



COMMENTARY

Finding an appropriate order for a hierarchy: a comparison of the I&SI and the BBS methods

HAN DE VRIES* & MICHAEL C. APPLEBY†

*Ethology and Socio-ecology Group, Utrecht University

†School of Agriculture, University of Edinburgh

(Received 12 July 1999; initial acceptance 20 August 1999;
final acceptance 30 September 1999; MS. number: SC-1168)

An important topic in the study of social behaviour of animals living in social groups is the description, analysis and explanation of dominance behaviour and relationships. Important questions in this context are (1) whether there are behaviours performed by the animals from which the existence of dyadic dominance relationships can be concluded, and (2) if dominance relationships exist, whether the set of these relationships fits a linear rank order. There are a bewildering number of methods aiming to produce a linear hierarchy from an observed dominance matrix, that is, a matrix containing the numbers of wins and losses of dyadic dominance encounters for each pair of animals. A review of most of these methods is given in [de Vries \(1998\)](#).

Overall, one can distinguish between two types of method, one in which some numerical criterion, calculated for the matrix as a whole, is minimized (or maximized), resulting in a reorganized matrix for which this criterion value is smallest (or largest). Thus, the result produced by this class of methods is a rank order of the animals, that is the most plausible one relative to the specific criterion used, and given the dominance encounters observed. These methods include: (1) [Slater's \(1961\)](#) method of minimizing the number of inconsistencies; (2) [de Vries's \(1998\)](#) I&SI method, which aims to find a rank order that is most consistent with a linear hierarchy by first minimizing the number of inconsistencies *I* and, subsequently, minimizing the total strength of the inconsistencies *SI*, subject to the condition that *I* does not increase; (3) [McMahan & Morris's \(1984\)](#) method of maximizing the likelihood under the assumption of a

paired comparison weak stochastic transitivity (WST) model; (4) [Brown's \(1975\)](#) method of minimizing the proportion of entries below the diagonal; (5) [Boyd & Silk's \(1983\)](#) method of maximizing the likelihood under the assumption of a Bradley–Terry model; and (6) other methods reviewed in [de Vries \(1998\)](#).

The second class of methods aims to provide a suitable measure of individual overall success in the group, from which a rank order can be directly derived. The appropriateness of the rank order thus found is a direct consequence of the suitability of the overall success measure. As yet, a generally accepted success measure appears not to exist, although many different candidates have been put forward. Measures provided by this class of methods include: (1) the well-known index, number of individuals dominated, or (if not all dominance relationships are known) the proportion of individuals dominated; (2) the proportion of total number of encounters won; (3) [Clutton-Brock et al.'s \(1979\)](#) index of fighting success, which takes the strength of the animals beaten into account as well as the strength of the animals lost to, and uses the numbers of dominance and subordination relationships in the definition of the index; (4) [David's \(1987, 1988\)](#) score, which equally reflects the strength of the animals defeated as well as the animals defeated by, but uses in its definition the summed proportions of wins of the individuals encountered, the weighted sum of the scores of the individuals beaten, the summed proportions of losses of the individuals encountered, and the weighted sum of the scores of the individuals lost to; and (5) [Jameson et al.'s \(1999\)](#) score, which takes into account the proportions of wins and losses with others as well as the scores of the others that an individual has met in encounters.

This dichotomy of ranking methods is not exclusive: [Boyd & Silk's \(1983\)](#) method also provides an individual

Correspondence: H. de Vries, Ethology and Socio-ecology Group, Utrecht University, Postbus 80.086, 3508 TB Utrecht, The Netherlands (email: j.devries@bio.uu.nl). M. C. Appleby is at the School of Agriculture Building, University of Edinburgh, West Mains Road, Edinburgh EH9 3JG, U.K.

overall success measure since this method defines a cardinal dominance index D_i in such a way that the binomial probabilities P_{ij} depend logistically on the difference between the dominance indices D_i and D_j .

We present an extended comparison of the two most recently developed methods, de Vries's I&SI method and Jameson et al.'s method. This last method for finding a rank order was introduced by Batchelder, Bershad and Simpson (Batchelder & Bershad 1979; Batchelder & Simpson 1989) and is based on Thurstone's (1927) method of paired comparisons. The method, called the BBS method, is a formalization of an idea originally developed by Elo (1978) for rating chess players. For each individual a scale value is calculated based on the proportion of wins in its encounters with others, its proportion of losses, and the scale scores of the others that it has met in dominance encounters. The equations for calculating these scale values require several assumptions with respect to the outcomes of the encounters to hold, which we discuss below. In contrast, the I&SI method makes use of information only at the level of the dyad, that is, it uses the dominance relationship of each dyad and not the number of dominance encounters. The outcomes (wins and losses) of the dominance encounters are used only to determine the dyadic dominance relationships in the standard way by asymmetries in these outcomes (cf. Appleby 1983; and see below). I&SI is applicable to any set of dominance relationships without making any assumption about the probability distribution of the wins and losses of the dominance encounters, and also allows for the presence of tied and/or unknown relationships.

We do not make a thorough comparison with other methods, since (part of) this comparison has already been done in the papers mentioned above, although we see the need for a more thorough comparison. Another reason to compare these two methods with each other is that Jameson et al. (1999) advised against ranking methods that use the assumption of transitivity of relationships such as I&SI and Brown's (1975) method. They stated that such methods 'may obscure irregularities' by 'tidying up' hierarchies in this way (Jameson et al. 1999, pp. 991–992). They presented Boyd & Silk's (1983) method as one method that addresses this problem, since 'their index of dominance does not assume transitivity' (page 992). However, as we show below, both Boyd & Silk's method and the BBS method do assume transitivity. A further reason is the following. At the end of their paper Jameson et al. wrote: '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'. Indeed, in their paper they presented results that indicate how well the observed dominance relationships can be predicted by the ranking obtained by the BBS method. This offered the opportunity to compare both methods with respect to their predictive strength. Finally, we detected, accidentally, that the BBS method can give rise to counterintuitive, anomalous results. We show this by some specific examples.

Comparing the I&SI and BBS Methods

The purpose of both the I&SI and BBS methods is to obtain a dominance rank order of individuals that is, in some sense, a good estimate of, or approximation to, a linear hierarchy in which there is one top-ranking individual dominating all others, a second one dominating all others except the top one, and so on down to the one at the bottom of the hierarchy which is dominated by all others. The result of applying the I&SI method is a rank ordering of the individuals, whereas applying the BBS method results in obtaining a dominance score for each individual, from which a rank order can be derived. The two methods differ in the following respects.

Type of data used

The difference in type of results obtained is a direct consequence of the difference between the type of data used by either method. The I&SI method uses the dominance relationships of the dyads, and therefore the result of this method can only be a rank ordering of the individuals. No explicit dominance score is provided for the individuals. The dominance relationships can be determined by the outcomes of the dominance encounters observed in some period of time in the following way. A is called dominant to B if A wins more encounters with B than B does against A. If A and B have an equal number of wins and losses the dominance relationship is tied: A and B are said to be equidominant. If no encounters have been observed their relationship is unknown. In both these cases the relationship is called undecided. Alternative ways for determining the dominance relationship are also possible. For instance, in some cases, a researcher might prefer to use stricter criteria for calling A dominant to B (resulting in more relationships being left undecided/unknown) or several different dominance-related behaviours might be combined in some weighted way to determine which individual dominates whom.

The BBS method, on the other hand, uses the outcomes of the individual dominance encounters. These outcomes are used to provide a dominance score for each individual on a numerical scale. So, the BBS method makes full use of the available dominance observations, whereas the I&SI method does not. Besides the implicit assumption that the behaviours by which a win or loss is measured are of equal weight, so that arithmetical operations such as adding and subtraction are justified, there are also distributional conditions (to be discussed below) that must be assumed to hold.

How are the data used?

The I&SI method uses the set of dominance relationships (derived in some way, usually the standard way, from the outcomes of the dominance interactions) in an iterative algorithm for finding a rank order with a minimum number of inconsistencies I (inconsistencies are dyads for which the actual dominance relationship does not agree with the relationship in the hierarchy found), and a minimum value of the total strength of the inconsistencies SI , subject to the condition that I is at its minimum. The strength of an inconsistency is defined as

the absolute difference between the ranks of the two individuals involved in this inconsistency (de Vries 1998, page 830). The algorithm is explained in de Vries (1998). In the final step of the procedure, pairs of animals adjacent to each other in the rank order found, and which have an undecided dominance relationship, are ordered according to the following rule. If $D_i - S_i > D_j - S_j$, then put i above j in the rank order (unless this increases the total strength of the inconsistencies), where D_i is the number of individuals dominated by i , and S_i is the number of individuals by which i is dominated.

The BBS method uses the outcomes of the dominance interactions in the following equation to provide initial estimates of scale values for each animal:

$$s(a_i) = \alpha(2W_i - N_i)/2N_i$$

where $\alpha = \sqrt{2\pi}$ is a constant, W_i is the number of encounters in which animal a_i won, and N_i the number of encounters in which a_i was involved. Subsequently a second equation is used recursively to rescale the animals until their scale scores become invariant:

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

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.

The two algorithms differ as follows. BBS ranks the animals in terms of individual overall success in the whole group, whereas in the I&SI method the criterion of minimizing the number of inconsistencies implies that the dyadic dominance relationship prevails over the 'group' success criterion. With ranking in 'group' terms the relative ranking of two animals is based on their respective wins and losses with all other animals, whereas with ranking in 'dyadic' terms two animals are switched in the current rank order if this decreases the number of inconsistencies. For instance, suppose that A and B are adjacent to each other in the rank order found thus far, and that the overall success of A is higher than that of B, but B dominates A. According to the 'group' criterion A should be ranked above B. According to the criterion of minimizing the number of inconsistencies, however, B should be ranked above A, resulting in a so-called Hamiltonian ranking (a Hamiltonian ranking is one in which for each pair of animals adjacent to each other in this ranking, the one that is ranked highest is actually dominant to the other one).

Assumptions

The I&SI method has been developed with the express intention of devising a method for rank ordering a dominance matrix that requires as few assumptions as possible. Many of the methods for ranking individuals on the basis of their wins and losses of dyadic dominance encounters have originally been developed in the context of a paired comparisons paradigm (David 1988). The necessary conditions for applying such methods are taken care of by the designer of the experiment. When animals living in a social group are observed such conditions may not be

fulfilled (see de Vries 1998 for a more extended discussion of these conditions). Therefore, instead of using the dominance encounter as the observational unit of analysis, the I&SI method is based on the dyad as the observational unit of analysis. In this way the only assumption that must be satisfied for validly applying the I&SI method is that the linearity in the set of dominance relationships is statistically significant, which can be tested with the linearity test (Appleby 1983; de Vries 1995). Although I&SI can also be used if the linearity is nonsignificant, and could then still provide useful insight into the dominance structure of the group, the ranking obtained underrepresents the information present in the full set of dominance relationships.

The equations used in the BBS method, on the other hand, are based on four specific assumptions with respect to the outcomes of the dominance encounters: (1) the underlying distribution of dominance is continuous; (2) the probability that a particular animal will defeat a particular other is constant; (3) the outcome of a particular encounter between any pair of animals is independent of the outcomes of their previous encounters; and (4) the distribution of dominance that underlies the observed behaviour is normal. With respect to the third assumption, Jameson et al. (1999) noted that, although the theory that produced the algorithm makes the assumption of independence of individual dominance encounters, the results of the algorithm can be tested directly without that assumption.

So the main difference between the two methods with regard to assumptions with respect to the data is that the I&SI method sacrifices precision in the result (it provides only a rank order, whereas BBS provides dominance scores), with the advantage that fewer assumptions have to be fulfilled.

The two methods also differ with respect to the type of transitivity that is assumed in the model. I&SI tries to fit a model in which the relationships satisfy a transitive, linear dominance hierarchy to the data. BBS (as well as Boyd & Silk's method for estimating dominance indices D_{ij} , which is based on the assumption that the Bradley-Terry model is appropriate: Boyd & Silk 1983, page 48) tries to fit a more restrictive model to the data, namely a model in which the dominance probabilities P_{ij} satisfy the condition of strong stochastic transitivity. In fact, an even stronger model is entertained, namely one in which the P_{ij} satisfy the condition that these probabilities are a monotonic function of the difference between the true dominance rankings of the animals (Jameson et al. 1999, page 994). This is one of the models in the chain of hierarchy models presented by Iverson & Sade (1990, page 66), for all of which transitivity is the underlying algebraic condition.

(Non)uniqueness of ranking

Slater's (1961) ranking method often does not lead to a unique ranking: several rankings have the same minimum number of inconsistencies. This is the more so when there are unknown dominance relationships present. By adding the extra criterion of minimizing the

total strength of the inconsistencies SI , subject to the condition that I does not increase, a further selection is possible. In addition, the criterion used in the final phase, for ordering the animals adjacent in the rank order that have an undecided relationship, limits the possibility of more than one ranking fulfilling the criteria. Nevertheless, the I&SI method does not necessarily yield a unique ranking. In particular, animals that have few interaction partners can have different positions in the rank order, without a change in the criterion values (or sometimes with only a small change in the value of SI). This means that the ranks of such animals cannot be estimated reliably. The method of McMahan & Morris (1984; i.e. maximize the likelihood under the assumption of a paired comparisons weak stochastic transitivity model) can also lead to more than one possible ranking, all being equally likely under the assumption of weak stochastic transitivity.

Since the BBS method yields a dominance score for each animal, this method necessarily yields a unique ranking. If two or more animals have exactly the same numbers of wins and losses against the other animals, they have the same dominance score, which means that they occupy the same position in the ranking. Nevertheless, the ranking obtained (which may include tied ranks) is unique. This is an advantageous feature of this method. None the less, one may expect here also that for animals having few dominance encounters the reliability of the estimated dominance position is low.

Fit of predicted relationships with observed ones

In this section we apply the I&SI and BBS methods to a specific dominance matrix containing the wins and losses of dyadic dominance interactions. The observed dominance relationships are defined in the standard way by asymmetries in success in the dyadic encounters (cf. Appleby 1983): A is said to dominate B if A has won more encounters from B than B from A. Applying the two methods to the observed dominance matrix yields two rankings from which two sets of predicted dominance relationships predicted by either method can be derived: A is predicted to dominate B if A is higher in the respective rank order than B.

An important question to be asked now is: how well are the observed dominance relationships predicted from the results obtained by either method? Jameson et al. (1999) illustrated the BBS method by applying it to the dominance interactions observed among 68 red deer, *Cervus elaphus*, stags on the Isle of Rum (Appleby 1982). The dominance scale ordering obtained through the BBS method predicts for every pair of animals which one is dominant (Table 3 in Jameson et al. 1999). They calculated a proportional reduction of error measure (Kendall's tau) to determine how well the observed relationships are predicted by the outcome of the BBS method. Since many of the animals did not meet each other and four pairs had an equal number of wins and losses, only the 642 pairs for which a dominance relationship could be established were used in the calculation of the Kendall's tau value. In this way a tau value of 0.853 was obtained (Table 4 in Jameson et al. 1999).

This result can be compared with the result obtained with the I&SI method. Before applying the I&SI method, we tested whether the amount of linearity present in the set of dominance relationships differs significantly from what would be expected by chance for a matrix of this size containing a random set of dominance relationships. Since the observation matrix contains 1632 pairs with an unknown dominance relationship and four pairs with a tied one, we used the randomization test described in de Vries (1995). The linearity index h' (corrected for unknown relationships) is 0.126, which differs significantly from the expected value of 0.043 ($P < 0.001$). Next, we applied the I&SI method to obtain a rank order that is most consistent with a linear hierarchy. The following order was obtained: 1 Maxi, 2 Pete, 3 Blt2, 4 Sx95, 5 Ferd, 6 Brea, 7 Slip, 8 Cona, 9 Torm, 10 Budy, 11 Hami, 12 Boss, 13 Junc, 14 Orph, 15 Cork, 16 Bl45, 17 Yest, 18 Gill, 19 Clyd, 20 Ltsp, 21 Dick, 22 Frod, 23 Sc01, 24 Yhoc, 25 Stul, 26 Sbbw, 27 Upt3, 28 Gips, 29 Talc, 30 Rgrc, 31 Hect, 32 Crtc, 33 Solo, 34 Sooc, 35 Clec, 36 Bdlt, 37 Bltc, 38 Feat, 39 Dor5, 40 Whic, 41 Choc, 42 Licc, 43 Redc, 44 Recc, 45 Cocc, 46 Trcc, 47 Fecc, 48 Clcc, 49 Upcc, 50 Broc, 51 Myrc, 52 Spri, 53 Rgcc, 54 Elsi, 55 M4, 56 Bst4, 57 Ta24, 58 Sco5, 59 Tr34, 60 Coc4, 61 Cr14, 62 Rg14, 63 Gre4, 64 Alto, 65 Tal5, 66 Bst5, 67 Che5, 68 Rgr4.

The number of inconsistencies for this rank order is 8 with a total strength of 64. The following eight pairs of stags have a dominance relationship that is inconsistent with the linear order found by the I&SI procedure. The strength of each inconsistency, that is, the absolute difference between the ranks of the two animals, is given in parentheses. Stul > Gill (7), Hect > Yhoc (7), Sooc > Stul (9), Choc > Sbbw (15), Solo > Talc (4), Choc > Hect (10), Myrc > Trcc (5), Rgcc > Trcc (7), where A > B means A dominates B. The order found by the I&SI method predicts for every pair of animals which one is dominant. Of the 642 dyads for which a dominance relationship could be established on the basis of the observed wins and losses, 634 are predicted correctly by the I&SI method. Similarly, as has been done above, Kendall's tau was calculated to express how well the observed 642 relationships are predicted by the linear order found by the I&SI method. With the I&SI method a Kendall tau value of 0.975 is obtained. The difference between the two proportions of correctly predicted relationships is statistically significant when tested with a chi-square test for a 2×2 contingency table containing the numbers of correctly and wrongly predicted relationships for both methods (I&SI: 634 correct, eight wrong; BBS: 595 correct, 47 wrong; $\chi^2_1 = 27.4$, $P = 0.0001$).

To examine the BBS method further, Jameson et al. (1999) excluded three sets of pairs with a predicted dominance probability near to 0.5, and then calculated the Kendall's tau for each of the set of pairs retained (see Table 1). Similarly, we calculated the Kendall's tau for each of these three sets of retained pairs to see how well the relationships predicted by the linear order found by the I&SI method agree with the observed (retained) ones. In every case the I&SI method predicts the observed relationships better than the BBS method (see Table 1).

Table 1. Agreement between dominance relationships predicted by either the BBS method or the I&SI method and those actually observed

Probability range excluded	Pairs excluded	Pairs retained	Kendall's tau	
			BBS method	I&SI method
None	0	642	0.853	0.975
0.45–0.55	56	586	0.925	0.973
0.40–0.60	111	531	0.958	0.981
0.30–0.70	195	447	0.991	0.996

The first row shows Kendall's tau expressing the agreement between 642 predicted and observed dominance relationships. Successive rows were calculated as pairs of animals with relatively poor predictability (as calculated by the BBS method) were excluded (cf. Table 4 in Jameson et al. 1999).

Predictive strength

Another appropriate question to ask is how well the observed dominance relationships are predicted from the results obtained by either method using only a subset of the dominance data. Jameson et al. (1999) continued to explore the predictive strength of the BBS method by excluding randomly 65 (10%) pairs from the original 642 pairs. These 65 pairs were set to zeros in their matrix cells as if no encounter had been recorded. Subsequently, scale values and a linear order were obtained with the BBS procedure. Again Kendall's tau was calculated as a measure of how well the observed 642 pairs can be predicted from the order found by applying the BBS method to 90% of the data. A tau value of 0.847 was obtained.

In the same way, we applied the I&SI method to 90% of the data, resulting in a linear order and predicted dominance relationships for all pairs of animals. The agreement between the 642 observed and predicted relationships is stronger than for the BBS method: Kendall's tau equals 0.956. To obtain a fuller impression of the predictive strength of the I&SI method we also calculated tau values for linear orders obtained through the I&SI method using subsets of the data varying between 80 and 50% (see Table 2). Even when using only 50% of the data as a basis, the resulting tau value is high, namely 0.882. This means that 604 of the 642 relationships are correctly predicted by the linear order found by applying the I&SI method to only 321 pairs.

Individual overall success

We showed above that the I&SI ranking predicts relationships better than the BBS ranking. Another criterion by which the rankings obtained by the two methods might be compared is in how far these rankings indicate overall success in the group. This, however, is problematic since there is, as yet, no generally accepted measure of individual overall success. Indeed, the BBS method producing a dominance score for each individual is an attempt to provide such a measure. None the less, to present some indication of how far these two methods differ in this respect, the rankings produced by I&SI and

Table 2. Agreement between dominance relationships predicted by either the BBS method or the I&SI method and those observed, using only a subset of all observed pairs

	Percentage of data used					
	100	90	80	70	60	50
I&SI*	0.975	0.956	0.947	0.928	0.894	0.882
BBS†	0.853	0.847				

*Agreement between dominance relationships predicted by the I&SI method and the 642 relationships actually observed. Values of Kendall's tau are presented for predictions based on the complete data set (100% of the pairs) and predictions based on only a subset of the observed pairs (90–50% data).

†The same using the BBS method for 100% and 90% data (cf. Table 4 in Jameson et al. 1999).

BBS can be compared with other methods that produce a ranking based on some measure of overall success in the group. One such method is based on the numbers of animals beaten and lost to, and another is based on the summed proportions of wins and losses per dyad. When these methods are applied to the stags' dominance matrix, the results turn out to be very similar. Rankings produced by I&SI and BBS have almost identical correlation coefficients with the ranking produced by the index of Clutton-Brock et al. (1979): $r_s=0.959$ and 0.956 respectively ($N=68$ stags). They also have similar correlation coefficients to those obtained with the method suggested by David (1987, 1988, pp. 107–108): $r_s=0.907$ and 0.877 , respectively. With cruder measures of success, the ranking produced by BBS tends to correlate slightly more closely than does that from I&SI. For example, the BBS ranking correlates with 'proportion of encounters won' with $r_s=0.871$; the coefficient for I&SI is 0.817 .

However, there are also matrices for which the ranking obtained by BBS diverges from those obtained by I&SI and the methods of Clutton-Brock et al. and David. Below we present an example matrix (see Table 4 in the next section) for which the rankings obtained by I&SI, Clutton-Brock et al.'s and David's methods all completely agree with each other. This rank order is clearly the one that is intuitively the best ranking for this matrix. In contrast, the ranking produced by BBS (as well as the one derivable from 'proportion of encounters won') clearly deviates from the intuitively best ranking.

Anomalous Results of the BBS Method

It turns out that the BBS method can give anomalous results. We show this with some specific examples. Table 3 presents the scale scores of the stags Pete and Maxi for three matrices that have been slightly modified from the original matrix. The first row presents Pete and Maxi's scores for the original matrix; Pete has a score of 3.0073, Maxi of 2.3447 (cf. Table 3 in Jameson et al. 1999). In the first modified matrix we have increased several numbers of wins by Maxi, in such a way that Maxi wins twice against Pete and wins more often or equally often against any animal that Pete beats. More

Table 3. Values of the scale scores of the stags Pete and Maxi resulting from applying the BBS method to the original matrix of dominance interactions and three modified matrices

Matrix	Dyad	Original wins	Changed into	Pete's score	Maxi's score
Original	Maxi Pete	0	0	3.0073	2.3447
1	Maxi Pete	0	2	2.8477	2.4669
	Maxi Torm	0	5		
	Maxi Blt2	7	13		
	Maxi Sx95	2	4		
	Maxi Budy	3	8		
	Maxi Upt3	0	7		
	Maxi Gips	0	1		
	Maxi Alto	0	1		
2	Maxi Alto	0	2	3.0125	2.3043
3	Pete Myrc	1	10	2.2793	2.3370
	Pete Tal5	1	10		
	Pete Alto	1	10		

specifically, in the modified matrix there are 28 stags that Maxi beats more often than Pete does; there are 13 stags that Maxi beats as often as Pete does; and there are 25 stags that do not interact with either Pete or Maxi. None of the 66 stags ever beats Pete or Maxi. Despite these changes, which mean that Maxi is more successful than Pete over a lot of stags and as successful as Pete over 25 other stags, Pete's score of 2.8477 is higher than Maxi's score of 2.4669. This counterintuitive result is due to the second equation containing the term Q_i , the mean scale score of those animals that animal i met in agonistic encounters. Indeed, if the original matrix is modified such that Maxi has two wins against the low-ranking Alto instead of none, Maxi's score decreases to 2.3043 which is lower than his original score; and Pete's score increases (Table 3, matrix 2). So Maxi decreases his score by beating a low-ranking animal. Similarly, if Pete wins 10 encounters with the low-ranking stags Myrc, Tal5 and Alto, this results in his score decreasing below that of Maxi (Table 3, matrix 3).

This anomaly means that Pete obtains a higher score than Maxi, where in fact the reverse would be more appropriate, given the distributions of Pete's and Maxi's wins (Table 3, matrix 1). This higher score implies that Pete is predicted to be dominant to Maxi while in fact Maxi dominates Pete (Maxi beats Pete twice and Pete never beats Maxi in this example matrix), so it is also an example of the general point made above, that BBS does not always predict relationships correctly. The fact that Pete's score, which is supposed to be an indication of Pete's overall success, is greater than Maxi's, while Maxi's overall success is clearly higher than Pete's, indicates a failure in BBS's aim of scoring overall success.

To show more clearly the possible impact of this failure on the results obtained by BBS, we applied this method to a smaller matrix containing fictitious numbers of wins and losses (Table 4).

The matrix in Table 4 is ordered according to the ranking found by the I&SI method. When the individuals are ranked according to the scores obtained by BBS, a quite different rank order b, a, e, c, d, f, g results. This completely counterintuitive rank order is due to animal g having a low score (−1.582), which causes a's score to fall below b's, and c's score to fall below e's, via the term Q_i in the second equation of BBS. As an extra comparison we also present the scores obtained by David's method and the values of the Clutton-Brock index as well as the proportions of encounters won. The rankings found by the first two methods both agree with the I&SI ranking.

These results raise the question of whether BBS can be improved, by a modification that avoids this anomaly. Such a modification is beyond the scope of the present paper, but we can note that the anomaly is avoided by three comparable methods aiming to define an appropriate measure of individual overall success in the group. The Kendall–Wei approach uses the sum of the scores of the individuals beaten, not those lost to (cited by David 1987). In the scoring formula of David (1987, 1988, pp. 107–108) the summed proportions of wins of the individuals encountered is added to the weighted sum of the scores of the individuals beaten, from which is then subtracted the summed proportions of losses of the individuals encountered and the weighted sum of the scores

Table 4. Fictitious dominance matrix ordered according to I&SI method

	a	b	c	d	e	f	g	BBS score	David's score	Clutton-Brock index	Proportion of encounters won
a	*	1	1	4	0	3	6	1.547	15.0	16.00	1.00
b	0	*	1	4	0	0	0	1.997	8.0	5.00	0.83
c	0	0	*	1	1	3	14	0.716	7.0	2.75	0.90
d	0	0	0	*	2	2	1	0.430	0.0	1.00	0.36
e	0	0	0	0	*	17	2	0.935	−4.0	0.50	0.86
f	0	0	0	0	0	*	12	−0.579	−10.0	0.17	0.32
g	0	0	0	0	0	0	*	−1.582	−16.0	0.06	0.00

Individual scores for three different scoring methods, and the proportion of encounters won are also shown. See text for details.

of the individuals lost to. The index of Clutton-Brock et al. (1979) adds the number of animals beaten to the total number they beat plus one, which is then divided by the sum of the number of animals lost to and the total number to which they lost plus one.

Conclusion

We compared two methods for finding a linear rank order, the I&SI and the BBS methods, in several respects. The I&SI method is based on the binary dominance relationships derivable from the wins and losses of each pair of animals, and provides a ranking of the animals. In contrast, the BBS method is based on the numbers of wins and losses of each animal against all others, and provides for each individual an explicit dominance score, from which a unique ranking can be derived. The more extended use of the available dominance data by the BBS method requires more assumptions to be satisfied than are necessary for application of the I&SI method. Indeed, I&SI was developed with the express intention of having a ranking method that demands as few conditions to be fulfilled by the data as possible. It sacrifices precision in the result (it provides a rank order that is not necessarily unique, whereas BBS provides dominance scores), with the advantage that fewer conditions have to be fulfilled. Even though I&SI makes fewer assumptions than BBS, it was significantly better at predicting observed relationships, when both methods were applied to a matrix of dominance encounters among 68 stags. When we used only 50% of the data as a basis, the number of relationships predicted correctly by I&SI (604 out of 642) was still numerically higher than that from BBS when we used 100% of the data as a basis (595 out of 642).

Besides comparing the predictive strength of the two methods, we also compared how far the rankings obtained by both methods indicate overall success in the group. This is problematic since there is no generally accepted measure of individual overall success. Indeed, BBS is an attempt to provide such a measure. Nevertheless, to give some indication how the methods may differ in this respect, we compared the rankings obtained by I&SI and BBS with the rankings obtained by two other methods that attempt to rank individuals in terms of an overall success score (Clutton-Brock et al. 1979; David 1987, 1988). It turned out that the correlations of the I&SI and BBS rankings with each of these rankings were strong and quite similar when applied to the dominance matrix of the 68 red deer stags. However, when applied to a small (fictitious) dominance matrix, the BBS ranking clearly deviated from the ranking obtained by I&SI. This I&SI ranking, which is intuitively best for this matrix, agreed completely with the rankings obtained by the Clutton-Brock index and David's score.

In conclusion, by applying BBS to several (fictitious) matrices, we found that BBS can give counterintuitive, anomalous results, in the sense that animal A can obtain a lower score than B although A beats B and also beats any other animal at least as often as B does. This counter-

intuitive result is due to the second equation of the BBS method containing the term Q_i , the mean scale score of those animals that animal i met in agonistic encounters. Whether this failure has a major impact on the outcome depends on the spread of the wins and losses in the matrix. Anyhow, this suggests that BBS could be modified to improve its aim of providing an appropriate measure of individual overall success in the group.

Acknowledgments

Jan van Hooff made helpful comments on the manuscript. We thank two referees for their stimulating comments.

References

- 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.
- 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. *UMAP Modules in Undergraduate Mathematics and its Applications: Module 698*. Reprinted in Campbell, P. J. (Ed.) 1988. *UMAP Modules 1988: Tools for teaching*, pp. 289–314. Arlington, Massachusetts: Consortium for Mathematics and its Applications.
- Boyd, R. & Silk, J. B. 1983. A method for assigning cardinal dominance ranks. *Animal Behaviour*, **31**, 45–58.
- Brown, J. L. 1975. *The Evolution of Behavior*. New York: Norton.
- Clutton-Brock, T. H., Albon, S. D. & Guinness, F. E. 1979. The logical stag: adaptive aspects of fighting in red deer (*Cervus elaphus* L.). *Animal Behaviour*, **27**, 211–225.
- David, H. A. 1987. Ranking from unbalanced paired-comparison data. *Biometrika*, **74**, 432–436.
- David, H. A. 1988. *The Method of Paired Comparisons*. New York: Hafner.
- Elo, A. 1978. *The Rating of Chess Players, Past and Present*. New York: Arco Publishing.
- 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., Appleby, M. C. & Freeman, L. C. 1999. Finding an appropriate order for a hierarchy based on probabilistic dominance. *Animal Behaviour*, **57**, 991–998.
- McMahan, C. A. & Morris, M. D. 1984. Application of maximum likelihood paired comparison ranking to estimation of a linear dominance hierarchy in animal societies. *Animal Behaviour*, **32**, 374–378.
- Slater, P. 1961. Inconsistencies in a schedule of paired comparisons. *Biometrika*, **48**, 303–312.
- Thurstone, L. L. 1927. A law of comparative judgment. *Psychological Review*, **34**, 273–286.
- 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.