) ‘the assumption that the random variables  $U_x$  [and  $U_y$ ] are normal appears quite arbitrary . . . Thurstone himself was well aware of the apparent arbitrariness involved in making such a strong distributional assumption . . . [However,] suppose the normality assumption is replaced by some other distributional hypothesis. What happens typically is that no simplification of the  $[P_{x,y}]$  equation ensues, even if the random variables  $U_x$ ,  $U_y$  underlying a comparison of  $x$  and  $y$  are assumed to be independent.’ Iverson concludes that the normality assumption is not critical to this general class of Thurstonian models. Furthermore, as it is used here, the normal assumption has been shown to provide a reasonable approximation of a number of other possible underlying distributions (Yellott 1977).

Intuitively it also makes sense that the random variables  $U_x$ ,  $U_y$  are approximately normally distributed for the case of animal dominance. In any reasonably large group of animals, factors that affect dominance, such as competitive ability, will display a range of variation, and these features of individual animals are typically normally distributed. Clearly, if dominance were deterministically given by the distribution of size in the herd, then a given animal’s dominance process (represented by  $U_x$ ) would be normally distributed with respect to every other animal in the herd.

Finally, the assumption is further supported by the usefulness of the model as presented here and elsewhere where it has been proven to be robust in empirical tests with psychological phenomena (Batchelder & Simpson 1989; Alvarado & Jameson 1996; Jameson 1996).

Despite its statistical appropriateness, there might nevertheless be concern that the assumption has implications for the distribution of scale values derived for animal dominance hierarchies. Any method using only ranks produces an ordinal scale and has no meaningful metric-distance function. In particular, for such ordinal ranking scales there is no distance information conveyed by adjacency in the ranking data. The BBS procedure, however, yields probabilities associated with pairs of ranks, so that for any two animals, information on how much one animal dominates the other is given in terms of probability. This represents an improvement over cardinal dominance scales for the purposes of characterizing and predicting closeness of ranks within a dominance structure.

Data exhibiting shifts in dominance over encounters present serious difficulties for deterministic linear dominance models. The BBS model was designed to account for such shifts. The ordering derived from it satisfies the following weak form of stochastic transitivity: if A dominates B in more than 50% of the encounters and B dominates C in more than 50% of the encounters, then A dominates C in more than 50% of the encounters. If the assumptions of the model were to hold and the data consisted of a very sparse sample (for example, a sample like the kind of data collected for this article), then, due to sampling, one would expect part of the data to exhibit violations of the above form of stochastic transitivity. (Of course violations of this form of stochastic transitivity as well as other kinds of intransitivities are taken into account in evaluating how well the stochastic ordering derived from the BBS algorithm fits the observed probabilistic data.)

Stochastic intransitivities that do not result from sampling are a realistic possibility for animal hierarchies: a hierarchy as an exact ordering of individuals is a theoretical construct (Appleby 1993) and methods that use transitivity of dominance to find a ranking (Brown 1975; de Vries 1998) are based on the assumption that the hierarchy is linear. Yet, there is no logical reason why dominance should be transitive; indeed, transitivity is not all-or-nothing but generally intermediate between these extremes (Tufto et al. 1998).

Although stochastic intransitivities not due to sampling are inconsistent with BBS model, the stochastic ordering derived from the BBS algorithm may still be very informative, either as a good linear approximation of the hierarchy or as an ordering that ranks animals in 'group' terms. In the 'group' case, the relative rank of two animals is interpreted in terms of their individual success or failure in encounters with other animals, rather than in terms of their dominating behaviour with respect to each other. In either case, the obtained stochastic ordering can be used to consider correlates of rank, including causes and effects.

Thurstone (1927) formulated a strong version of the Case V model with assumptions of independence and an underlying normal distribution, which involves explicit estimation of a few parameters. As a model of dominance relations, the Case V model can be rejected. This is an important point because the goal in modelling animal

interactions is not so much to determine which of the many possible models might underlie dominance relations, but rather which models best predict and characterize animal behaviours.

### Acknowledgments

This work was in part supported by a National Research Council Research Associateship granted to K.A.J. We thank Louis Narens who made helpful suggestions on the manuscript.

### References

- Alvarado, N. & Jameson, K. A. 1996. New findings on the contempt expression. *Cognition and Emotion*, **10**, 379–407.
- Appleby, M. C. 1982. The consequences and causes of high social rank in red deer stags. *Behaviour*, **80**, 259–270.
- Appleby, M. C. 1983. The probability of linearity in hierarchies. *Animal Behaviour*, **31**, 600–608.
- Appleby, M. C. 1993. How animals perceive a hierarchy: reactions to Freeman et al. *Animal Behaviour*, **46**, 1232–1233.
- Batchelder, W. H. & Bershad, N. J. 1979. The statistical analysis of a Thurstonian model for rating chess players. *Journal of Mathematical Psychology*, **19**, 39–60.
- Batchelder, W. H. & Simpson, R. S. 1989. Rating systems for human abilities: the case of rating chess skill. In: *UMAP Modules in Undergraduate Mathematics and its Applications: Module 698*. Reprinted in *UMAP Modules 1988: Tools for Teaching* (Ed. by P. J. Campbell), pp. 289–314. Arlington, Massachusetts: Consortium for Mathematics and its Applications.
- Batchelder, W. H., Bershad, N. J. & Simpson, R. S. 1992. Dynamic paired-comparison scaling. *Journal of Mathematical Psychology*, **36**, 185–212.
- Beilharz, R. G. & Mylrea, P. J. 1963. Social position and behaviour of dairy heifers in yards. *Animal Behaviour*, **11**, 522–528.
- Boyd, R. & Silk, J. B. 1983. A method for assigning cardinal dominance ranks. *Animal Behaviour*, **31**, 45–58.
- Bradley, R. A. & Terry, M. E. 1952. The rank analysis of incomplete block designs. I. The method of paired comparisons. *Biometrika*, **39**, 324–345.
- Brown, J. L. 1975. *The Evolution of Behaviour*. New York: Norton.
- Chase, I. D. 1974. Models of hierarchy formation in animal societies. *Behavioural Science*, **19**, 374–382.
- 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.
- Clutton-Brock, T. H., Albon, S. D., Gibson, R. M. & Guinness, F. E. 1979. The logical stag: adaptive aspects of fighting in red deer (*Cervus elaphus* L.). *Animal Behaviour*, **27**, 211–225.
- Clutton-Brock, T. H., Guinness, F. E. & Albon, S. D. 1982. *Red deer: Behaviour and Ecology of Two Sexes*. Chicago: University of Chicago Press.
- Cole, B. J. 1981. Dominance hierarchies in *Leptothorax* ants. *Science*, **212**, 83–84.
- Drews, C. 1993. The concept and definition of dominance in animal behaviour. *Behaviour*, **125**, 283–311.
- Elo, A. 1978. *The Rating of Chess Players, Past and Present*. New York: Arco.
- Freeman, L. C., Freeman, S. C. & Romney, A. K. 1992. The implications of social structure for dominance hierarchies in red deer (*Cervus elaphus* L.). *Animal Behaviour*, **44**, 239–245.
- Iverson, G. J. 1987. Thurstonian psychophysics: case III. *Journal of Mathematical Psychology*, **31**, 219–247.

- Iverson, G. J. & Sade, D. S.** 1990. Statistical issues in the analysis of dominance hierarchies in animal societies. *Journal of Quantitative Anthropology*, **2**, 61–83.
- Jameson, K. A.** 1996. Numerical scaling techniques for evaluating semiotic models of orthographies. *Acta Psychologica*, **92**, 169–208.
- Kendall, M. G.** 1955. Further contributions to the theory of paired comparisons. *Biometrics*, **11**, 43–62.
- Roberts, J. M., Jr.** 1990. Modeling hierarchy: transitivity and the linear ordering problem. *Journal of Mathematical Sociology*, **16**, 77–87.
- Sade, D. S.** 1992. Dominance hierarchies as partial orders: a new look at old ideas. In: *Aggression and Peacefulness in Humans and Other Primates* (Ed. by J. Silverburg & J. P. Gray), pp. 57–71. Oxford: Oxford University Press.
- Schein, M. W. & Fohrman, M. H.** 1955. Social dominance relationships in a herd of dairy cattle. *British Journal of Animal Behaviour*, **3**, 45–55.
- Thurstone, L. L.** 1927. A law of comparative judgment. *Psychological Review*, **34**, 273–286.
- Tufto, J., Solberg, E. J. & Ringsby, T.-H.** 1998. Statistical models of transitive and intransitive dominance structures. *Animal Behaviour*, **55**, 1489–1498.
- de Vries, H.** 1998. Finding a dominance order most consistent with a linear hierarchy: a new procedure and review. *Animal Behaviour*, **55**, 827–843.
- Wagnon, K. A., Loy, R. G., Rollins, W. C. & Carrol, F. D.** 1966. Social dominance in a herd of Angus, Hereford and shorthorn cows. *Animal Behaviour*, **14**, 474–479.
- Wei, T. H.** 1952. The algebraic foundations of ranking theory. Ph.D. thesis, University of Cambridge.
- Wood-Gush, D. G. M.** 1955. The behaviour of the domestic chicken: a review of the literature. *British Journal of Animal Behaviour*, **3**, 81–110.
- Yellott, J. I., Jr.** 1977. The relationship between Luce's choice axiom, Thurstone's theory of comparative judgment, and the double exponential distribution. *Journal of Mathematical Psychology*, **5**, 109–144.