

A likely ranking interpolation for resolving dominance orders in systems with unknown relationships

G. Wittemyer¹⁾ & W.M. Getz

(University of California at Berkeley, Department of Environmental Science, Policy, and Management, 201 Wellman Hall, Berkeley, CA 94720-3112, USA)

(Accepted: 6 June 2006)

Summary

In many animal systems agonistic interactions may be rare or not overt, particularly where such interactions are costly or of high risk as is common for large mammals. We present a technique developed specifically for resolving an optimized dominance order of individuals in systems with transitive (i.e. linear) dominance relationships, but where not all relationships are known. Our method augments the widely used I&SI method (de Vries, 1998) with an interpolation function for resolving the relative ranks of individuals with unknown relationships. Our method offers several advantages over other dominance methods by enabling the incorporation of any proportion of unknown relationships, resolving a unique solution to any dominance matrix, and calculating cardinal dominance strengths for each individual. As such, this method enables novel insight into difficult to study behavioural systems.

Keywords: dominance hierarchy, algorithm, social agonistic interactions.

Introduction

Dominance relationships are an important construct of social interactions, offering insight into behavioural and fitness related differences among individuals (Krebs & Davies, 1987). Differentiation in fecundity and survival rates in relation to social rank have been demonstrated in a variety of species

¹⁾ Corresponding author's e-mail address: georgew@nature.berkeley.edu

	A	B	C	D	E	F	G	H	I	J	K	L	N	M	O	P	Q	R	S	T
A	*	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1
B	0	*	1	1	1	0	0.9	1	1	1	1	1	1	1	1	1	1	1	1	1
C	0	0	*	1	1	0.7	1	1	1	1	1	1	1	1	1	1	1	0.9	1	1
D	0	0	0	*	1	1	1	1	1	1	0	1	1	0.9	1	1	1	1	1	1
E	0	0	0	0	*	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
F	0	0	0.3	0	0	*	1	1	1	1	1	1	1	0	1	1	1	1	1	1
G	0	1	0	0	0	0	*	1	1	0.7	1	1	1	1	1	1	1	1	1	1
H	0	0.1	0	0	0	0	0	*	1	1	1	1	1	1	0.7	1	1	1	1	1
I	0	0	0	0	0	0	0	0	*	1	1	1	1	0	1	1	1	1	1	1
J	0	0	0	0	0	0.3	0	0	0	*	1	1	1	1	0	1	1	1	1	1
K	0	0	0	1	0	0	0	0	0	0	*	1	0	1	1	1	0.8	1	1	1
L	1	0	0	0	0	0	0	0	0	0	0	*	1	0.9	1	1	1	1	1	1
N	0	0	0	0	0	0	0	0	0	0	0	0	*	1	1	0.7	1	1	1	1
M	0	0	0	0.1	0	0	0	0	1	0	1	0	0	*	1	1	1	1	1	1
O	0	0	0	0	0	1	0	0	1	0	0.1	0	0	*	1	1	1	1	1	1
P	0	0	0	0	0	0	0.3	0	0	0	0.3	0	0	0	*	0.6	1	0	1	1
Q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.4	*	1	1	1	1
R	0	0	0.1	0	0	0	0	0	0	0.2	0	0	0	0	0	0	*	1	0.8	1
S	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	*	1	1
T	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.2	0	*

Figure 1-A. A sample interaction matrix with a 5% chance of circularity randomly imposed upon each dyad. The proportion of wins by an individual is denoted by the number between 0 and 1 in the individual's row and correspondingly the proportion of losses is denoted by the number between 0 and 1 in the individual's column. Initially this 20×20 matrix had 9 inconsistencies (I) with the sum of inconsistencies (S_I) equalling 47, but the relative ranks of individuals N and M were switched resulting in $I = 8$ and $S_I = 48$.

(reindeer: Holand et al., 2004, clown anemonefish: Mitchell, 2005, rabbits: von Holst et al., 2002, mountain goats: Côté & Festa-Bianchet, 2001, black-capped chickadees: Otter et al., 1998, spotted hyaena: Holekamp et al., 1996). Dominance structure is also a key factor influencing the spatial properties of animal populations, such as partitioning of resources through differentiation in home range size and movement (Boydston et al., 2003; Say & Pontier, 2004; Hansen & Closs, 2005). Social rank can even impact the health of individuals (Sapolsky, 2005). Thus the quantification of dominance interactions is an important aspect of behavioural ecology and the development of quantitative methods to resolve dominance relations has been the focus of numerous studies (as reviewed in de Vries, 1998).

Standard methods for deriving the dominance hierarchy of interacting individuals rely on the scoring of pair-wise competitive interactions (Boyd & Silk, 1983; Martin & Bateson, 1993). The results of these interactions, wins or losses, are arranged in a dyadic interaction matrix where the ij^{th} element represents the observed dominance relationship of i to j (Figure 1-A). In a discrete formulation, this element is a 1-0 win-lose metric. In a continuous formulation, it may take on any value between 0 and 1 and represent the probability of dominance (number of wins/number of total contests per pair). Conventions for resolving dominance ordering from dyadic interaction matrices can be broadly grouped into two categories (Albers & de Vries, 2001;

	A	B	C	D	E	F	G	H	I	J	K	L	N	M	O	P	Q	R	S	T
A	*	1				1	1	1					1						1	1
B	0	*					1		1					1						
C			*			1	1	1	1	1	1	1						1	0.9	1
D				*			1		1		1	1	1	1	1			1	1	1
E	0	0			*			1								1				
F						*			1	1	1	1	1	1	0	1		1		
G	1	0	0				*		0.7				1	1	1			1		
H	0	0	0					*	1	1	1	1	1	1	1			1	1	1
I	0	0	0						*	1	1	1	1	1	1			1	1	1
J			0	0		0	0.3	0		*	1	1	1	1	1			1	1	1
K								0	0		*	1	0	1				1	1	1
L				0	0	0	0	0	0	0	*	1	1	0.9	1	1	1	1	1	1
N	0	0	0	0	0	0	0	0	0			*	1	1	1	1	1	1	1	1
M	0	0	0	0	0	0							*	1	1	1	1	1	1	1
O		0		1	0	0	0				0	0.1	0			*	1	1	1	1
P			0	0										0			*	1	0	1
Q		0	0								0			0	0			*	1	1
R	0	0.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	*	1	0.8	1
S	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	*	1	1
T	0									0	0	0	0	0	0	0	0	0.2	*	1

Figure 1-B. Unknown relationships are depicted as empty elements in an interaction matrix. In order to test our method, 60% of elements from the complete matrix were randomly removed and the results of our analysis compared with the known rank order (note the four inconsistencies (L, A), (K, D), (M, I) and (O, J) that were randomly eliminated in the data set).

Gammell et al., 2003): *optimisation methods* generally resolve the rank order of individuals by solving for the minimum number of inconsistencies or some other numerical criterion solved for the dominance matrix as a whole (Crow, 1990; de Vries & Appleby, 2000) and *strength methods* that derive a cardinal rank measure for an individual’s overall position in the hierarchy (Clutton-Brock et al., 1979; David, 1987). Optimisation methods generally resolve the rank order of individuals by solving for the minimum number of circular dominance interactions (i.e., interactions of the form A dominates B who dominates C who dominates A). These methods result in ordinal ranks requiring non-parametric methods of analysis. Strength methods use the difference between the sum of wins and the sum of losses as a measure of an individual’s rank, providing a quantitative numerical dominance value that offers insight into the differences in relative dominance strengths of individuals. Such dominance values can be employed in parametric analyses, but may offer sub-optimal ordinal rankings in the presence of circular interactions. Neither of these rank methods appear to be generally accepted (de Vries & Appleby, 2000; Langbein & Puppe, 2004), and both methods are challenged by sparse data sets that do not contain information on the relationships among all dyads.

Agonistic interactions may be rare, subtle, or cryptic despite the importance of dominance hierarchies in a system. Infrequent dominance interactions tend to be common in species where agonistic interactions are poten-

	C	A	D	G	B	E	H	F	I	J	M	K	L	N	O	S	P	Q	R	T
C	*	0.52	0.55	1	0.6	1	1	0.67	0.69	1	0.74	0.76	1	1	0.83	1	0.88	1	0.9	0.95
A	0.48	*	0.52	0.55	1	0.6	1	1	0.67	0.69	0.71	0.74	0.76	1	0.81	0.83	0.86	0.88	1	1
D	0.45	0.48	*	1	0.55	0.57	0.6	0.62	0.64	1	0.69	0.71	0.74	1	1	1	0.83	1	0.88	0.9
G	0	0.45	0	*	1	0.55	0.57	0.6	0.62	0.67	1	0.69	0.71	1	1	1	0.81	0.83	0.86	0.88
B	0.4	0	0.45	0	*	1	0.55	0.57	1	0.62	1	0.67	0.69	0.71	0.74	0.76	0.79	0.81	0.83	0.86
E	0	0.4	0.43	0.45	0	*	1	0.55	0.57	0.6	0.62	0.64	0.67	0.69	0.71	0.74	1	0.79	0.81	0.83
H	0	0	0.4	0.43	0.45	0	*	0.52	0.55	1	0.6	1	1	1	1	1	0.74	0.76	1	0.81
F	0.33	0	0.38	0.4	0.43	0.45	0.48	*	1	1	0.57	0.6	1	1	0	0.69	1	0.74	1	0.79
I	0.31	0.33	0.36	0.38	0	0.43	0.45	0	*	1	0.55	0.57	1	0.62	1	0.67	0.69	0.71	1	0.76
J	0	0.31	0	0.33	0.38	0.4	0	0	0	*	0.52	1	0.57	1	0.62	0.64	0.67	0.69	1	0.74
M	0.26	0.29	0.31	0	0	0.38	0.4	0.43	0.45	0.48	*	1	0	0.57	0.6	0.62	1	0.67	1	1
K	0.24	0.26	0.29	0.31	0.33	0.36	0	0.4	0.43	0	0	*	1	0.55	1	0.6	0.62	0.64	0.67	1
L	0	0.24	0.26	0.29	0.31	0.33	0	0	0	0.43	1	0	*	1	0.9	1	0.6	1	0.64	1
N	0	0	0	0	0.29	0.31	0	0	0.38	0	0.43	0.45	0	*	1	0.55	0.57	0.6	1	1
O	0.17	0.19	0	0	0.26	0.29	0	1	0	0.38	0.4	0	0.1	0	*	1	0.55	1	0.6	0.62
S	0	0.17	0	0	0.24	0.26	0	0.31	0.33	0.36	0.38	0.4	0	0.45	0	*	1	0.55	0.57	0.6
P	0.12	0.14	0.17	0.19	0.21	0	0.26	0	0.31	0.33	0	0.38	0.4	0.43	0.45	0	*	0.52	1	1
Q	0	0.12	0	0.17	0.19	0.21	0.24	0.26	0.29	0.31	0.33	0.36	0	0.4	0	0.45	0.48	*	1	0.55
R	0.1	0	0.12	0.14	0.17	0.19	0	0	0	0	0.33	0.36	0	0.4	0.43	0	0	*	0.8	
T	0.05	0	0.1	0.12	0.14	0.17	0.19	0.21	0.24	0.26	0	0	0	0	0.38	0.4	0	0.45	0.2	*

Figure 1-C. The unknown elements in the matrix are filled in using a likely-ranking interpolation (equation 7). The best rank order has been solved for the sparse data set by minimizing the values of I and S_I (equations 2 and 5), yielding $I = 2$ and $S_I = 9$ for this example, then resolving the relative ranks of unknown dyads using a strength metric (equation 9), and finally minimizing the sum of elements below the matrix diagonal. Note the changes in the position of individuals B & G and P & S as a result of resolving their inconsistencies in the previous matrix. Also, the strength of the inconsistency between individuals K & M has decreased from 3 to 2 by altering the ranks of L, K, and M with the result that the inconsistency was assigned to M & L.

tially costly or life threatening, as has been observed in elephants (Lee, 1987; Archie et al., 2006) and grizzly bears (Gende & Quinn, 2004). As a result, dominance relationships among all pairs of individuals may not be easily ascertained, complicating definition of ranks (see discussion in de Vries, 1995). As the number of individuals included in a study increases, the number of unknown relationships tends to increase (relationships increase quadratically with the number of individuals, thus sampling must increase quadratically to sample a fixed proportion of dyads). This problem is often encountered in uncontrolled field studies where dyads cannot be forced to compete (Tarvin & Woolfenden, 1997; Gende & Quinn, 2004; Roden et al., 2005; Archie et al., 2006). In many studies, ascertaining the general position of individuals with few observed interactions is essential for the realization of a dominance hierarchy and subsequent analysis of behavioural differentiation among in-

Individual	Rank Strength
C	13.18
A	11.20
D	10.32
G	7.48
B	5.90
E	4.00
H	5.52
F	4.72
I	2.66
J	-1.26
M	-0.08
K	-1.60
L	-1.00
N	-5.84
O	-5.88
S	-7.76
P	-7.18
Q	-8.28
R	-12.92
T	-13.18

Figure 1-D. The optimised individual rank orders and their strengths calculated from matrix 1 C, as described in the text. Note that the individual strength metrics are not completely monotonic. This is a consequence of initially deferring to the I&SI minimization procedure as producing the best ordering. For example, the strength of individual M is greater than that of individual J but J is ranked higher as a result of minimizing the strength of the inconsistency of dyad (M, L).

dividuals. Few methods, however, are designed to resolve the rank order of a system with few observations and multiple unknown relationships.

Here we present a *likely ranking interpolation* approach specifically designed for transitive (i.e. linear) dominance systems, the quantitative assessment of which is imperative prior to implementation (see discussion below), where agonistic interactions occur infrequently resulting in multiple unknown relationships. For such systems, statistically based methods for assigning dominance ranks such as Bradley-Terry models (1952) are untenable because far too little interaction data are available. Our technique builds directly upon the widely used I&SI method (de Vries, 1998) that was designed to determine the rank order of social systems containing unknown and/or circular relationships among individuals (I refers to the number of inconsistencies and SI to the strength of inconsistencies – see description of the I&SI method below), but with the limitations of not resolving a unique rank order

for all systems with unknown relationships nor enabling the quantification of individual dominance strength metrics (often preferred for statistical analyses of rank driven behavioural impacts). Our method offers the advantages of both the optimisation and strength categories of methods to overcome the problems associated with multiple unknown relationships, thereby simultaneously yielding an optimised rank order solution, a unique rank order, and cardinal strengths of individuals in the hierarchy. Hence it is a more informative method for ranking individuals when data are sparse, and it provides insights beyond those offered by either type of method alone.

Description of method

Our method enhances the functionality of the I&SI method (de Vries, 1998) for addressing the problem of unknown relationships and enables the formulation of cardinal dominance strengths from the optimised rank order. The expanded method entails a five step process to resolve the nearest linear rank order from a set of dyadic interaction data. For clarity, we summarize the steps of the I&SI method (de Vries, 1998) in steps 1 and 2 below. The third step applies a likely-rank interpolation to quantitatively estimate unknown relationships, necessary for the fulfilment of step four. The fourth step combines observed and interpolated data to solve individual dominance strength metrics, which are then employed to resolve the relative ranks of unknown dyads. And the fifth step solves a minimum matrix metric to arrive at the unique, near linear rank ordering of individuals.

The optimised rank ordering: the I&SI method

In systems with full or partial knowledge, a rank order is termed Hamiltonian when no individual dominates the individual directly above him (McMahan & Morris, 1984). Assuming the elements of a dominance matrix are composed of dominance probabilities or 1-0 metrics, an optimised rank order can be derived by minimizing the sum of the matrix elements to the left of the matrix diagonal or maximizing the sum of the matrix elements to the right of the matrix diagonal.

Formally, let a_{ij} denote the degree (or probability) to which individual i dominates individual j . Since $a_{ij} + a_{ji} = 1$, it follows that

$$a_{ji} = 1 - a_{ij}, \quad i, j = 1, \dots, n, \quad (1)$$

where n is the number of individuals in the system. Define the upper and lower triangular sums of the dyadic interaction matrix A of elements a_{ij} respectively as

$$H = \sum_{i=1}^n \sum_{j=i+1}^n a_{ij} \quad \text{and} \quad I = \sum_{i=2}^n \sum_{j=1}^{i-1} a_{ij}. \quad (2)$$

From relationship (1), this implies

$$H = \frac{n(n-1)}{2} - I. \quad (3)$$

The optimised rank order occurs when H is a maximum or equivalently I is a minimum. Under conditions of perfect knowledge and linearity (i.e., all relationships are known and higher ranking individuals always defeat those of lower rank), $I = 0$ and the rank order solved using a strength method will be the same as the Hamiltonian rank order obtained using an optimisation method. When conditions of perfect linearity and knowledge are not met the results of these two classes of methods diverge.

The presence of circular polyads (e.g., A dominates B, which dominates C, which dominates A) in the data result in a number of inconsistencies I in the rank order (given by equation 2). In the I&SI method, a flipping algorithm is implemented that switches the relative positions of individuals in the rank order (Johnson et al., 1982; Roberts, 1990) in a direction that ultimately leads to the local minimum for I (equation 2; Figure 1-A). We adopt this algorithm in **Step 1** of our method.

Once this minimum is obtained, the I&SI method further restricts solutions to the rank order through the minimization of the sum of differences in the ranks of all inconsistent dyads. Specifically, if the $i - j$ dyad is inconsistent then $a_{ij} = 1$ (or > 0.5 in the case of probabilistic element values) even though $i > j$. Define

$$\psi_{ij} = \begin{cases} i - j & \text{whenever } a_{ij} = 1 \\ 0 & \text{otherwise} \end{cases} \quad \text{when } i > j. \quad (4)$$

de Vries' approach is to minimize the total strength of the inconsistencies

$$S_I = \sum_{i>j}^n \psi_{ij} \quad (5)$$

in a procedure that he refers to as the I&SI method (Figure 1-A). Thus, **Step 2** of our method finds the ordering that minimizes S_I , subject to the condition

that I does not increase. De Vries (1998) compared the I&SI method with a variety of published ranking methods, demonstrating the I&SI method offered the most robust dominance ordering in a near linear system.

The I&SI approach finds a local minimum. To ensure a good solution, the procedure is repeated for multiple, different initial orderings of individuals organized according to the following two-step procedure: (a) individuals are first ordered by their win-loss ratio, where pairs with equal ratios are ordered by numbers of wins; (b) a permutation technique is implemented that randomly switches individuals a random number of times from the initial ordering such that a spectrum of initial conditions are searched to ensure that the final solution is not a local optimum. The rank order for each starting position is then solved by inserting dominant individuals above their subordinates. If the total change in the number of inconsistencies (equation 2) followed by the total rank sum difference of these inconsistencies (equation 5) increases as a result of the row insertion, the insertion is reversed. Finally, the rank order is regarded as having converged to an acceptable solution when the value of I followed by the value of S_I remains unchanged after repeated permutation runs. Generally, 100 runs is regarded as sufficient for locating the minimum of both the I and S_I metrics (de Vries, 1998). We stress, however, that minimizing I and S_I does not in general produce a unique rank order when the data are incomplete (unknown dyadic relationships exist) or when multiple orderings of individuals may offer the same I and S_I values. Our method offers a solution to this problem through the implementation of an interpolation function to resolve unknown relationships in combination with the use of strength metrics to order closely ranked individuals.

Resolving unknown relationships

Treatment of unknown relationships has been handled in different ways: some methods assign the unknown dyad the equivalent value for a tie (e.g., 0.5 if the dominance is scaled between 0 and 1; Appleby, 1983) while others ignore the relationship between such dyads in their overall ranking system, except for adjacently ranked individuals (de Vries, 1998). In strength-based methods both elements of unknown dyads may be assigned zero (Clutton-Brock et al., 1979; David, 1987). Assessing dominance relationships across groups of individuals with large proportions of unknown relationships, however, is difficult using either optimisation or strength methods.

Here we present a linear interpolation technique to quantify a likely dominance relationship between unknown dyads. Our method is simple, quick, and pragmatic and offers a first cut at the problem when data are insufficient to provide statistically robust approaches: Bayesian or other probabilistic methods of estimation are problematic when dominance matrices are sparse because of the huge confidence intervals that arise and associated numerical computation issues of optimizing maximum likelihood measures. After execution of de Vries' I&SI approach, if interaction values a_{ij} are not known for dyad $i - j$, (where i and j are the rankings after Step 2) we use 2nd and higher order relationships of the individuals in the unknown dyad to derive the probable relationship between the individuals. In **Step 3** of our method, unknown elements are assigned interpolated values

$$\hat{a}_{ij} = 0.5 - (i - j)/2n \quad (6)$$

where it is clear that $\hat{a}_{ij} + \hat{a}_{ji} = 1$ (the caret denotes that this is an interpolated element of the matrix A rather than a value obtained from data). Our interpolation is based on the assumption that the greater the separation in ranks between two individuals with an unknown relationship the more likely the higher-ranking individual is to dominate the lower ranking individual (Bradley & Terry, 1952; Crow, 1990), with values ranging from almost one if i is the top individual and j is the bottom individual to almost zero for the reverse situation (see the discussion of our assumptions below). After using equation 6 to fill in all the missing values (unknown relationships) in the interaction matrix A (Figure 1-B), **Step 4** of our method calculates each individual's relative dominance strength as described next (Figure 1-C).

Calculating dominance strengths

Strength methods are used to obtain a unique rank order for the individuals in a hierarchical system, modulo those individuals that have identical interaction histories. This uniqueness is an attractive feature for biologists studying the relationship between dominance and other biological measures in a population. In Step 4, we apply a strength method to our most current version of A (which in Steps 1-3 has been minimized with respect to inconsistencies and out filled using the interpolating equation 6). In this step all known dominant-subordinate dyads and the optimised rank order of circular polyads are conserved. These are conditions that are not met when using

	A	B	C	D	E
A	*	1	1	0.8	0.8
B	0	*	0.6	1	1
C	0	0.4	*	1	0.7
D	0.3	0	0	*	1
E	0.2	0	0.3	0	*

Figure 2-A. Our dominance strength metric δ_i subtracts the sum of i -th row elements (wins) from the sum of i -th column elements (losses) for individual i . The dominance strength for individual B = 1.16 (= 2.58 – 1.42) and C = 0.18 (= 2.09 – 1.91).

	A	B	C	D	E
A	*	1	1	0.8	0.8
B	0	*	0.6	1	1
C	0	0.4	*	1	0.7
D	0.3	0	0	*	1
E	0.2	0	0.3	0	*

Figure 2-B. The half-matrix metric may be used when the biology of a study organism affects the distribution of interactions, where an individual is prone to interact more frequently with individuals near their rank order. The dominance strength calculated from interactions below the individual in the rank order for individual B = 2.16 (= 0.58 + 1 + 1 – 0.42 – 0 – 0) and C = 1.34 (= 1 + 0.67 – 0 – 0.33).

strength metrics alone (Figure 1-D). The resulting strength metric enables most unknown relationships to be resolved and can be used for parametric analysis.

In contrast to other strength methods, which explicitly incorporate a contestant's second order relations (i.e., the relative rank of those defeated or victorious during pairwise interactions) in the algorithm, second order relations are implicitly incorporated in our algorithm through the interpolation procedure, which is dependent on the derived order (Step 3). As a result, the dominance strength of individuals for both possible rank orders for the unknown dyad (when B is above C and when C is above B or $i = j - 1$ and when $i = j + 1$) must be assessed prior to assigning relative ranks. This entails recalculating the likely rank metric for each unknown element after switching the position of the dyad members, since these interpolated matrix element values will change with the position of the dyad member (see equation 6). The individual with the greatest dominance strength when in the higher rank position is designated as the actual dominant individual. Specifically, our dominance strength metric δ_i for individual i is generated simply by subtracting the column sum from the row sum (Figure 2 A): define $a'_{ij} = a_{ij}$ or $a'_{ij} = \hat{a}_{ij}$ depending on whether the element is respectively calculated from data or interpolated using equation 6. Further define 'above'

(superscript a) and ‘below’ (superscript b) the matrix diagonal, and row (ρ) and column (κ) sums:

$$\rho_i^a = \sum_{j=i+1}^n a'_{ij}, \quad \rho_i^b = \sum_{j=1}^{i-1} a'_{ij}, \quad \kappa_i^a = \sum_{j=1}^{i-1} a'_{ji}, \quad \kappa_i^b = \sum_{j=i+1}^n a'_{ji} \quad (7)$$

$$\delta = \delta^a + \delta^b, \quad \text{where } \delta^a = \rho_i^b - \kappa_i^a \text{ and } \delta^b = \rho_i^a - \kappa_i^b. \quad (8)$$

Note that δ linearly scales between $n - 1$ and $1 - n$, thereby representing the full spectrum of possible defeats and wins; δ can be easily scaled between 0 and 1 by $(n + \delta)/2n$. This step in our method allows tied ranks to be incorporated into an individual’s dominance strength (de Vries, 1995). Further, in circumstances where likelihood assessment of dominance ranks results in equal ranking, our method assigns tied rank values (0.5 in the case of a 0 to 1 rank system) to the unknown dyad’s elements.

The application of this metric may be problematic in circumstances where the number of unknown interactions is skewed across individuals. In such cases, an individual’s strength will be artificially boosted when relationships with higher ranked individuals are unknown, resulting from the high proportion of likely ranking metrics in its row score (sum of wins) instead of potential 0 values. Correspondingly, an individual with a larger number of interactions with higher ranks will have 0 scores in its win row and thus have a lower strength score. Such scaling issues are particularly salient when the data have a relatively high proportion of unknown relationships or where the social organization of the study organism results in naturally skewed interaction matrices (e.g., in species which tend to interact only with close ranking individuals). Under such conditions, implementation of our method using δ may not offer an acceptable solution to the relative rankings of unknown dyads. Rather, an individual’s rank strength may be better quantified by δ^b (i.e., its interactions with only those below it in the rank order) or by δ^a (i.e., its interactions with only those above it in the rank order) (equations 9; Figure 2-B). Since this half-matrix approach only includes ‘wins’ or ‘losses’ but not both, circular interaction components are excluded from dominance strength calculations. Thus the half-matrix approach should only be employed in highly skewed data sets in which no circular relationships are evident. Differences in the rank order derived by δ^a , δ^b , and δ can be used to identify dyadic relationships subject to skew and those areas in the

dominance matrix where dominance relationships are sufficiently complicated that the accuracy and applicability of assigning discrete rank values to individuals is limited. In systems where highly skewed relationships are recognized and these metrics differ substantially, the use of ordinal ranks or rank categories is recommended since strength metrics will be biased. The full-matrix method should be employed for all other scenarios.

Minimizing the sum of elements below the matrix diagonal

The resolution of the relative ranks of unknown dyads in Step 4 offers a unique rank order to the initial sequence of individuals derived from Steps 1 and 2. As the number of unknown dyads increase, however, the possibility of multiple, equal solutions to Steps 1 and 2 increases. To resolve the 'best' rank order for the systems interaction matrix, **Step 5** of our method solves the minimum sum of the matrix elements below the matrix diagonal (equation 2), ensuring a unique solution for dominance matrices regardless of the number of inconsistencies. Thus the preferred, unique rank order in a system with unknown relationships is one containing the minimum number of inconsistencies (Step 1), the minimum strength of those inconsistencies (Step 2), all unknown relationships ordered according to their dominance strength values (Step 3 and 4), and finally the minimum sum of elements below the matrix diagonal (Step 5).

Assumptions

Few assumptions are required to implement the I&SI method (de Vries, 1998) and, hence, our method which uses the I&SI for the initial sorting procedure. As mentioned previously, the primary assumption of our interpolation function is the greater the separation in ranks between two individuals with an unknown relationship, the more likely it is that the higher ranking individuals dominates the lower ranking individual. By implementing a linear interpolation (equation 6) to out-fill unknown elements in our matrix, we pragmatically assume the differentiation in ranks among individuals is simply linear and guided purely by the dimension of the data set (which is known and hence there are no additional parameters to be estimated). Non-linear interpolation equations, such as sigmoidal or hyperbolic sine, may offer better approximation of rank differentiation between higher and lower individuals in some systems where additional information demonstrates such a

characteristic of dominance relationships (e.g., where strength metrics indicate non-linear differentiation among individuals as if rank relationships are skewed).

The other major assumption, implicit in the first assumption, is that the study system has a near linear dominance hierarchy. Thus, we suggest two measures of linearity be assessed prior to employment of our method. First, statistical significance of linearity for the dominance matrix should be tested using a modified Landau's linearity index h (Landau, 1951) derived from randomization tests (de Vries, 1995). Second, the degree of symmetry in dyadic agonistic interactions (with more than a single agonistic interaction observed) should be assessed using the directional consistency (DC) index (Noë et al., 1980; van Hooff & Wensing, 1987), whereby the number of times interactions occur in the higher frequency direction (H) is subtracted from the number of time agonistic interactions occur in the lower frequency direction (L) and then divided by the total number of interactions: $DC = (H - L)/(H + L)$. In social systems where linear hierarchies cannot be assumed, our method is not appropriate and different techniques such as controlled experiments are needed to assess the context specific relationship of dominance outcomes between individuals (e.g., territoriality and dominance in resident versus non-resident birds). Additionally, the individuals composing the data matrix to be analysed must be from a single, intermixing population in which all individuals can potentially interact. Results obtained for systems that fail to meet the assumptions of near-linearity and intermixing could well turn out to be spurious.

The degree of independence in units of observation need be considered prior to the implementation of our method. Some optimisation methods avoid the assumption of independence in dominance encounters, a condition not often met in animal systems (Kramer & Schmidhammer, 1992), by making the dyad the observational unit of analysis and assigning the rounded 1 or 0 value to dyad members with variable numbers of wins and losses. Unfortunately, this procedure may under-represent the available dominance information if numerous interactions with variable results have been observed. Probability values for the elements a_{ij} may be preferred over simplified 1-0 values in studies that employ strength metrics in parametric analyses, since results using strength metrics are more accurate when employing probabilities. The illustrative example presented here (Figure 1) demonstrates the versatility in data structure that our method can use, depending on the user's assumptions

regarding the independence of the observational unit. The employment of probabilities instead of rounded values is simply implemented by assigning dominance to the i^{th} individual in the flipping algorithm of steps 1 and 2 whenever $a_{ij} > a_{ji}$. In contrast with the assumptions of a Bradley-Terry model (Boyd & Silk, 1983), where the outcome of a dominance encounter is considered the analogue to a comparison between two objects that comply with a Bernoulli trial sampling scheme, our assumption of linearity implies that random observations of interactions between individuals will be skewed as a result of rank differentiation. As such, even few or single observations of dyadic interactions offer important insight and are useful in resolving dominance hierarchy orders. In systems containing high numbers of dominance interactions with multiple observations of dyad interactions, statistically robust methods are developed to distinguish dominance ranks (Boyd & Silk, 1983; Heinze & Schemper, 2002). It is important to note that the method presented here is designed for systems in which collection of numerous dominance interactions per dyad, generally requiring at least five observations per dyad for adequate statistical power (de Vries, 1998), is not feasible.

Analysis of simulated data

In order to test the accuracy of our method, we created two complete dominance matrices containing 20 individuals where the first had a completely linear dominance hierarchy and the second contained a non-linear dominance hierarchy with 5% probability of circular polyads imposed across all dyads (Figure 1 A is the non-linear matrix). The probability of circular polyads was set to 5% because this is a rather high proportion of circular dyads for a near-linear dominance hierarchy. The most likely rank order was found for the complete matrices by minimizing the number of inconsistencies (I) and sum of inconsistencies (S_I). We then randomly removed 20%, 40%, 60% and 80% of elements from both the linear and non-linear dominance matrix (Figure 1 B shows 60% removal of elements in the non-linear matrix). This resulted in matrices which differed in the number and, correspondingly, structure of inconsistencies. Our dominance rank algorithm was run on each of the resulting matrices containing unknown dyads. This process was repeated 10 times on each category of matrix, e.g., 60% removal of the non-linear matrix, in order to get an estimate of the accuracy of the algorithm's

performance with the different data regimes. The relative difference in position between the known rank and the solved rank for each individual in each dominance matrix category was calculated ($n = 200$ per category), from which the average and standard deviation of differences across all individuals was determined.

Not surprisingly, our technique ranked individuals from linear matrices more accurately across all categories of data removal than individuals from non-linear matrices (Table 1). Generally, the algorithm performed well on both linear and non-linear data sets when at least 60% of dyads were known, with the average difference in ranks being less than 1 and the maximum difference in rank being 6 among the permuted 40% removed non-linear matrices (Table 1). The accuracy of rankings in matrices with 60% and 80% removal of data decreased markedly. However, the average (\pm SD) difference between the solved rank and known rank was still only 3.03 ± 2.59 positions for the non-linear case with 80% missing data, and no constraints on the minimum number of interactions per individual (Table 1). Such resolution is potentially suitable for categorizing individuals in course rank-order groups like high, medium, and low ranking. The maximum and 75th quartile differences in an individual's ranking, respectively equal to 13 and 4 positions, demonstrate that large errors may occasionally arise when data is sparse (Table 1). The use of rank classes instead of individual ranks is therefore more conservative and will reduce ranking errors.

This lowered performance in the sparser data sets is due in part to the increased number of individuals for which no or little dyad interaction data are available. In reality, individuals with no known relationships should be excluded from analyses. By implementing a simple constraint whereby each individual in the matrix is required to have a minimum of 3 known relationships (the maximum possible in a 20×20 matrix with 80% of dyads unknown), the performance of our method improves (Table 1). Particularly relevant to studies containing few known relationships is our result that the maximum difference in rank order declined by around a quarter when a minimum constraint of three dyads per individual was implemented for the data set containing 5% circular relationships. The bottom line appears to be that, for a group of 20 individuals, the relationships among more than half the dyads need to be known to ensure the maximum placement error is not more than 3 ranks.

Table 1. The accuracy of the interpolation method is tested on simulated dominance matrices containing 20 individuals with 20%, 40%, 60%, and 80% of data missing and matrices containing 10 and 50 individuals with 20% and 40% of data missing. Results from runs on matrices containing strictly linear and 5% circular relationships are compared.

	Average difference in rank position	SD	Maximum difference in rank position	75 th percentile
Linear				
Matrix $n = 20$				
20%	0.28	0.49	2	1
40%	0.55	0.66	3	1
60% no min. ¹	1.19	1.16	8	2
60% min. 3 ²	1.16	1.21	7	2
80% no min. ¹	2.28	2.01	9	3
80% min. 3 ²	2.10	1.97	8	3
5% Circularity				
Matrix $n = 20$				
20%	0.42	0.79	3	1
40%	0.89	1.19	7	1
60% no min. ¹	1.74	1.74	12	2
60% min. 3 ²	1.47	1.58	9	2
80% no min. ¹	3.03	2.59	13	4
80% min. 3 ²	2.61	2.12	10	4
Matrix $n = 10$				
20%	0.24	0.62	3	0
40%	0.62	0.85	3	1
Matrix $n = 50$				
20%	0.52	0.75	4	1
40%	1.28	1.39	8	2

¹ No constraint was implemented on the minimum number of relationships per individual.

² The minimum number of relationships per individual was constrained to 3.

We also explored the impact of the number of individuals in the system by running our method on a linear system comprising 10 and 50 individuals with 20% and 40% removal of data. The ranking performance of matrices comprising 10 individuals demonstrated improvements over those composed of 20 individuals (20% removed averaging 0.24 ± 0.62 and 40% removed averaging 0.62 ± 0.85), though the improvement in accuracy was slightly less

Table 2. Comparison of de Vries I&SI (1998) to our method with linear interpolation of unknown relationships

Proportion unknown	Method	Average difference	SD	Maximum	75 th percentile	*Unique solution
20%	Interpolation	0.40	0.76	3	1	100%
	I&SI	0.41	0.76	3	1	80%
40%	Interpolation	1.05	1.26	7	1	100%
	I&SI	1.06	1.24	7	2	50%
60%	Interpolation	1.38	1.56	9	2	100%
	I&SI	1.54	1.56	10	2	30%
80%	Interpolation	2.46	2.29	11	3	100%
	I&SI	2.66	2.47	11	4	10%

* The proportion of runs in which a unique solution for the dominance rank was found

than the proportional decrease in the number of individuals simulated. Interestingly, the ranking performance in the larger data matrices (20% removed averaging 0.52 ± 0.75 and 40% removed averaging 1.28 ± 1.39 for matrices composed of 50 individuals) was similar to that of the matrices composed of 20 individuals, indicating error rates do not increase in direct proportion to the number of individuals in a system for larger data matrices. The maximum displacement in rank also remained similar to the results found for the smaller matrix (*viz.* 4 on the 20% and 8 on the 40% removed in contrast to 3 and 7 in the 20 individual system). These results suggest that for sparse data sets, useful levels of resolution are going to be obtained, even for large systems.

A separate set of simulations were run to compare the accuracy of de Vries (1998) I&SI method, as originally published, to the interpolation method presented here. The difference between these two methods lies in how they resolve the relative rankings of adjacent pairs of individuals with unknown relationships. The sum of wins and the sum of losses for each of the unknown individuals are compared in the I&SI method, and dominance is assigned to the individual with the greater value. Unique dominance matrices with approximately 5% circularities (ranging from 3.2% to 9.0%) were created and a proportion of matrix elements removed (20%, 40%, 60% and 80%) as described above. The two methods were then run on the same matrix so direct comparison of their accuracy could be conducted. Simulation results indicate that the two methods perform similarly for matrices with 20-40% unknown

relationships. As the proportion of unknown relationships increases, however, the interpolation method appears to rank individuals with greater accuracy (Table 2). Interestingly, across all four of the matrix categories, some matrices were more accurately solved by I&SI and other matrices by our interpolation method, though our interpolation method generally did better than the I&SI method. The I&SI method did not always find unique solutions to the randomly generated matrices across all categories (Table 2), with this phenomenon increasing as the proportion of unknown relationships increased. Therefore, comparison to the interpolation method was arbitrarily selected as the solution from the 400th iteration of de Vries I&SI method.

Discussion

Our likely dominance ranking method combines the strengths of two existing and complementary approaches to ranking hierarchies of individuals. First, our method uses an optimisation paradigm to find the rank order nearest to a linear hierarchy, following the sorting procedure of the I&SI method (de Vries, 1998), and then employs a strength paradigm to ensure the final order is a unique solution. As with the I&SI method, this offers an advantage over other methods by not weakening or undermining the results of known interactions.

Simulated data were used to compare the I&SI method (de Vries, 1998) to our interpolation function. The simulations demonstrated several advantages of the interpolation method over the I&SI method. First, the interpolation method derives a unique solution to any dominance matrix regardless of the number and structure of unknown relationships. The I&SI method failed to find unique solutions in all categories of matrices tested, even occasionally in those with 20% unknown relationships (Table 2). Second, the accuracy of the interpolation method and the I&SI are similar when the majority of relationships are known; however, as the proportions of unknown dyads increases the interpolation method is more accurate than the I&SI method. Third, our technique can use multiple types of data to resolve the rank order: the observational unit can be designated at the categorical 1-0 (win-lose) dyad level or the graded, probabilistic value arising from scoring multiple interactions of the same pair allowing all information on the dominance relationships between individuals to be incorporated into the assessment of

rank. Finally, the interpolation method derives a strength metric for each individual which can be used in subsequent analyses – a characteristic similar to the often used David's Score (David, 1987; Gammell et al., 2003).

The calculation of continuous dominance strength metrics, not offered by the I&SI method, is of relevance to many studies of dominance. The benefits of such a metric include resolving the rank order among closely ranked individuals, identifying dyads with truly tied ranks, and providing a continuous variable with which to compare behavioural correlates. Employing only the strength metric from our method in parametric analyses should be conducted with caution because circular polyads can result in a non-linear relation between a strength-based order and the order derived using an optimisation method (Fig. 1 D). The relative strengths of linearly ranked portions of the interaction matrix, however, can be used in parametric comparisons. The incorporation of our strength metric, the sum of elements in an individual's row (wins) minus the sum of elements in an individual's column (losses), to solve unresolved components of the rank order consistent with the I&SI result offers a novel combination of the two conventions of ranking techniques thereby drawing upon the advantages of both approaches.

The accuracy of our method was tested on simulated dominance matrices with different proportions of unknown and circular relationships. Solved rank orders were compared with known rank orders, providing a quantified assessment of algorithm performance (Table 1). The algorithm provided useful information on relative ranks even when large proportions of the matrix elements were empty. We found that relatively simple rules, such as setting limits to the number of unknown dyads per individual, can greatly increase the accuracy of rank orders, particularly as the size of the system increases. Furthermore, our results indicate that our approach can be used effectively as an aid in determining key unknown dyads whose relationship, when known, will disproportionately improve the solved rank order. Thus, in addition to solving the best rank order, our method may be usefully employed to explore the structure of behavioural data and improve data collection regimes. While this algorithm will solve the most likely linear dominance order for the available data, data quality and quantity and skew in inter-individual dominance relationships can strongly impact results. Dominance strength metrics are particularly subject to error from skewed relationships. When data quantity is low and skew occurs, results should be applied to analyses with caution and we recommend using dominance categories based on the error rates from our simulated data rather than individual based metrics.

Acknowledgements and note on software

Funding for this work was provided by the National Science Foundation, the Department of Environmental Science, Policy, and Management at UC Berkeley, and a James S. McDonnell Foundation 21st Century Science Initiative Award to WMG. The comments from anonymous reviewers, Dr. William E. Wagner Jr. and Dr. Charles Snowdon greatly helped to improve this manuscript.

As a courtesy, our algorithm is available at: <http://nature.berkeley.edu/getz/dominance/ODO.pl>. A time cap on the server at this site, limits its application to relatively small problems. The code can be provided on request. We take no responsibility for the accuracy of this code or results obtained through use of our website.

References

- Albers, P.C.H. & de Vries, H. (2001). Elo-rating as a tool in the sequential estimation of dominance strengths. — *Anim. Behav.* 61: 489-495.
- Appleby, M.C. (1983). The probability of linearity in hierarchies. — *Anim. Behav.* 31: 600-608.
- Archie, E.A., Morrison, T.A., Foley, C.A.H., Moss, C.J. & Alberts, S.C. (2006). Dominance rank relationships among wild female African elephants, *Loxodonta africana*. — *Anim. Behav.* 71: 117-127.
- Boyd, R. & Silk, J.B. (1983). A method for assigning cardinal dominance ranks. — *Anim. Behav.* 31: 45-58.
- Boydston, E.E., Kapheim, K.M., Szykman, M. & Holekamp, K.E. (2003). Individual variation in space use by female spotted hyenas. — *J. Mammal.* 84: 1006-1018.
- Bradley, R.A. & Terry, M.E. (1952). Rank analysis of incomplete block designs. 1. The method of paired comparisons. — *Biometrika* 39: 324-345.
- Clutton-Brock, T.H., Albon, S.D., Gibson, R.M. & Guinness, F.E. (1979). Logical stag – Adaptive aspects of fighting in red deer (*Cervus elaphus* L.). — *Anim. Behav.* 27: 211-225.
- Côté, S.D. & Festa-Bianchet, M. (2001). Reproductive success in female mountain goats: the influence of age and social rank. — *Anim. Behav.* 62: 173-181.
- Crow, E.L. (1990). Ranking paired contestants. — *Comm. Stat. Simul. Computat.* 19: 749-769.
- David, H.A. (1987). Ranking from unbalanced paired-comparison data. — *Biometrika* 74: 432-436.
- 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. — *Anim. Behav.* 66: 601-605.
- Gende, S.M. & Quinn, T.P. (2004). [The relative importance of prey density and social dominance in determining energy intake by bears feeding on Pacific salmon.](#) — *Can. J. Zool.* 82: 75-85.
- Hansen, E.A. & Closs, G.P. (2005). [Diel activity and home range size in relation to food supply in a drift-feeding stream fish.](#) — *Behav. Ecol.* 16: 640-648.
- Heinze, G. & Schemper, M. (2002). A solution to the problem of separation in logistic regression. — *Stat. Medic.* 21: 2409-2419.

- Holand, O., Weladji, R.B., Gjostein, H., Kumpula, J., Smith, M.E., Nieminen, M. & Roed, K.H. (2004). Reproductive effort in relation to maternal social rank in reindeer (*Rangifer tarandus*). — *Behav. Ecol. Sociobiol.* 57: 69-76.
- Holekamp, K.E., Smale, L. & Szykman, M. (1996). Rank and reproduction in the female spotted hyaena. — *J. Reprod. Fert.* 108: 229-237.
- von Holst, D., Hutzelmeyer, H., Kaetzke, P., Khaschei, M., Rodel, H.G. & Schrutka, H. (2002). Social rank, fecundity and lifetime reproductive success in wild European rabbits (*Oryctolagus cuniculus*). — *Behav. Ecol. Sociobiol.* 51: 245-254.
- van Hooff, J.A.R.A.M. & Wensing, J.A.B. (1987). Dominance and its behavioral measures in a captive wolf pack. — In: *Man and wolf* (Frank, H.W., ed.). Junk Publishers, Dordrecht, p. 219-252.
- Johnson, D.F., Modahl, K.B. & Eaton, G.G. (1982). Dominance status of adult male Japanese macaques – Relationship to female dominance status, male mating-behavior, seasonal-changes, and developmental-changes. — *Anim. Behav.* 30: 383-392.
- Kramer, M. & Schmidhammer, J. (1992). The Chi-squared statistic in ethology – Use and misuse. — *Anim. Behav.* 44: 833-841.
- Krebs, J.R. & Davies, N.B. (1987). *An introduction to behavioural ecology*. — Blackwell Scientific, Boston.
- Landau, H.G. (1951). On dominance relations and the structure of animal societies: I. Effect of inherent characteristics. — *Bull. Math. Biophys.* 13: 1-19.
- Langbein, J. & Puppe, B. (2004). Analysing dominance relationships by sociometric methods – a plea for a more standardised and precise approach in farm animals. — *Appl. Anim. Behav. Sci.* 87: 293-315.
- Lee, P.C. (1987). Allomothering among African elephants. — *Anim. Behav.* 35: 278-291.
- Martin, P. & Bateson, P. (1993). *Measuring behaviour. An introductory guide*. — Cambridge University Press, Cambridge.
- McMahan, C.A. & Morris, M.D. (1984). Application of maximum likelihood paired comparison ranking to estimation of a linear dominance hierarchy in animal societies. — *Anim. Behav.* 32: 374-387.
- Mitchell, J. (2005). Queue selection and switching by false clown anemonefish, *Amphiprion ocellaris*. — *Anim. Behav.* 69: 643-652.
- Noë, R., de Waal, F.B.M. & van Hooff, J.A.R.A.M. (1980). Types of dominance in a chimpanzee colony. — *Folia Primatol.* 34: 90-110.
- Otter, K., Ratcliffe, L., Michaud, D. & Boag, P.T. (1998). Do female black-capped chickadees prefer high-ranking males as extra-pair partners? — *Behav. Ecol. Sociobiol.* 43: 25-36.
- Roberts, J.M. (1990). Modeling hierarchy – Transitivity and the linear ordering problem. — *J. Math. Sociol.* 16: 77-87.
- Roden, C., Vervaecke, H. & Van Elsacker, L. (2005). Dominance, age and weight in American bison males (*Bison bison*) during non-rut in semi-natural conditions. — *Appl. Anim. Behav. Sci.* 92: 169-177.
- Sapolsky, R.M. (2005). The influence of social hierarchy on primate health. — *Science* 308: 648-652.
- Say, L. & Pontier, D. (2004). Spacing pattern in a social group of stray cats: effects on male reproductive success. — *Anim. Behav.* 68: 175-180.
- Tarvin, K.A. & Woolfenden, G.E. (1997). Patterns of dominance and aggressive behavior in blue jays at a feeder. — *Condor* 99: 434-444.

- de Vries, H. (1995). An improved test of linearity in dominance hierarchies containing unknown or tied relationships. — *Anim. Behav.* 50: 1375-1389.
- de Vries, H. (1998). Finding a dominance order most consistent with a linear hierarchy: a new procedure and review. — *Anim. Behav.* 55: 827-843.
- de Vries, H. & Appleby, M.C. (2000). Finding an appropriate order for a hierarchy: a comparison of the I&SI and the BBS methods. — *Anim. Behav.* 59: 239-245.
-