The construction of dominance order: comparing performance of five methods using an individual-based model

C.K. Hemelrijk[1,2,3)], **J. Wantia**[2)] **& L. Gygax**[2,4)]

([1] Theoretical Biology, Centre for Ecological and Evolutionary Studies, Biological Centre, University of Groningen, Kerklaan 30, 9751 NN Haren, the Netherlands; [2] Department of Informatics, University of Zurich, Andreasstrasse 15, 8050 Zurich, Switzerland)

(Accepted: 28 June 2005)

Summary

In studies of animal behaviour investigators correlate dominance with all kinds of behavioural variables, such as reproductive success and foraging success. Many methods are used to produce a dominance hierarchy from a matrix reflecting the frequency of winning dominance interactions. These different methods produce different hierarchies. However, it is difficult to decide which ranking method is best. In this paper, we offer a new procedure for this decision: we use an individual-based model, called DomWorld, as a test-environment. We choose this model, because it provides access to both the internal dominance values of artificial agents (which reflects their fighting power) and the matrix of winning and losing among them and, in addition, because its behavioural rules are biologically inspired and its group-level patterns resemble those of real primates. We compare statistically the dominance hierarchy based on the internal dominance values of the artificial agents with the dominance hierarchy produced by ranking individuals by (a) their total frequency of winning, (b) their average dominance index, (c) a refined dominance index, the David's score, (d) the number of subordinates each individual has and (e) a ranking method based on maximizing the linear order of the hierarchy. Because dominance hierarchies may differ depending on group size, type of society, and the interval of study, we compare these ranking methods for these conditions. We study complete samples as well as samples randomly chosen to resemble the limitations of observing real animals. It appears that two methods of medium complexity (the average dominance index and David's score) lead to hierarchical orders that come closest to the hierarchy based on internal dominance values of the agents. We advocate usage of the average dominance index, because of its computational simplicity.

[3)] Corresponding author's e-mail address: c.k.hemelrijk@rug.nl

[4)] Present address: Centre for proper housing of ruminants and pigs, Swiss Federal Veterinary Office, Agroscope FAT Taenikon, 8356 Ettenhausen, Switzerland

© Koninklijke Brill NV, Leiden, 2005

Keywords: dominance order, ranking method, agent-based model, statistical method, aggression

Introduction

Dominance hierarchy is a central theme in the study of animal behaviour. It is correlated with many behavioural variables, such as social activities, success at foraging and reproduction. The dominance hierarchy is thus a central feature in many studies of animal behaviour. Yet, in observing animal behaviour, several hierarchical orderings of individuals are usually possible and it is unclear which is the more realistic one. Whereas much has been written on the question how to measure dominance rank, no unanimous conclusion as to what method should be preferred has been reached. It is difficult to decide this, because we have no independent information about real internal power or dominance values, that influence the winning tendencies; besides, winning tendencies change in time and depend on the social situation (Chase et al., 2003). In the present paper, we compare statistically the efficiency of several ranking methods in a model consisting of artificial agents that group and compete. This method is efficient, since here we have unique access both to the internal power or dominance values of the agents (that represent their winning tendencies) and to the frequency of their victories and defeats in a model that shows large resemblances to behaviour of real animals (Hemelrijk, 1999).

The evaluation method we use is new, because we determine the degree with which the outcome of a ranking method matches the real dominance hierarchy among agents based on their internal dominance values or power. We test the agreement of the internal 'true' dominance values and the dominance values calculated according to the different methods. The model, called DomWorld, is based on self-organisation that arises from the effects of self-reinforcement and spatial structure (Hemelrijk, 1999, 2000, 2002). Here, agents merely group and perform competitive interactions in which the effects of winning and losing are self-reinforcing (winning/losing an interaction increases the chance of doing so again). These self-reinforcing effects have been observed in many animal species, ranging from invertebrates such as insects (Theraulaz et al., 1992), crustaceans (Karavanich & Atema, 1998) via fish (Hsu & Wolf, 1999) and birds (Drummond & Canales, 1998) to mammals, such as deer (Thouless & Guinness, 1986), primates (Barchas

& Mendoza, 1984) and humans (Mazur, 1985). It has appeared that merely by increasing the intensity of aggression in this model the society switches from an egalitarian society (with a weakly developed dominance hierarchy) to a despotic one in which the hierarchy is steep (Hemelrijk, 1999, 2000). This arises via a complex feedback between the developing hierarchy and the spatial structure in which dominants are in the centre and subordinates at the periphery. This feedback develops only at a high intensity of aggression, not at a low one. At a high intensity several phenomena originate as side-effects: e.g., grouping becomes looser, aggression less frequent and less symmetrical, the hierarchy becomes steeper and social behaviour more dominance-oriented. All these differences between artificial societies with a low intensity and a high one, resemble those between egalitarian and despotic types of societies in macaques (see, e.g., Caldecott, 1986; de Waal, 1989; de Waal & Luttrell, 1989; Thierry, 1990, 2000; Thierry et al., 1990a, b).

Because of the resemblance of artificial societies in the model to those of real systems, such as primates (Hemelrijk, 1999) and because it is better to have an explicit model than arbitrary assumptions, we use this model as a device to think about what ranking method of dominance order is best. Because in some species, but not in others, individuals memorise experiences with others individually, in the discussion (data not shown) we compare results of both cases, direct dominance perception and experience-based dominance estimation (Hemelrijk, 2000). In order to derive the dominance order, many methods have been used (for a review, see de Vries, 1998). These are nearly always based on a matrix of the summed frequency of the victories and defeats (or status signals, yielding, supplants or initiation of aggression) in dyadic dominance encounters over a certain period (for an exception, see Albers & de Vries, 2001). Simple methods of determination of the dominance order are based on the frequency of aggression won and the percentage of aggression directed against others (Zumpe & Michael, 1986) or on reshuffling data in the matrix until the maximum number of zeros is found below the diagonal (Roberts, 1990). However, there are also more complex methods, such as an iterative procedure to calculate cardinal ranks by Boyd & Silk (1983). Such methods are, however, based on assumptions that are clearly refuted. For example it is often assumed that the probability of encounter among all group-members is the same. Further, it is assumed that the probability of victory for each individual of a pair is unchangeable (thus ignoring the winner/loser effect).

Another complex method is based on ordering individuals in a maximal linear order (e.g., see de Vries, 1998; de Vries & Appleby, 2000). First, linearity is tested for, subsequently inconsistencies of a full linear rank are minimised. This is done by repeatedly re-ordering a set of individuals into a (near-) linear dominance hierarchy, while minimizing both the number and strength of inconsistencies (also called 'triangles', such as A>B, B>C and C>A). After comparing it to a large number of other methods de Vries concluded that this method is to be preferred. Although experiments by Chase et al. (e.g., see 2002) have repeatedly shown that there is a significant tendency for hierarchies to become linear, de Vries' method ignores that the change in dominance that accompanies victory and defeat differs in degree depending on the identity of both opponents and the moment at which the interaction takes place. Therefore, this method may be misleading.

Thus, we have assumed that the dominance order may be derived from a simpler ranking method just as well as from maximisation of linearity. In order to investigate this, we compare this method of maximised linear ordering to a number of simpler ones that do not assume linearity of hierarchy. We order individuals in a hierarchy by (a) their total frequency of attack/winning per individual (Chase et al., 1994), (b) by their 'average dominance index' (i.e., the average percentage with which an individual wins in interactions with each of its group members), (c) a weighted version of the average dominance index, the David's score, whereby the relative success is weighted by the power of the opponent (Gammell et al., 2003), (d) by the number of individuals against whom they win more often than lose from, i.e., that are subordinate to them and (e) the method based on the maximum linear order of the hierarchy by de Vries (1998). To study the performance of these ranking methods for different kinds of hierarchies as may arise depending on group size, type of societies, and time interval of the study we compare them between these conditions. Further, we compare them for random samples of the observed interactions and for the complete set of observations of all dominance interactions.

Methods

DomWorld (Dominance World)

Here, a brief summary of DomWorld must suffice (for a more complete description see Hemelrijk, 1999, 2000, 2002). This mechanistic model is based

on two essential aspects of social life. It consists of a homogeneous virtual world inhabited by agents provided with no more than two tendencies: (1) to group and (2) to perform dominance interactions. Why agents actually group (whether to avoid predators or because resources are clumped) is not specified and irrelevant to the model. The same holds for dominance interactions. They reflect competition for resources (such as food and mates), but these resources are not explicitly specified.

Whenever an individual does not see another agent close by (within its personal space, PersSpace), grouping rules come into effect. The agent starts looking for others at greater and greater distances (NearView = 24 and MaxView = 50 units). If, even then, no one else is in sight, it turns over a SearchAngle (of 90 degrees) in order to rejoin its group. In this way individuals tend to remain in a group.

If, however, an agent spots another agent close by, within its personal space (PersSpace = 4 units), a dominance interaction may take place.

The likelihood that an agent begins an aggressive interaction depends on the risks involved: it increases with its chance to defeat its opponent. This chance depends on the relative capacities of both agents to defeat each other, i.e., the relative dominance values. This is the so-called 'risk-sensitive attack strategy' (Hemelrijk, 2000).

If a dominance interaction actually takes place, agents i and j observe each other's capacity of winning, i.e., their dominance values $\mathrm{Dom_i}$ and $\mathrm{Dom_j}$. The probability of winning is greater for whoever is higher in rank, and this is proportional to the Dom-value of agent i relative to that of its opponent j (see (1)). To allow for dominance reversals, a stochastic effect is introduced, so that if the relative dominance value of an interacting agent is greater than a random number (drawn from a uniform distribution), then agent i wins ($w_i = 1$), else it loses ($w_i = 0$):

$$w_i = \begin{cases} 1 & \dfrac{\mathrm{DOM}_i}{\mathrm{DOM}_i + \mathrm{DOM}_j} > \mathrm{RND}(0,1) \\ 0 & \text{else} \end{cases} \tag{1}$$

To reflect the self-reinforcing effects of victory and defeat, dominance values are updated by increasing the dominance value of the winner and decreasing that of the loser by the same amount:

$$\mathrm{DOM}_i := \mathrm{DOM}_i + \left(w_i - \frac{\mathrm{DOM}_i}{\mathrm{DOM}_i + \mathrm{DOM}_j} \right) * \mathrm{STEPDOM} \tag{2}$$

$$\mathrm{DOM}_j := \mathrm{DOM}_j - \left(\mathrm{w}_i - \frac{\mathrm{DOM}_i}{\mathrm{DOM}_i + \mathrm{DOM}_j}\right) * \mathrm{STEPDOM}$$

This positive feedback is 'damped' because a victory of the higher ranking opponent reinforces its relative Dom-value only slightly, whereas (unexpected) success of the lower ranking agent increases its relative dominance value by a greater change, thus equalising their dominance values. This conforms to detailed behavioural studies of bumble bees by Honk & Hogeweg (1981). To keep Dom-values positive, their minimum value is, arbitrarily, set at 0.01. The change in Dom-values is multiplied by a scaling factor, called StepDom, which varies between 0 and 1 and represents the intensity of aggression (see experiments). High values imply a great change in Dom-value when updating it, and thus indicate that single interactions (e.g., involving biting) may strongly influence the future outcome of conflicts. Conversely, low StepDom-values represent low impact (e.g., threats or slaps).

Winning an interaction includes chasing the opponent over a distance of one unit and then turning randomly 45 degrees to right or left in order to reduce the chance of repeated interactions between the same opponents. The loser responds by fleeing under a small random angle over a predefined FleeingDistance (of 2 units).

Experiments

Groups of agents consist of 4, 8, or 16 individuals that are completely identical at the start (with initial dominance values, initDom of 24). Group sizes reflect those of the adults of one or both sexes in many groups of primate species in captivity or under natural conditions. Further, primates and other mammals usually live in bi-sexual groups. Females often have a smaller body size, weaker muscles and their aggression is less intense than that of males (e.g., see Bernstein & Ehardt, 1985). In artificial groups we implement both sexes by an equal number of agents of two types that differ in their capacity of winning. One type ('females') initially has a lower dominance value than the other (initDom = 16 instead of 32) and also a lower intensity of aggression (10% or 80% of the intensity of aggression of the 'males'). To resemble species with a different intensity of aggression, such as found in primates, we test all settings for a high and a low intensity of aggression (StepDom = 1 and StepDom = 0.1). For all parameter settings see Table 1.

At the start of a run, when individuals have only been activated a few times, the hierarchy is hardly differentiated and there are many dominance reversals (Hemelrijk & Gygax, 2004). Later on, after many activations, the hierarchy is qualitatively more stable (as measured by the Kendall rank correlation between dominance positions of individuals for an interval of two periods, whereby one period consists of the number of individuals * 20 activations). This is the case approximately from period 200 onwards. The run ends at period 260. We study dominance during the supposedly stable phase from period 200 till 260, thus over 60 periods. We may use this interval completely, but because shorter intervals are more stable, we also study shorter intervals around the mid-point of period 230. Thus, we look at an interval of 20 periods ranging from 220 till 240 and one of 10 periods ranging from 225 till 235.

Because in the observation of real animals many events are usually missed, we also build matrices from a random sample of all interaction events next to analysing all dominance interactions that occur. We chose our random sample for the intervals of 60 and 20 periods such that we observe 20 interactions per individual, thus 80 interactions for group size of 4, 160 interactions for group size 8 and 320 for group size 16. For the shortest interval of 10 periods interactions are so few, however, that we only used 10 interactions per individual, thus, 40, 80 and 160 interactions for group sizes 4, 8 and 16 respectively. For all group sizes this amounts to approximately 20% of all interactions if the interval of observation is 60 periods and approximately 50-60% of all interactions for the shorter intervals of 10 and 20 periods.

Per setting, we have conducted 40 replicates. In total, this study is based on (3 group size * 2 StepDom + 4 StepDom * 2 sex difference in StepDom) * 40 = 14 * 40 = 560 runs (see Table 1).

Ranking methods

At each moment, the Dom-values of the agents in DomWorld give the real order of the agents in the hierarchy.

Using the events of aggression-initiation and of victory (which are significantly positively correlated) we evaluate the effectiveness of the different ranking methods to produce a dominance hierarchy. In case of ties, the individuals with the same value get the mean rank of these tied values.

We use the following five methods: we rank individuals by (a) the total frequency of their attacks (or victories), called AttFr, (b) by their average

Table 1. The experimental conditions. For each StepDom value (or combination of two values) 3 intervals (of 60, 20 and 10 periods) are studied, by means of a random sample and a complete analysis. Per condition 40 replicas have been carried out.

#Type	Group size	initD(s)	StepD(s)	Total
One type	4, 8, 16	24	1 or 0.1	3*2*40 = 80
Two types				
4, 4	8	32, 16	1, 0.8 and 1, 0.1 and 0.1, 0.08 and 0.1, 0.01	4*40 = 160
8, 8	16	32, 16	1, 0.8 and 1, 0.1 and 0.1, 0.08 and 0.1, 0.01	4*40 = 160
			Total	560 runs

initD = initDom, StepD = StepDom (i.e., the intensity of aggression).

individual dominance index (called ADI), (c) a refined version, David's score (called DS), (d) the number of individuals against whom they win more often than lose from (called Netto) and (e) we construct a hierarchy by minimising the number and strength of hierarchical inconsistencies, the so-called I&SI method (de Vries & Appleby, 2000).

For the first ranking method (AttFr, see Table 2A), the frequency with which an individual, i, attacks or wins from an opponent, j, x_{ij}, is summed per individual ($\sum_j w_{ij}$) over the phase of the run-time in which the hierarchy is 'stable' (Hemelrijk & Gygax, 2004). A higher frequency is supposed to indicate a higher dominance-rank (Zumpe & Michael, 1986).

Second, the average individual dominance index (ADI, Table 2B) is calculated as follows. The dominance index per pair of individuals, w_{ij}, is the number of times an individual has beaten (or attacked) a certain opponent divided by the total number of fights in which the pair was involved with each other, thus $w_{ij} = \frac{x_{ij}}{x_{ij}+x_{ji}}$. If a pair of individuals were not involved in a fight with each other, it was excluded from the analysis. The average dominance index of an individual is the average of all its dominance indices with all its interaction partners, thus $1/N \sum_j w_{ij}$. A higher value indicates a higher dominance in the group. This index is a simplified version of the dominance index by Zumpe & Michael (1986). Zumpe & Michael calculated the average of the sum of the two percentages of attack and of submission received per pair and they averaged these per individual.

Third, the David's score is a sophisticated dominance index (Gammell et al., 2003). It is calculated as shown in Table 2B. Like in the dominance index, per dyad the ratio of winning (losing) is calculated, w_{ij}, and these ratios are summed per row (called w) and summed per column (called l). Further, as an indication of the power of the victim in each pair, each ratio is weigted by the summed winning ratio of the loser and these products are summed per row-individual (called $w2$). For instance, for individual a, $w2 = [(1.0*1.66)+(1.0*1.55)+(0.89*1.16)+(1.0*0.72)] = 4.98$. A higher value indicates that ego itself wins more often and/or that the winning tendency of egos victims is high. Similarly, ratios are summed per column (called l) and the product with the total winning ratio of each opponent is calculated and summed per column (called $l2$). The David's score is the sum of the measures of winning in the rows minus those in the columns ($DS = w + w2 - l - l2$).

The fourth ranking method (Netto, see Table 2C) implies a count, per individual, of the number of interaction-partners it beats more often than it is defeated by, in other words the number of individuals over which it is dominant, or that are subordinate to it, s_i. In case of ties, nothing is added to the count. Higher values indicate higher dominance positions. Note that Netto does not control for the total number of interactions.

The fifth method (see Table 2A) is the so-called I&SI method by de Vries (1998). Within each pair, the individual that wins more often is considered dominant and cells are reshuffled such as to produce the hierarchy with the lowest number and degree of inconsistencies (i.e., intransitivity and triangles). For instance, in Table 2A initially, there is one inconsistency, I, in the matrix: individual d wins over c more often than the other way around; the strength of this inconsistency, SI, is 1. After swapping individuals c and d, the inconsistency is gone. Our calculation follows the algorithm that is described in the appendix of the paper by de Vries. In DomWorld complete linearity arises because the dominance values are perceived by others without error and everyone has a different history and therefore, a different dominance value. Consequently, we do not need to measure linearity separately as is done in the I&SI method.

Note that we did not include other methods of individual overall success, such as the fighting index of Clutton-Brock et al. (1979), because Gammell et al. (2003) show that the David's score is more appropriate.

In the model we study the correspondence between, on the one hand, the hierarchy derived by these methods from matrices of victory (and of

Table 2. Computational example. The matrices of proportion of winning (B) and of binary dominance (C) are derived from the matrix of winning (attack) frequencies in (A).

A) Attack/win frequency (Attfr) and I&SI

The total frequency of attack / wins, x_{ij}:

	a	**b**	**c**	**d**	**e**	**AttFr** $= \sum x_{ij}$	**Rank$_{\text{AttFr}}$**	**I&SI**	**Rank$_{\text{I\&SI}}$**
a	–	6	9	8	5	28	5	initial I:1	5
b	0	–	4	6	0	10	3	initial SI:1	4
c	0	2	–	4	7	13	4	final I:0	2
d	1	0	5	–	3	9	2	final SI:0	3
e	0	0	2	3	–	5	1		1

B) Average Dominance Index (ADI) and David's score (DS)

The proportion of winning, w_{ij}. DS $= w + w2 - l - l2$:

	a	**b**	**c**	**d**	**e**	**ADI** $= 1/N \sum w_{ij}$	**Rank$_{\text{ADI}}$**	**W**, $\sum_j w_{ij}$	**w2**	**DS**	**Rank$_{\text{DS}}$**
a	–	1	1	0.89	1	0.972	5	3.88	4.98	8.44	5
b	0	–	0.67	1	–	0.555	4	1.66	2.20	1.61	4
c	0	0.33	–	0.44	0.78	0.388	3	1.55	1.64	−2.33	3
d	0.11	0	0.56	–	0.5	0.292	2	1.16	1.66	−3.66	2
e	0	–	0.22	0.5	–	0.240	1	0.72	0.93	−4.05	1
L, $\sum_i w_{ij}$	0.11	1.33	2.44	2.83	2.28						
l2	0.31	0.93	3.08	3.66	3.43						

C) Netto

Subordinates to a row individual, s_{ij}:

	a	**b**	**c**	**d**	**e**	**Netto**, $\sum s_{ij}$	**Ranks**
a	–	1	1	1	1	4	5
b	0	–	1	1	0	2	4
c	0	0	–	0	1	1	2.5
d	0	0	1	–	0	1	2.5
e	0	0	0	0	–	0	1

Ad B)'–': individuals that did not meet (b and e) are excluded from the calculation. w2, l2 are sums of the winning proportions weighted by the opponent. For further explanation, see text.

attack) and on the other, the 'real' hierarchies, by ranking agents by their internal Dom-values. For this we use a Kendall rank correlation between the hierarchical order at the middle of the stable phase (at period 230) and the hierarchical order derived by each of the five ranking methods.

Results

The correlation between dominance hierarchy and hierarchy based on each of the five ranking methods are shown in Figure 1. The overall difference in performance among all five measures is tested by the Friedman test. On the X-axis the five measures are shown and on the Y-axis the value of the Kendall rank correlation between the rankings according to each method and the Dom value of the agent. The higher this value the greater the correspondence between the value of the rank estimated from a matrix of winning frequencies and the rank of the Dom-value of the agent. In the first column, we see results of a sample size of $N = 16$ of a complete sample (Figure 1A, D) and a random part of it (Figure 1B, E) for a high intensity of aggression (StepDom = 1, Figure 1A, B) and a low one (0.1, see Figure 1D, E). In the complete sample (Figure 1A, D), the correlation is strongest between hierarchy based on the internal Dom-values and the average dominance index (ADI) and its refined form David's score (DS). The next highest correlation is when the hierarchy is based on the number of subordinate group-members each individual dominates over (Netto). Table 3 shows that for all conditions tested only few differences between these methods (ADI, DS and Netto) appear to be significant. All significant cases show that the average dominance index and David's score lead to better results than Netto (the number of one's subordinates). The frequency of attack (AttFr) and maximal-linear-ordering method (I&SI) produce hierarchies that correlate in a similar way, but significantly less with the real dominance positions than the other ranking methods (ADI, DS and Netto). Further, correlations between real and inferred hierarchies and differences between ranking methods (as indicated by the significances of the Wilcoxon matched pairs signed ranks test above each figure) are stronger at a high intensity of aggression (Figure 1A) than at a low intensity of aggression (Figure 1D). Besides, correlations are strengthened and differences between ranking methods are greater when the hierarchy is inferred from a complete sample instead of smaller sample randomly drawn from all data (compare Figure 1A vs 1B, and 1D vs 1E). If we evaluate the five ranking methods over a shorter time interval (of 20 and 10 periods, compare among three columns of Figure 1A, D), correlations may become weaker and the difference between ranking methods may diminish somewhat.

For a smaller group size of 8 individuals significant differences are generally similar but fewer (Table 3A). In groups of four individuals significant

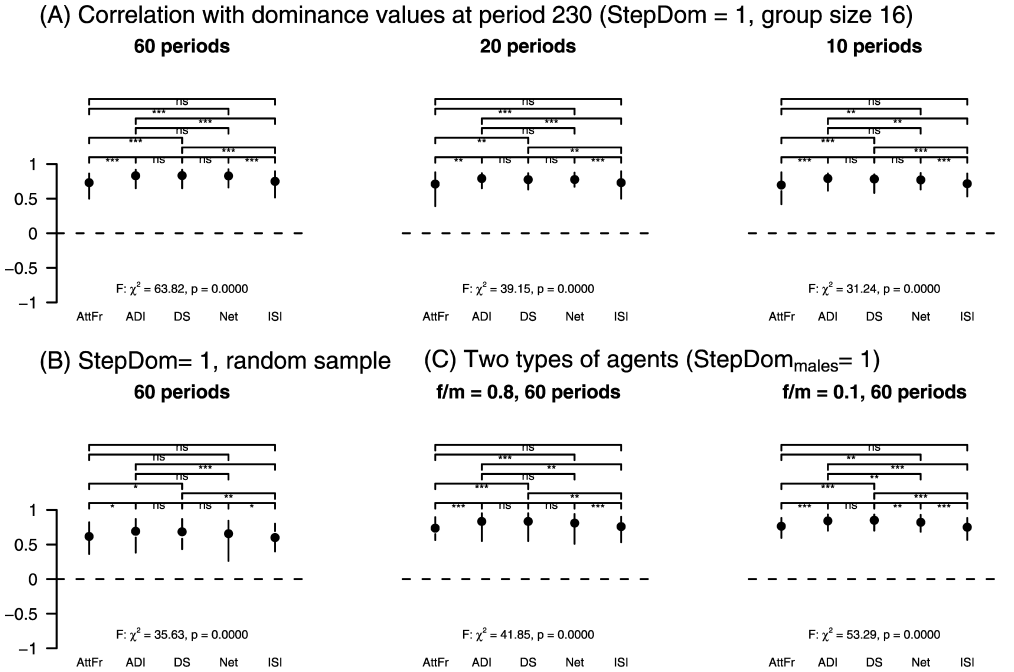

Figure 1. Kendall rank correlation between the hierarchy based on Dom-values and that derived by five ranking methods (the point indicates the median of 40 replica's, vertical bars range from the minimum to the lower quartile and from the upper quartile to the maximum value). The five ranking methods are: attack frequency (AttFr), average dominance index (ADI), David's score (DS), number of subordinates (Net) and the linearity-maximising method (I&SI). In the top of the figure the efficiency of pairs of methods is compared by the Wilcoxon matched-pairs signed-ranks test; significances are corrected by the Bonferroni-Holm procedure. In the bottom of the figure the Friedman-test is used to compare the efficiency of all five procedures. f/m indicates the relative stepDom (intensity of aggression) of females compared to males. 60, 20 and 10 periods indicate over which interval the performance is studied (see text).

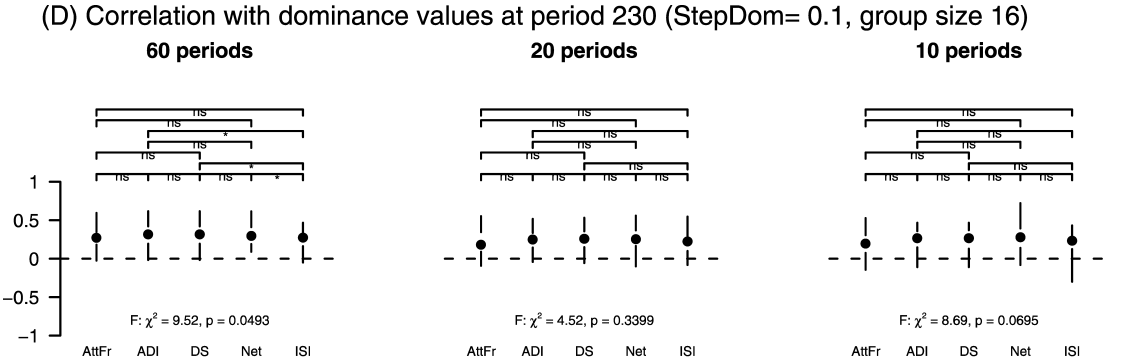

Figure 1. (Continued).

Table 3. Pair-wise comparison (using the Wilcoxon matched-pairs signed-ranks test with a Bonferroni-Holm correction) of the efficiency of five ranking methods to rank agents in a hierarchical order. (A) Group with one type; (B) group with two types, 'males' and 'females'; (C) Hierarchy per sex in bi-sexual group. Results are based on 40 replicas. Data of complete stable phases of 60 periods were used. The five ranking methods are: attack frequency (AttFr), average dominance index (ADI), David's score (DS), number of subordinates (Net) and the linearity-maximising method (I&SI). NS = not significant, $\sim = p < 0.10$, $* = p < 0.05$, $** = p < 0.01$, $*** = p < 0.001$; StepD = StepDom; InitD = InitDom. N = group size. 'f/m' indicates the relative StepDom of females compared to males. '<' and '>' indicates that the first performs worse or better, respectively, than the second method.

Wilcoxon matched pairs signed rank test	AttFreq vs ADI	AttFreq vs DS	AttFreq Netto	AttFreq I&SI	ADI vs DS	ADI vs Netto	ADI vs I&SI	DS vs Netto	DS vs I&SI	Netto I&SI
A) One Type, initD = 24										
$N = 16$										
StepDom = 1.0	<***	<***	<***	NS	NS	NS	>***	NS	>***	>***
StepDom = 0.1	NS	NS	NS	NS	NS	NS	>**	NS	>**	>**
$N = 8$										
StepDom 1.0	<***	<***	<**	NS	NS	>*	>***	>*	>***	>*
StepDom 0.1	<*	<*	NS	NS	NS	>*	>*	>*	>*	NS
$N = 4$										
StepDom = 1.0	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
StepDom = 0.1	NS	NS	<~	<~	NS	NS	NS	NS	NS	NS
B) Bisexual group, $N = 16$, InitD = 16 / 32										
StepD = 1, f/m = 0.8	<***	<***	<***	NS	NS	NS	>***	NS	>***	>***
StepD = 1, f/m = 0.1	<***	<***	<***	NS	NS	>**	>***	>***	>***	>***
StepD = 0.1, f/m = 0.8	<~	<~	NS	>**	>*	NS	>***	NS	>***	>***
StepD = 0.1, f/m = 0.1	NS	NS	<~	>~	NS	NS	>**	NS	>**	>***

Table 3. (Continued).

Wilcoxon matched pairs signed rank test	AttFreq vs ADI	AttFreq vs DS	AttFreq Netto	AttFreq I&SI	ADI vs DS	ADI vs Netto	ADI vs I&SI	DS vs Netto	DS vs I&SI	Netto I&SI
C) Hierarchy per sex, $N = 8$										
StepD = 1.0, f/m = 0.8										
‘females’, initDom = 16	<***	<***	<***	<**	NS	NS	NS	NS	NS	NS
‘males’, initDom = 32	<**	<**	<**	<*	NS	NS	>∼	NS	>*	NS
f/m = 0.1										
‘females’, initDom = 16	<***	<***	<**	<*	NS	NS	>*	NS	>*	>*
‘males’, initDom = 32	<**	<**	<***	<**	NS	NS	NS	NS	NS	>**
StepD = 0.1, f/m = 0.8										
‘females’, initDom = 16	<*	<∼	NS	NS	NS	NS	NS	NS	NS	>∼
‘males’, initDom = 32	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
f/m = 0.1										
‘females’, initDom = 16	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
‘males’, initDom = 32	NS	NS	NS	NS	NS	NS	NS	NS	NS	>∼

differences between ranking methods disappears and all ranking methods give approximately the same results.

To reflect sex differences in fighting capacity, we also study groups of 16 agents consisting of an equal number of two types of agents that differ in their fighting capacity thus reflecting 'males' and 'females' (Figure 1C, F; Table 3B). Results remain similar. At a low intensity, however, the ranking methods perform better for societies with two types of agents than with a single type of agent (compare Figure 1D, first graph, to F, last graph). This is due to the steeper hierarchy that results from the larger difference in dominance between the sexes ('males' and 'females', initDom of 32 and 16 respectively).

Since in studies of real animals, dominance hierarchies are usually established per sex, we compare the five ranking methods also for the hierarchies per sex in bi-sexual groups (Figure 2, Table 3C). Results of the data per sex show that, as before, the hierarchy based on the average dominance index (ADI), David's score (DS) and number of subordinates (Netto) correspond closest to the real hierarchy and that the ranking methods based on attack frequency and the I&SI method perform (significantly) less well.

Results are the same for matrices of attack (as shown here) and matrices of winning, because both matrices are significantly correlated in the model (data not shown).

Discussion

The efficiency of the simplest ranking method, the attack (or winning) frequency (AttFr) and that of the most complex one, the linearity-maximising method (I&SI), is minimal. Unexpectedly, the ranking methods of medium complexity, the average dominance index (ADI), its refined version (David score, DS) and the number of subordinates each individual dominates over (Netto) reproduce the 'real' Dom-value based hierarchy among the agents with significantly greater precision than others. Further, in a few cases the average dominance index and David's score are significantly more efficient than the method based on the number of subordinates an individual dominates over.

The bad performance of the simplest ranking method by the total frequency of attack/winning arises, because it does not take relational aspects

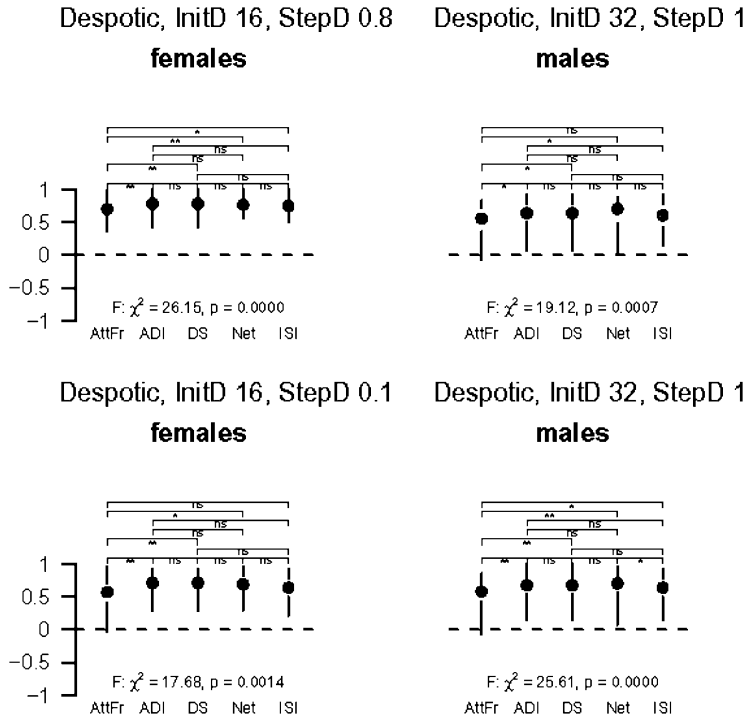

Figure 2. Kendall rank correlation between the real hierarchy and that derived using five ranking methods for male-agents and for female-agents for different values of StepDom and initDom. Groups contain 16 individuals with two types of agents (8 individuals each sex).

into account (it does not show to whom an individual is dominant) and it is influenced by the frequency with which an individual encounters others.

As regards the most complex ranking method, in which the linear order is maximised, its bad result is unexpected and in contradiction with results by de Vries & Appleby's comparative study (2000). In the model its low efficiency compared to that of the other three is due to the fact that during the process of maximising of the linear ordering of the hierarchy, it allows single events of winning and losing within a pair to change a dominance relation. This procedure may be misleading, because it incorrectly assumes that all events of winning produce the same effect, whereas in the model the impact may differ between events depending on the aggression intensity and the relative power of both agents (as in the real world). Although the other methods, i.e., average dominance index and number of subordinates, also use this assumption incorrectly, they are not unduly influenced by it, because

they use a single average value or summed count per individual (over all interaction partners) on the basis of which they are ranked. Therefore, they produce better results.

Of these methods the results of ranking by the number of subordinates an individual is dominant to (Netto) are slightly weaker than by the average dominance index and David's score. This is due to the fact that one event of winning (or losing) can have a great effect on the final count of the number of subordinates. For instance, if an individual A wins in one more interaction with B than vice versa, B counts as an additional subordinate to A and heightens A's count by one. In case of the average dominance index and David's score, a single event of winning and losing is unlikely to have so great an effect because this index is based on the averaging of winning proportions with all its opponents. Average dominance index and David's score perform equally well. The advantage of the average dominance index over David's score is that it is easier to understand and computationally simpler.

Note that the results also hold if individuals in the model do not perceive the dominance of each other directly (when e.g., rank is indicated through pheromones or body posture), but estimate dominance of others from their memorised experiences with them (data not shown). This experience-based dominance perception has been implemented as described by Hemelrijk (2000). It implies that everyone has a slightly different impression of the dominance of other group members, because of its different experience with each group member (Hemelrijk, 2000).

These results are robust, because they hold for different hierarchical structures as they occur in several group sizes and types of societies and types of agents (as regards 'sex' and dominance perception). Results are clearer for the complete data (that is if each single act of interaction is observed) than for random samples of all acts taken from it, and for larger than for smaller groups, because in both cases the discriminatory power of the comparative method increases with sample size. If we reduce the interval length, however, results sometimes become slightly more marked despite the smaller sample size. This arises because the hierarchy is more stable over a shorter interval than a longer one.

All five ranking methods are better at a high intensity than at a low one. This is due to the steeper hierarchy, which implies that the hierarchy is more marked (Hemelrijk, 1999). For the same reason correlations in Figure 2 are stronger among females than males. The hierarchy among females is steeper

despite the female's lower intensity of attack, because the same changes in dominance values have a greater effect on female dominance values than on those of males, because of the lower dominance of females (Hemelrijk & Gygax, 2004).

It needs to be noted, however, that these results were observed in a virtual world. This comprises a useful device for thinking (better than a simple randomisation procedure that is usually applied) but it differs from the phenomenon in the real world and is simpler. There are, however, several reasons why these results are relevant for studies of real animals.

First, the agents of the model are equipped with behavioural tendencies (grouping and competitive interactions) that are found in many animal species.

Second, the model reproduces many characteristics that are similar to those found in natural societies with different dominance styles, namely egalitarian and despotic societies (Vehrencamp, 1983) and thus, it catches some of the essentials of dominance style (Hemelrijk, 1999, 2000).

Third, data collection of the behaviour of artificial agents in the model is performed by artificial observers in a similar way as it is done in studies of real animals and the data of the frequencies of dominance interactions and their outcome are cast into the interaction matrices also in a similar way.

Fourth, the model shows us that the efficiency of the dominance index and David's score is highest under almost all circumstances: for different social systems, group sizes, hierarchical gradients and periods over which data are collected. This is important, because such aspects (social systems, group sizes) vary between groups of animals, and it seems possible that the efficiency of different ranking methods actually depends on them. However, results show that, either, the average dominance index and David's score are significantly better than the others or that, for instance, at a low intensity, results do not differ significantly between all methods.

In short, models like DomWorld is a useful device to evaluate statistically the efficiency of ranking methods that produce a dominance hierarchy among individuals in a group on the basis of matrices containing the frequency of winning/attack. In the model, the ranking methods of medium complexity, in particular the average dominance index and David's score, correlate with dominance order best. Therefore, and because the average dominance index is also computationally simple and easy to understand, results of this model lead us to recommend this ranking method of average dominance index for

estimating the hierarchical order of real individuals from the matrix of their agonistic interactions.

Acknowledgements

We would like to thank Rolf Pfeifer, Franjo Weissing / the University of Groningen for continuous support. We are grateful for the financial support provided to Lorenz Gygax by the Swiss National Science Foundation project number 3100-0654444. We thank Han de Vries and Pascal Girod for comments on a former version.

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.

Barchas, P.R. & Mendoza, S.D. (1984). Emergent hierarchical relationships in rhesus macaques: an application of Chase's model. — In: Social hierarchies: Essays towards a sociophysiological perspective (Barchas, P.R., ed.). Greenwood Press, p. 81-95. Westport, CT.

Bernstein, I.S. & Ehardt, C.L. (1985). Intragroup agonistic behavior in Rhesus monkeys (*Macaca mulatta*). — Int. J. Primatol. 6: 209-226.

Boyd, R. & Silk, J.B. (1983). A method for assigning cardinal dominance ranks. — Anim. Behav. 31: 45-58.

Caldecott, J.O. (1986). Mating patterns, societies and ecogeography of macaques. — Anim. Behav. 34: 208-220.

Chase, I.D., Bartelomeo, C. & Dugatkin, L.A. (1994). Aggressive interactions and intercontest interval: how long do winners keep winning? — Anim. Behav. 48: 393-400.

Chase, I.D., Tovey, C. & Murch, P. (2003). Two's company, three's a crowd: Differences in dominance relationships in isolated versus socially embedded pairs of fish. — Behaviour 140: 1193-1218.

Chase, I.D., Tovey, C., Spangler-Martin, D. & Manfredonia, M. (2002). Individual differences versus social dynamics in the formation of animal dominance hierarchies. — Proc. Natl. Acad. Sci. USA 99: 5744-5749.

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.). — Anim. Behav. 27: 211-225.

Drummond, H. & Canales, C. (1998). Dominance between booby nestlings involves winner and loser effects. — Anim. Behav. 55: 1669-1676.

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.

Hemelrijk, C.K. (1999). An individual-oriented model on the emergence of despotic and egalitarian societies. — Proc. Roy. Soc. London B 266: 361-369.

Hemelrijk, C.K. (2000). Towards the integration of social dominance and spatial structure. — Anim. Behav. 59: 1035-1048.

Hemelrijk, C.K. (2002). Despotic societies, sexual attraction and the emergence of male 'tolerance': an agent-based model. — Behaviour 139: 729-747.

Hemelrijk, C.K. & Gygax, L. (2004). Dominance style, differences between the sexes and individuals: an agent-based model. — Interaction studies: Social behavior and communication in biological and artificial systems 5: 131-146.

Honk, C.v. & Hogeweg, P. (1981). The ontogeny of the social structure in a captive *Bombus terrestris* colony. — Behav. Ecol. Sociobiol. 9: 111-119.

Hsu, Y. & Wolf, L.L. (1999). The winner and loser effect: integrating multiple experiences. — Anim. Behav. 57: 903-910.

Karavanich, C. & Atema, J. (1998). Individual recognition and memory in lobster dominance. — Anim. Behav. 56: 1553-1560.

Mazur, A. (1985). A biosocial model of status in face-to face primate groups. — Social Forces 64.

Roberts, J.M. (1990). Modeling hierarchy, transitivity and the linear ordering problem. — J. Math. Sociol. 16: 77-87.

Theraulaz, G., Gervet, J., Thon, B., Pratte, M. & Semenoff-Tian-Chansky, S. (1992). The dynamics of colony organization in the primitively eusocial wasp *Polistes dominulus* (Christ). — Ethology 91: 177-202.

Thierry, B. (1990). Feedback loop between kinship and dominance: the macaque model. — J. theor. Biol. 145: 511-521.

Thierry, B. (2000). Covariation of conflict management patterns across *Macaque* species. — In: Natural conflict resolution (Aureli, F. & de Waal, F.B.M., eds). University of California press, Berkeley, p. 106-128.

Thierry, B., Anderson, J.R., Demaria, C., Desportes, C. & Petit, O. (1990a). Tonkean macaque behaviour from the perspective of the evolution of Sulawesi macaques. — In: Current primatology: Social development; learning and behaviour (Roeder, J.J., Thierry, B., Anderson, J.R. & Herrenschmidt, N., eds), p. 83-103.

Thierry, B., Gauthier, C. & Peignot, P. (1990b). Social grooming in Tonkean macaques (*Macaca tonkeana*). — Int. J. primatol. 11: 357-375.

Thouless, C.R. & Guinness, F.E. (1986). Conflict between red deer hinds: the winner always wins. — Anim. Behav. 34: 1166-1171.

Vehrencamp, S.L. (1983). A model for the evolution of despotic versus egalitarian societies. — Anim. Behav. 31: 667-682.

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 hieryrchy: a comparison of the I&SI and the BBS method. — Anim. Behav. 59: 239-245.

de Waal, F.B.M. (1989). Dominance style and primate social organisation. — In: Comparative socioecology (Standon, V. & Foley, R.A., eds). Blackwell, Oxford, p. 234-264.

de Waal, F.B.M. & Luttrell, L.M. (1989). Towards a comparative socioecology of the genus *Macaca*: different dominance styles in rhesus and stumptail monkeys. — Am. J. Primatol. 19: 83-109.

Zumpe, D. & Michael, R.P. (1986). Dominance index: A simple measure of relative dominance status in primates. — Am. J. Primatol. 10: 291-300.