Reducing the bias due to unknown relationships in measuring the steepness of a dominance hierarchy

Authors: Saccà, Tommaso[1*], Gort, Gerrit[2], van de Waal, Erica[3,4], Hemelrijk, Charlotte K[1]

[1] Groningen Institute for Evolutionary Life Sciences, GELIFES; University of Groningen, the Netherlands

[2] Biometris; Wageningen University, The Netherlands

[3] Department Ecology & Evolution; University of Lausanne, Switzerland

[4] Inkawu Vervet Project; Mawana Game Reserve, KwaZulu Natal, South Africa

*Corresponding author: t.sacca@rug.nl

Abstract

Measuring the steepness of a dominance hierarchy is important for classifying a social system in a continuum between egalitarian and despotic. For this, often the steepness-slope from de Vries and colleagues (*Animal Behaviour*, 2006, **71**, 585-592) is used. It compares the cardinal and ordinal dominance rank of each individual using the slope of the linear regression. The disadvantage of this measure is that the slope becomes lower the higher the proportion of unknown relationships (dyads without interactions). In the present paper, we investigate what causes this bias, and propose a solution. We show: (1) that the bias is due to the treatment of unknown relationships by the dominance index currently used in this methodology, the David's score (namely by assuming, among other things, an equal number of wins and losses for both members of the pair). (2) Instead, using the Average Dominance Index (the average proportion of wins by each individual from all its opponents) reduces the bias due to unknown relationships, because it excludes unknown relationships, and (3) the standard error of the steepness slope based on the Average Dominance Index is smaller. (4) The two indices (David's score and Average Dominance Index) result in similar steepness-slopes when all relationships are known. For comparing the two indices we use empirical data (from four group-years of wild vervet monkeys) and data from a computational model on dominance interactions in a group (DomWorld). We conclude that the Average Dominance Index (compared to the David's score) is preferable for measuring the steepness-slope.

Keywords: Steepness of hierarchy; David's score; Average Dominance Index; Dominance interactions; Egalitarian society; Despotic society

In social animals, the steepness of a group's dominance hierarchy classifies its dominance style along a continuum from "egalitarian" at one end to "despotic" at the other (e.g., Vehrencamp, 1983; van Schaik, 1989). Conceptually, the hierarchy's steepness represents the difference in ability to win agonistic interactions between individuals of adjacent rank: when the difference is large, the hierarchy is steep and the society is despotic, and when it is small, the hierarchy is shallow and the society egalitarian (Vehrencamp, 1983). The most widely used measure of steepness of the hierarchy is that proposed by De Vries et al. (2006), with over 300 citations to date (in Web of Science Core Collection). De Vries et al. (2006) quantified steepness as the absolute value of the slope of the (least squares straight) line describing the relationship between the group member's ordinal rank (on the X axis) and its cardinal rank (on the Y axis). The ordinal rank of an individual represents its relative position in the hierarchy (e.g., first, second … n^{th} position, with n equal the number of individuals), whereas its cardinal rank represents a continuous measure of its competitive ability calculated using a dominance index. As dominance index, de

Vries et al. (2016) used the David's Score, normalized for group size. This index of steepness is problematic, however, because its value decreases with the percentage of pairs of group members (dyads) without competitive interactions, the so-called 'unknown relationships' (Klass & Cords, 2011). In empirical datasets unknown relationships are common. For example, a meta-analysis of 101 interaction matrices from 55 published studies found unknown relationships to comprise, on average, 26% of the total number of dyads (standard deviation = 23%), reaching values up to 70% in few cases (Shizuka & McDonald, 2012). Therefore, a steepness measure that is less influenced by unknown relationships is desirable.

In the present work, we propose a modification of the steepness-slope that is less influenced by unknown relationships. We argue that the sensitivity of the steepness-slope to unknown relationships is caused by the David's Score's treatment of unknown relationships as being perfectly egalitarian, i.e., as if the two individuals had the same number of wins and losses (de Vries, Stevens, & Vervaecke, 2006). We minimize this sensitivity by excluding unknown relationships from the calculation of the dominance index as is done in the Average Dominance Index, ADI (Hemelrijk, Wantia, & Gygax, 2005). The Average Dominance Index ranks individuals according to their average proportion of winning from all their opponents (excluding group members with whom they did not fight). Thus, we expect the steepness-slope to be less affected by unknown relationships when we use the Average Dominance Index in the measure. A steepness measure that is influenced by unknown relationships less is preferable in empirical datasets, because here unknown relationships are common.

We first explain the mathematical formula of the David's Score and why it is particularly susceptible to unknown relationships in the group. Next, we explain the formula of the Average Dominance Index and show that it is less impaired by these.

The David's Score

The David's Score consists of four elements, the first two increasing the final score, while the last two decrease it, with higher scores representing better competitive ability, thus higher dominance (de Vries et al., 2006):

$$\text{David's score} = W + W_{\text{weighted}} - L - L_{\text{weighted}}$$

The calculation of the four elements for an individual i is as follows: the first element, W, is the sum of dyadic proportions of fights won by the individual i with all other group individuals $j, j \neq i, j \leq N$, where *N* is the group size.

$$W_i = \sum_{j=1}^{N(j \neq i)} \frac{\text{wins}_{ij}}{\text{n of fights}_{ij}} \quad \text{(equation 1)}$$

The second element, W_{weighted}, is the sum of the dyadic proportion of fights won by the individual i with each other individual $j, j \neq i$, weighted by the combat capabilities of each interacting partner, represented by the interacting partner's W value, as in equation 1.

$$W_{i\ \text{weighted}} = \sum_{j=1}^{N(j \neq i)} W_j \frac{\text{wins}_{ij}}{\text{n of fights}_{ij}} \quad (equation\ 2)$$

The third element, L, is the sum of the dyadic proportion of fights lost by individual i with each other individual $j, j \neq i$. The proportion of losses for each dyad is complementary to the proportion of victories.

$$L_i = \sum_{j=1}^{N(j \neq i)} \frac{\text{losses}_{ij}}{\text{n of fights}_{ij}} \quad (equation\ 3)$$

The fourth element, L_{weighted}, is the sum of the dyadic proportion of fights lost by the individual i with each other individual $j, j \neq i$, weighted by the propensity of the interacting partner to lose a fight, represented by the partner's L value.

$$L_{i\ \text{weighted}} = \sum_{j=1}^{N(j \neq i)} L_j \frac{\text{losses}_{ij}}{\text{n of fights}_{ij}} \qquad (equation\ 4)$$

The David's Score for each individual is the sum of the first two elements (equation 1 and 2) minus the last two (equation 3 and 4). When there are no interactions between two individuals, i and j, (n of $\text{fights}_{ij} = 0$), the four parts of the David's Score for that dyad_{ij} are zero. This is also true in a perfectly egalitarian relationship, i.e., where the two individuals i and j have (1) the same number of wins (wins_{ij}) and losses (losses_{ij}) from each other and (2) the same number of total wins (W_i and W_j, equation 1) as losses (L_i and L_j, equation 4) with all group members (W_i=W_j=L_i=L_j). In this theoretical case, the first two elements of the David's Score would be equal to the second two and the sum would be zero for that dyad_{ij}. This zero value is true for the David's Score of both individuals, i and j.

For example, we can take two interaction matrices: one in which individuals A and B have the same number of wins and losses with each other and with two other individuals, C and D (interaction matrix I, Table 1), and another one in which individuals A and B never interact with each other, and have non-egalitarian relationships (i.e., different number of wins and losses) with C and D (interaction matrix II, Table 1). In the egalitarian case, the contribution of the single dyad_{ab} to the David's Score of individual A is zero (0.5 + 0.25 – 0.5 – 0.25 = 0, equations 1 to 4). The same is true for the David's Score of individual B (0.5 + 0.25 – 0.5 – 0.25 = 0). In the second matrix, the contribution of the single dyad_{ab} to the David's Score of individual A is also zero (0 +

0 – 0 – 0 = 0). The same is true for the David's Score of individual B (0 + 0 – 0 – 0 = 0). The effect of $dyad_{ab}$ in the calculation of the David's Score of both individuals (A and B) is the same in both hierarchies. This example shows that even if the two individuals (A and B) have different competitive abilities (A is higher ranking than B in the second interaction matrix), a possible unknown relationship between the two is treated as if they had the same number of wins and losses in their dyad ($dyad_{ab}$) as well as with the other individuals in the hierarchy ($dyad_{ac}$, $dyad_{ad}$, $dyad_{cb}$, $dyad_{db}$). This result is illogical and highlights further how the treatment of unknown relationships in the David's Score is a methodological limitation.

So far, the values of the David's Score range between -N(N-1)/2 to +N(N-1)/2, where N is the group size, and this causes the steepness-slope to range between 0 and N. To make the values of the steepness-slope range between 0 and 1, de Vries et al. (2006) normalized the David's Score, such that its values range between 0 and N-1 (equation 5). The maximum value of steepness is calculated for a perfectly linear dominance hierarchy in which all individuals consistently defeat others ranking below them, and lose all interactions from others above them (de Vries et al., 2006).

$$\mathrm{normDS} = \frac{\mathrm{DS} + N(N-1)/2}{N} \quad (equation\ 5)$$

The Average Dominance Index

The Average Dominance Index, ADI (Hemelrijk et al., 2005), is calculated for each individual i as its average proportion of winning from each opponent j with whom it had an interaction (equation 6).

$$\mathrm{ADI}_i = \frac{W_i}{\mathrm{n\ of\ opponents}_{ij}(*)} \quad (equation\ 6)$$

$$(*)\ \text{opponents}_{ij} \text{ with at least one interaction}$$

When a relationship between two individuals i and j is unknown ($\text{n of fights}_{ij} = 0$), that partner does not contribute to W_i nor to the $\text{n of opponents}_{ij}$, and therefore is omitted from the calculation of its average proportion of winning, ADI_i.

When used to calculate a steepness-slope that should range between 0 and 1, the Average Dominance Index needs to be normalised with group size, because its un-normalised values range between 0 and 1, regardless of group size. Without normalising the Average Dominance Index, when the index is used in the steepness-slope, the values of the steepness-slope become smaller the larger the group size, because the maximal cardinal rank equals the group size (N = group size) while the maximal Average Dominance Index is independent of group size and remains 1. Therefore, to get a steepness-slope between 0 and 1, the Average Dominance Index needs to be normalised with group size. When multiplying the Average Dominance Index by N-1 (the group size minus one), the normalised Average Dominance Index ranges between 0 to N-1, the same range as that of the normalized David's Score and consequently, the steepness values using the Average Dominance Index range between 0 and 1 too. The range of the values of the steepness-slope is the same when calculated with the normalized David's Score and normalized Average Dominance Index. From now on (apart from in tables, in abbreviations, and in the Conclusion section, in which the normalization is expressly mentioned) when referring to these dominance indices, we imply their normalized version.

Empirical data

Our empirical data concern aggressive interactions of vervet monkeys (*Chlorocebus pygerythrus*). collected at the Inkawu Vervet Project (IVP) in the Mawana game reserve, South Africa, between

2011 and 2018. Three groups of habituated wild vervet monkeys took part in the study: 'Ankhase' (AK), 'Kubu' (KB), and 'Noha' (NH). Habituation began in 2010 in AK and NH, and in 2013 in KB. Data were collected by several trained observers. All IVP observers are trained to identify each monkey by individual bodily and facial features (eye-rings, scars, color, shape etc.) and have to pass an identification test for all the individuals of a group. Before beginning the collection of behavioural data, observers had to pass an inter-observer reliability test with Cohen's kappa > .80 for each data category between two observers. Data were initially collected on handheld computers (Palm Zire 22) using Pendragon software version 5.1 and, from the end of August 2017, on tablets (Vodacom Smart Tab 2) and smartphones (Runbo F1) equipped with the Pendragon version 8.

The aggressive data were collected through *ad libitum* (Altmann, 1974) sampling because conflicts happened infrequently. In our matrices, we noted only conflicts that were dyadic, between adults, and had a clear outcome: the last behaviour of the winner was aggressive (namely, stare, chase, attack, hit, bite, take place) and the last behaviour of the loser was submissive (namely, retreat, flee, leave, avoid, jump aside).

Ethics guidelines: Our study adhered to the "Guidelines for the use of animals in research" of Association for Study of Animal Behavior and was approved by the relevant local authority, Ezemvelo KZN Wildlife, South Africa.

Monte Carlo simulations

We started by selecting four group-years with the smallest percentage of unknown relationships in the interaction matrices (Table 1). We monitored the performance of the steepness-slope using both indices (the David's Score and the Average Dominance Index) in each group-year, starting with the original percentage of unknown relationships and progressively increasing the percentage of unknown relationships. For this, we iteratively excluded one dyad at a time from the interaction matrix using Monte Carlo simulations. This is the same approach used by Klass and Cords, (2011). After the exclusion of each dyad, we calculated the steepness-slope using both indices. For each group-year, we calculated the median value of hierarchical steepness for both indices for 100 runs of the Monte Carlo simulation with each run using a new, random sequence of excluding dyads. We omitted runs in which, after the exclusion of the last dyad, one or more individuals were completely removed from the interaction matrix because they did not have any remaining interactions, because their absence would change the group size. For each Monte Carlo run, we calculated the Kendall Rank correlation between hierarchical steepness and proportion of unknown relationships. Applying the Wilcoxon signed rank test, we investigated whether the distribution of these Tau values was symmetrical around 0.

We compared how the two indices changed depending on the percentage of unknown relationships. For this, we calculated their percentage change in steepness as the difference in steepness between that of the original social interaction matrix and the median of the steepness when the percentage of unknown relationships approached one-fourth (25%) and two-thirds (approximately 66.7%). We chose these particular percentages to match the average (26%) and highest (up to 70%) percentage of unknown relationships found in published empirical datasets (Shizuka & McDonald, 2012). For each group-year, we took the percentage of unknown relationships closest to these thresholds, and calculated the median value of hierarchical steepness over the 100 Monte Carlo runs.

Using a computational model to create hierarchies without unknown relationships

To understand the performance of the two indices in the steepness-slope in the absence of unknown relationships, we produced artificial data on dominance interactions in the computational model, DomWorld (Hemelrijk, 1999). We use a computational model because, compared to empirical data, it is easier to create a large number of dominance hierarchies and social interaction matrices in which all interactions are known. In DomWorld, individuals group and interact competitively. We derive the dominance hierarchy from the social interaction matrix of winners and losers in the group. To validate the computational model, DomWorld, against the empirical data, we investigated, for both the David's Score and the Average Dominance Index, whether an increasing percentage of unknown relationships results in values of steepness-slope that resembles those in the empirical data. To compare against empirical data we used 100 runs of Monte Carlo simulations for ten hierarchies obtained from the model DomWorld and calculated the median values and percentage decrease of hierarchical steepness at one-fourth and two-third unknown relationships for each run. We omit Monte Carlo runs in which some individuals were completely excluded from the interaction matrix because all their dyads were randomly excluded, not to affect group size. This approach is the same as in empirical data. We compared both indices also for 100 hierarchies in DomWorld when all relationships were known. For each matrix, we calculated the hierarchical steepness for both indices, and compared their steepness-slopes with a linear regression.

Monte Carlo simulations and statistical analyses were conducted in R 4.0.3 (R Core Team, 2021). For the calculation of hierarchical steepness with the David's Score, we used the package EloRating (Neumann & Kulik, 2020) with the correction for the number of interactions in a dyad from de Vries *et al.* (2006), called D_{ij}. D_{ij} is a corrected proportion of wins for each dyad that adjusts the original proportion (P_{ij}) by subtracting a value linked with how likely the proportion of wins (P_{ij}) is given the amount of interactions in the dyad (equation 7: de Vries et al, 2006). The probability to observe each proportion of wins (P_{ij}) is calculated from a uniform distribution, and is equal to $1/(n_{dyad} + 1)$, with n_{dyad} being the number of interactions in the dyad. Thus, the value that is subtracted from the original proportion of wins for the dyad (P_{ij}) is smaller the greater number of interactions in the dyad; the value that is subtracted approaches zero when the number of interactions tends to infinite, but can never be zero.

$$D_{ij} = P_{ij} - \{(P_{ij} - 0.5) \text{ x likelihood}[P_{ij}]\} \quad (equation\ 7)$$

In our present, we use D_{ij} over the simpler P_{ij} for the proportion of wins in a dyad in calculating the David's Score, as was suggested by de Vries et al. (2006). However, Balasubramaniam et al. (2013) suspected that with an increasing percentage of unknown relationships D_{ij} might artificially decrease steepness-slopes more than P_{ij}. Thus, we first study in the computational model DomWorld whether D_{ij} and P_{ij} differ in their bias depending on the percentage of unknown relationships.

Testing the performance of David's Score's Dij versus Pij in the steepness-slope

We used a dominance hierarchy obtained from DomWorld to compare the performance of using in the David's Score the simple proportion of wins (P_{ij}) versus the proportion of wins corrected for chance (D_{ij}). When we used D_{ij} we found a smaller steepness-slope than when we used P_{ij}, when all relationships were known. However, the trend for the steepness-slope with an increasing percentage of unknown relationships was the same whether we used D_{ij} or P_{ij} in calculating the steepness-slope. When the percentage of unknown relationships was close to 25%, the steepness-slope decreased by 38.9% (from 0.712 to a median value of 0.435) when we used D_{ij}, compared to a decrease of 38.8% (from 0.735 to a median value of 0.450) when we used P_{ij}. When the percentage of unknown relationships was close to 66.7%, the steepness-slope decrease for both D_{ij} and P_{ij} by 83.4% (for D_{ij}: from 0.712 to a median of 0.118; for P_{ij} from 0.735: to a median of

0.122). Thus, from here onward we calculate the David's Score using the correction for the number of interactions in a dyad (D_{ij}), as originally proposed by de Vries et al. (2006).

Performance of the two indices in empirical data

The percentage of unknown relationships in the matrices of the four group-years of vervet monkeys from the Inkawu Vervet Project was significantly, negatively correlated with the steepness-slope calculated with the David's Score but significantly, positively when calculated with the Average Dominance Index (Table 2). The correlation was stronger with the David's Score than with the Average Dominance Index. At approximately 25% unknown relationships, steepness-slopes changed 62.1 times more due to unknown relationships when based on the David's Score than on the Average Dominance Index and at approximately 66.7% they changed 3.8 times more. At one-fourth unknown relationships (~25%), on average steepness-slopes calculated with the David's Score decreased by 28.8% (min -17.6%, max -39.2%), and with the Average Dominance Index decreased by 0.02% (min -0.05%, max +1.96%). At two-thirds unknown relationships (~66.7%), the steepness-slopes calculated with the David's Score decreased on average by 78% (min -77.5%, max -80.6%), and that by the Average Dominance Index increased by 20.3% (min +16.6%, max +22.6%).

Validating the computational model against empirical data

Results in DomWorld resemble those from the empirical data on vervet monkeys. With the David's Score, the steepness-slopes decreased on average by 38.7% at one-fourth unknown relationships, and by 83.0% at two-thirds. With the Average Dominance Index, it increased by 2.1% at one-fourth unknown relationships, and by 15.7% on average at two-thirds. Therefore, when 25% of the relationships were unknown the steepness-slopes changed approximately 18.4 times more due to unknown relationships when based on the David's Score than the Average Dominance Index and 5.3 times more when unknown relationships were approximately 66.7%. These results are for interaction matrices of groups of ten males in DomWorld. In the model, males are more intense in their aggression than females. From here onwards the DomWorld results are based on interaction matrices of ten males. This is because results do not change when we match sex composition and group size to those in the empirical group-years (data not shown, available on request) and we want to have a single baseline for interaction matrices across all analyses, with a large group size and moderately steep hierarchies (which we get by only including the sex with the most intense aggression in the model).

Testing the performance of the two indices in matrices without unknown relationships

In matrices where all relationships are known (there are interactions in all dyads), steepness-slopes that use the David's Score and Average Dominance Index give similar values (linear regression of the David's Score steepness on the Average Dominance Index steepness results in intercept = 0.002, slope = 0.966, $R^2 = 0.9994$, $P < 0.001$, Figure 2). The slope in this regression was slightly below one, and the steepness-slopes using the David's Score were on average 0.022 lower (standard deviation = 0.00026) than with the Average Dominance Index.

Testing the combination of bias and variability

The quality of an estimator also depends on its variability, with a smaller variability generally preferred among predictors with the same bias. To evaluate the contribution of the variability to the total deviation, we use the root mean squared deviation (RMSD). It is the square root of the sum of the square of the bias (difference between sample mean and value to be estimated) plus the variance of the sample. A smaller value of the RMSD is preferable. We used data from the computational model, DomWorld, to have initial matrices where all relationships are known to use as a reference for the unbiased value of steepness. At each step, the deviation of steepness was calculated per dominance index as the difference between the average steepness-slopes at that step and the steepness-slopes without unknown relationships. We used ten matrices obtained from DomWorld where all relationships were known and made 100 runs of the Monte Carlo simulation for each matrix. In DomWorld, the average value of RMSD at 25% unknown relationships was 0.27 for the David's Score and 0.05 for the Average Dominance Index. At 66.7%, the RMSD was 0.58 for the David's Score and 0.16 for the Average Dominance Index. Therefore, steepness-slopes based on the Average Dominance Index outperform those based on the David's Score by 5.4 times at one-fourth unknown relationships and by 3.6 times at two-thirds.

Discussion

The steepness-slope from de Vries et al. (2006) has been used widely to measure the steepness of a dominance hierarchy, even though this measure is known to be negatively correlated with the percentage of unknown relationships (Klass and Cords, 2011) and unknown relationships are common in empirical datasets (Shizuka & McDonald, 2012). Already in their initial publication, de Vries et al. warned about interpreting results from matrices with "observational zeroes" (i.e., unknown relationships). We argue that the problem arises because the David's Score treats unknown relationships as if both partners have the same ability to win competitive encounters. To address this limitation, de Vries et al. propose for each unknown dyad to use other behavioural observations outside of competitive interactions ("circumstantial observations", de Vries et al. 2006) to see whether an egalitarian relationship is a good approximation. For example, two individuals actively avoiding each other show an unresolved dominance relationships, in which case the unknown relationships can effectively be considered egalitarian. This is also true when two individuals tolerate each other by spending time in close proximity without physical aggression, which means that a dominance relationship between the two is missing. However, when one individual actively avoids the other, because a clear dominance relationship was already established in the past, the dyad should not be treated as egalitarian, but how the dyad should be treated instead (in the mathematical calculation of the index) is unclear. In addition to this problem, the use of circumstantial observations to define dominance relationships is problematic, as it requires additional data which are not always collected (or properly codified). To minimize the effect that unknown relationships have on the steepness-slope we propose to base it on the Average Dominance Index rather than the David's Score, because the Average Dominance Index excludes unknown dyads from the calculation, preventing researchers from relying on circumstantial observations and reducing the deviation of steepness values when the percentage of unknown relationships increases (Table 2 and Figure 1).

Our results confirm the negative relationship between the steepness-slope based on the David's Score and the percentage of unknown relationships in a group as shown by Klass and Cords (2011). We show that when the Average Dominance Index is used in the steepness-slope, this greatly reduces the effect of unknown relationships on the steepness measure. Although the steepness-slope based on the Average Dominance Index shows a significant, positive trend with an increasing percentage of unknown relationships, this trend is considerably smaller than that caused by the David's Score. Also for the combination of bias (the intensity of trend) and variation in values, tested with the root mean squared deviation of the two indices, the Average Dominance Index performs between 5.4 and 3.6 times better than the David's Score, even though the use of the Average Dominance Index results in a greater variation in steepness-slope values. The larger variability of the steepness-slope when using the Average Dominance Index when dyads are excluded could arise from the effect of missing dyads being larger and more random than the egalitarian, homogeneous treatment of unknown relationships in the David's Score. In the computational model DomWorld, we show a) the same trend of the steepness-slope with the percentage of unknown relationships as in empirical data, b) the similarity of the two indices, the David's Score and the Average Dominance Index, when all relationships are known, and c) that even when combining bias and variability the Average Dominance Index performs better than the David's Score. We also show that the David's Score's corrected for the number of interactions in a dyad (D_{ij}) performs similarly to its un-corrected version (P_{ij}) regarding the bias of the steepness-slope with an increasing percentage of unknown relationships.

The range of values of the steepness-slopes differ for the two dominance indices. When the David's Score is used, steepness-slopes never exceed the value of one, which is the theoretical maximum of a perfectly linear dominance hierarchy. The exclusion of any random dyads does not result in steeper values as each unknown relationship is treated as perfectly egalitarian and therefore, only decreases the general steepness of the hierarchy. When the Average Dominance Index is used, however, the steepness-slopes may reach values higher than one because the exclusion of unknown relationships may cause the hierarchical structure to become steeper (Figure 1).

In this work, for the calculation of the steepness-slope we use the David's Score with its correction for the number of interactions in a dyad (D_{ij}), as originally proposed by de Vries et al. (2006). We show that when the percentage of unknown relationships increases, the trend of the steepness-slope is virtually identical when we use the David's Score with the correction for the number of interactions in each dyad (D_{ij}) or without it (P_{ij}). We find that our use of D_{ij} returns a steepness-slope smaller than that of P_{ij}, similarly to what was found by de Vries et al. (2006). It is worth noting that, although we do not find a difference between our use of D_{ij} and P_{ij} in relation to increasing percentages of unknown relationships, this might be caused by our treatment of unknown relationships. Specifically, we iteratively excluded interactions between selected dyads of individuals to increase the percentage of unknown relationships in our data, but the number of interactions in the dyads that are not excluded is not affected. Thus, the correction for the number of interactions in a dyad (D_{ij}) does not have any effect. In some species, if groups with a higher percentage of unknown relationships also have a lower number of interactions in each dyad, this may still artificially reduce steepness-slope values when using D_{ij} instead of P_{ij}, as supposed by Balasubramaniam et al. (2013). However, we show that the percentage of unknown relationships alone cannot account for any difference between David's Score's D_{ij} and P_{ij} in the calculation of the steepness-slope.

Strauss and Holekamp (2019) proposed another modification of the David's Score, that considers previous interactions to the period analysed to better calculate David's Scores for each individual, and this slightly improves the performance of the index when unknown relationships increase. We do not test this modification in our present study, because we only test the methodology originally proposed. Although this modification is unlikely to significantly change the balance between the use of the Average Dominance Index and the David's Score, future work might consider investigating its effect on the performance of the David's Score in the steepness-slope.

In our present study, we reported results for two empirically relevant proportions of unknown relationships (25% as an average value, and 66.7% as the upper threshold: Shizuka & McDonald, 2012). This choice was made to highlight the trend occurring in the steepness-slope in relation to unknown relationships. It was not meant to indicate specific steepness values at any particular percentage of unknown relationships. When the steepness-slope was calculated with the Average Dominance Index, the difference in values between the steepness-slope of the original group and that of the group with an increased percentage of unknown relationship was not significantly different from zero until, on average for the four wild groups and the example group of DomWorld, 30.2% unknown relationships. Even so, values varied widely among groups and over time, with Kubu in 2017 having the lowest proportion (20%) and Noha in 2016 the highest (43%). When the steepness-slope was calculated with the David's Score, differences in steepness values were significantly different from zero for all percentages of unknown relationships and all groups. Thus, specific indications are not possible because of the difference among groups in trend for steepness-slopes in relation to unknown relationships.

In future, other methodologies of measuring the steepness of hierarchy (such as proposed by Sánchez-Tójar et al. 2017), should be tested for their performance in relation to unknown relationships using a similar study design.

Conclusion

We show that the steepness-slope suffers less from unknown relationships when calculated based on the normalized Average Dominance Index than the normalized David's Score. Specifically, the performance of the two indices is almost identical when unknown relationships are absent, but with increasing percentages of unknown relationships the hierarchical steepness is substantially less biased and its overall standard error is substantially smaller when it is calculated with the normalized Average Dominance Index than with the normalized David's Score.

Author Contributions

Tommaso Saccà: Conceptualization, Software, Formal analysis, Writing - Original Draft. **Gerrit Gort:** Supervision, Writing - Review & Editing. **Erica van de Waal:** Resources, Writing - Review & Editing. **Charlotte K Hemelrijk:** Supervision, Writing - Review & Editing.

Acknowledgments

This study was financed by grants to E.W. from the Swiss National Science Foundation (31003A_159587 and PP00P3_198913) and the Branco Weiss Fellowship Society in Science.

References

Altmann, J. (1974). Observational study of behavior: Sampling methods. *Behaviour*, *49*, 227–267. https://doi.org/10.1163/156853974 X00534

Balasubramaniam, K. N., Berman, C. M., De Marco, A., Dittmar, K., Majolo, B., Ogawa, H., Thierry, B., de Vries, H. (2013). Consistency of dominance rank order: A comparison of David's scores with I&SI and Bayesian methods in macaques. *American Journal of Primatology*, *75*(9), 959–971. https://doi.org/10.1002/ajp.22160

de Vries, H., Stevens, J. M. G., & Vervaecke, H. (2006). Measuring and testing the steepness of dominance hierarchies. *Animal Behaviour*, *71*(3), 585–592. https://doi.org/10.1016/j.anbehav.2005.05.015

Hemelrijk, C. K. (1999). An individual-orientated model of the emergence of despotic and egalitarian societies. *Proceedings of the Royal Society B: Biological Sciences*, *266*(1417), 361–369. https://doi.org/10.1098/rspb.1999.0646

Hemelrijk, C. K., Wantia, J., & Gygax, L. (2005). The construction of dominance order: Comparing performance of five methods using an individual-based model. *Behaviour*, *142*(8), 1043–1064. https://doi.org/10.1163/156853905774405290

Klass, K., & Cords, M. (2011). Effect of unknown relationships on linearity, steepness and rank ordering of dominance hierarchies: Simulation studies based on data from wild monkeys. *Behavioural Processes*, *88*(3), 168–176. https://doi.org/10.1016/j.beproc.2011.09.003

Neumann, C., & Kulik, L. (2020). *EloRating: Animal Dominance Hierarchies by Elo Rating*. Retrieved from https://cran.r-project.org/package=EloRating

R Core Team. (2021). *R: A language and environment for statistical computing*. Retrieved from https://www.r-project.org/

Sánchez-Tójar, A., Schroeder, J., & Farine, D. R. (2018). A practical guide for inferring reliable dominance hierarchies and estimating their uncertainty. *Journal of Animal Ecology*, *87*(3), 594–608. https://doi.org/10.1111/1365-2656.12776

Shizuka, D., & McDonald, D. B. (2012). A social network perspective on measurements of dominance hierarchies. *Animal Behaviour*, *83*(4), 925–934. https://doi.org/10.1016/j.anbehav.2012.01.011

Strauss, E. D., & Holekamp, K. E. (2019). Inferring longitudinal hierarchies: Framework and methods for studying the dynamics of dominance. Journal of Animal Ecology, 88(4), 521–536. https://doi.org/10.1111/1365-2656.12951

van Schaik, C. P. (1989). The ecology of social relationships amongst female primates. In V. Standen & R. Foley (Eds.), *Comparative Socioecology of Mammals and Man* (pp. 195–218). Oxford: Blackwell scientific.

Vehrencamp, S. L. (1983). A model for the evolution of despotic versus egalitarian societies. *Animal Behaviour*, *31*(3), 667–682. https://doi.org/10.1016/S0003-3472(83)80222-X

Table 1

Two fictional competitive interaction matrices for four individuals.

Interaction matrix I					**Interaction matrix II**				
	A	**B**	**C**	**D**		**A**	**B**	**C**	**D**
A	-	2	2	2	**A**	-	0	4	4
B	2	-	2	2	**B**	0	-	0	0
C	2	2	-	5	**C**	2	2	-	5
D	2	2	4	-	**D**	2	2	4	-

Two fictional competitive interaction matrices for four individuals (A, B, C and D) with winners in rows and losers in columns. Left: egalitarian case of A and B (same number of wins and losses in all dyads they are involved in); right: non-egalitarian case of A and B (unknown relationships between A and B, A wins all interactions, B loses all interactions.)

Table 2

Summary of results for the Monte Carlo simulations.

			normalized David's Score		**Kendall's Rank Correlation**	
Group (size)	Original # of unknown relationships (total, %)	Original steepness	Median steepness (% change) at 25% unkn. relationships	Median steepness (% change) at 66.7% unkn. relationships	Average Tau (± SD)	Wilcoxon on signed rank test P-value
KB2017 (6)	1 (15, 7)	0.68	0.46 (-31.8)[a]	0.15 (-77.8)	-0.999 ± 0.008	< 0.001
AK2011 (7)	3 (21, 14)	0.55	0.45 (-17.5)[b]	0.13 (-76.0)	-0.998 ± 0.008	< 0.001
NH2016 (8)	3 (28, 11)	0.54	0.41 (-24.7)	0.12 (-77.5)[d]	-0.997 ± 0.004	< 0.001
NH2011 (10)	3 (45, 7)	0.75	0.52 (-30.9)[c]	0.14 (-80.6)	-0.999 ± 0.002	< 0.001
DomWorld example (10)	0 (45, 0)	0.74	0.45 (-39.2)[c]	0.12 (-83.9)	-1 ± 0.0004	< 0.001
			normalized Average Dominance Index			

KB2017 (6)	1 (15, 7)	0.97	0.99 (+1.96)[a]	1.19 (+22.6)	0.43 ± 0.34	< 0.001
AN2011 (7)	3 (21, 14)	0.87	0.87 (-0.09)[b]	1.09 (+25.1)	0.43 ± 0.38	< 0.001
NH2016 (8)	3 (28, 11)	0.85	0.85 (-0.54)	0.99 (+16.6)[d]	0.31 ± 0.38	< 0.001
NH2011 (10)	3 (45, 7)	0.93	0.93 (+0.04)[c]	1.09 (+17.2)	0.35 ± 0.38	< 0.001
DomWorld example (10)	0 (45, 0)	0.76	0.77 (+0.95)[c]	0.86 (+12.9)	0.32 ± 0.40	< 0.001

[a]Median steepness (% change) at 26.7% unknown relationships

[b]Median steepness (% change) at 23.8% unknown relationships

[c]Median steepness (% change) at 24.4% unknown relationships

[d]Median steepness (% change) at 67.9% unknown relationships

Results for the Monte Carlo simulations for one-fourth and two-third unknown relationships in the social interaction matrix and Kendall Rank statistics for the correlation between the steepness-slope and the percentage of unknown relationships.

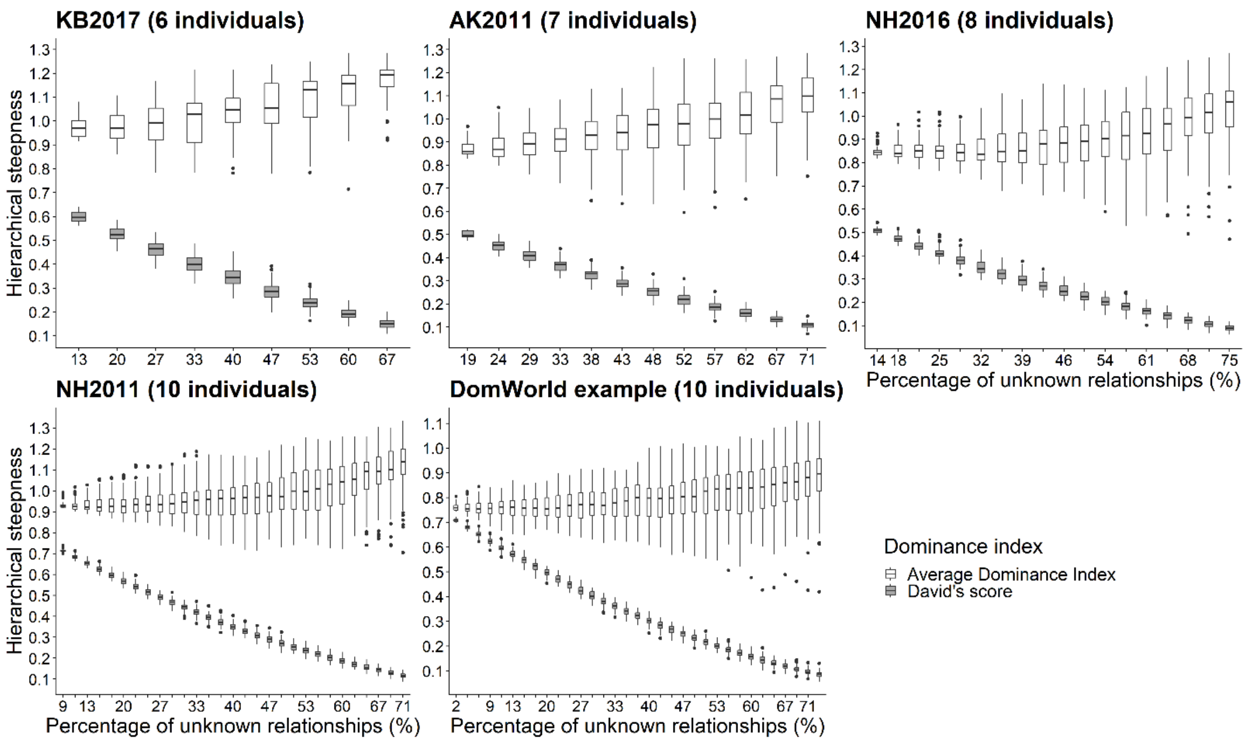

Figure 1: The steepness-slope (calculated with the Average Dominance Index and David's Score) and the percentage of unknown relationships in matrices of social interaction of four group-years of vervet monkeys (KB2017, AK2011, NH2016, NH2011) and one DomWorld run. Names of group-years of vervet monkeys indicate the abbreviation (two letters) of the name of the group and the year of observation. Group size is indicated in brackets for each group-year in empirical data of groups of vervets and for a DomWorld example run (see Table 2). For the boxplots, edges of boxes represent first and third quantile, internal lines represent median values, external lines show minimal and maximal values, while dots depict outliers. (2-column fitting image)

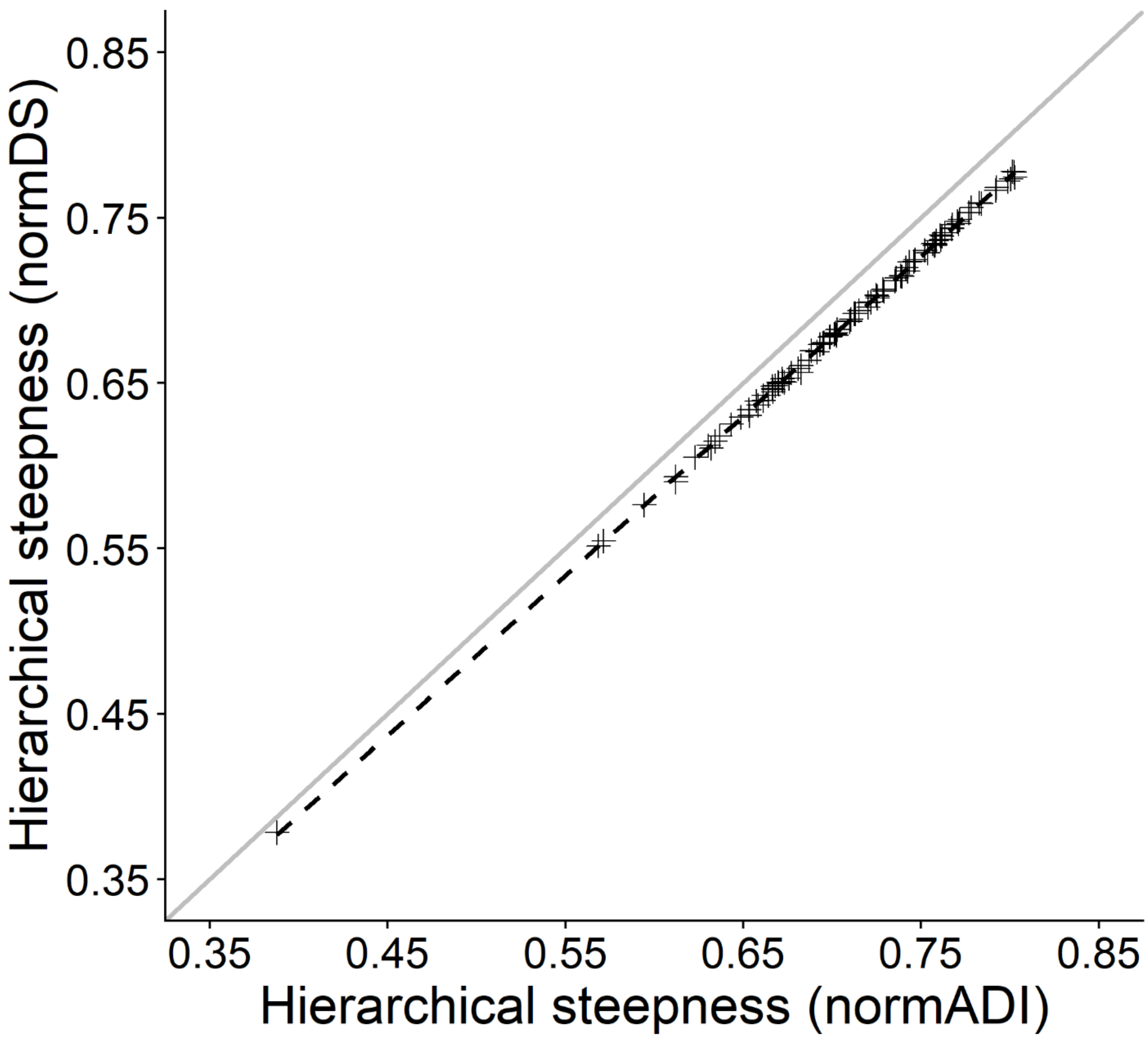

Figure 2: The steepness-slope calculated with the normalized Average Dominance Index (normADI, on the x axis) versus with the normalized David's Score (normDS, on the y axis) for 100 runs of DomWorld. The line in grey shows a regression line in which the use of the two indices results in identical values for hierarchical steepness. (single column fitting image)