PNAS
www.pnas.org

Supplementary Information for

Aggression heuristics underlie animal dominance hierarchies and provide evidence of group-level social information

Elizabeth A. Hobson, Dan Mønster, Simon DeDeo

Corresponding Author Elizabeth A. Hobson.
E-mail: elizabeth.hobson@uc.edu

This PDF file includes:

Supplementary text
Figs. S1.1 to S5.2
Tables S2.1 to S5.4
References for SI reference citations

Supporting Information Text

SI 1. Setting the regularization term for the measurement of PageRank in animal conflict

A basic step in the calculation of focus and average peak position is the estimation of the transition matrix, T_{ij}, a collection of probabilities, from the data. The "naive" way to estimate a probability of an event occurring from a finite number of observations is

$$\tilde{p} = \frac{n_i}{N}. \qquad [1]$$

While attractive in its simplicity, this estimator has a number of problems (see Ref. (1)); a Bayesian analysis leads to the correction

$$\hat{p} = \frac{n_i + \epsilon}{N + m\epsilon}, \qquad [2]$$

where m is the number of event types, and ϵ a regularization parameter (sometimes called a "teleportation term"). When ϵ is equal to unity, we have Laplace's rule; more generally, we can think of ϵ as parametrising a Dirichlet distribution that serves as the prior for the possible values of the underlying probabilities p (2, 3).

In the case we have here, T_{ij} is the estimate of the probability that i against j; by stipulation, the individual i can not aggress against itself. We can then adapt equation 2 to the estimate of the probability distributions in the matrix T_{ij}.

How do we choose ϵ? A natural way to do so is to learn ϵ from the data itself; we do so here using k-fold cross validation, with k set to five. For each dataset, in other words, we compute the probabilities T_{ij}, for some particular choice of ϵ, based on a randomly chosen sample of only 4/5 of the data. We then compute the log-probability per data-point of the remaining "held out" 1/5 of the data, n_{ij}, using those estimated T_{ij}s,

$$L(\epsilon) = \frac{1}{N_h} \sum_{i,j=1}^{N} n_{ij} \log T_{ij}(\epsilon), \qquad [3]$$

where N_h is the number of observations in the held-out set (*i.e.*, 1/5 of the total number of observations). In words, $L(\epsilon)$ is how well that particular choice of ϵ "predicts" the held-out data; the optimal choice of ϵ is that which best predicts.

We repeat this process many times, choosing a different hold-out set each time, to get an estimate of the average log-probability of the held-out data. We then choose ϵ to maximize this average of $L(\epsilon)$. Fig. S1.1 shows an example of this process for the data of Ref. (4). The peak of this function allows us to pick the optimal epsilon to be around 0.3 for this dataset, although values between 0.2 and 0.6 are largely indistinguishable. Fig. S1.2 shows a scatter plot of the L-maximizing ϵ for all 161 aggression matrices in our data, as a function of both total number of observations, and number of individuals.

We find that most matrices have optimal values of ϵ between 0.1 and 1.0, and that there is no strong correlation between optimal ϵ and system size or total number of observations. The average value of epsilon across all datasets is 0.694.

Little hinges on the exact value of ϵ; indeed, using the average value in place of the optimal choice for any particular dataset leads to an average (absolute value) shift in the focus measure of only 0.027, and in the average peak position of only 0.017; over our data, the two choices have a Pearson correlation of 0.95 (Focus) and 0.97 (Average Peak Position). Since finding the optimal ϵ is computationally intensive, and since the final results are largely insensitive to this choice, we suggest the average value, 0.694, is appropriate for ordinary use, and (for simplicity) we present our analyses here using this choice.

Fig. S1.1. Determining optimal ϵ through k-fold cross validation; an example of equation 3 applied to the data of Ref. (4). An ϵ value of approximately 0.3, in this case, best predicts held-out data, but a range of ϵ values between 0.1 and 1 perform similarly well.

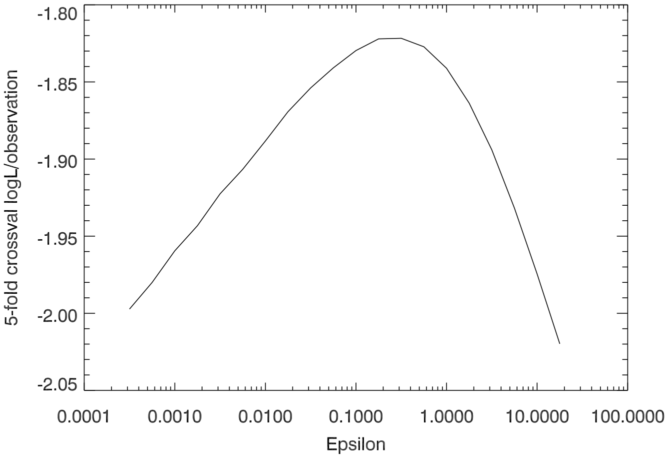
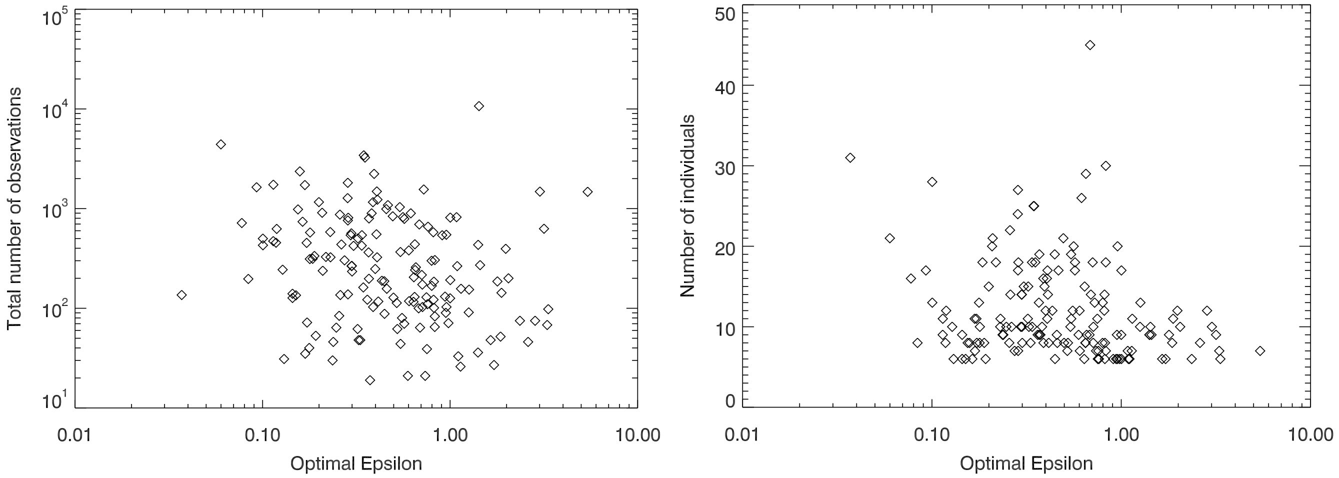


Fig. S1.2. A scatter plot of optimal epsilons found using equation 3, as a function of total number of observations (left), and total number of individuals in the data (right). The optimal value shows no strong trends with either variable; the average optimal value for ϵ is 0.694 and we use this for simplicity in the calculations in the main text.



SI 2. Observed summary data

Table S2.1. Observed social dominance pattern, focus, and position values for each empirical group in our dataset (focus and position values rounded). Social dominance patterns: downward heuristic (DH), close competitors (CC), bullying (BL), and undefined (UN). Records sorted by species name then by file name.

Species	Pattern	Focus	Position	File name
Addax nasomaculatus	CC	0.78	0.34	Reason1988-1.csv
Anas acuta	DH	0.69	0.36	Poisbleau2005-1b.csv
Anas platyrhynchos	DH	0.74	0.46	Poisbleau2005-1a.csv
Anas platyrhynchos	BL	0.45	0.38	Poisbleau2006-2a.csv
Anolis aeneus	CC	0.86	0.30	Stamps1978.csv
Antilocapra americana	DH	0.60	0.29	Bromley1991-1.csv
Antilocapra americana	DH	0.71	0.37	Fairbanks1994-5.csv
Antilocapra americana	DH	0.41	0.22	Fairbanks1994-6.csv
Antilocapra americana	BL	0.66	0.58	Fairbanks1994-7.csv
Bison bison	DH	0.44	0.30	Lott1979-1.csv
Bison bison	DH	0.64	0.34	Lott1987-1.csv
Bison bison	DH	0.24	0.28	Rutberg1986-2.csv
Branta bernicla	DH	0.55	0.31	Poisbleau2006-2b.csv
Callosciurus erythraeus	DH	0.62	0.37	Tamura1988-1.csv
Canis lupus	DH	0.55	0.36	Cafazzo2010-5.csv
Cercopithecus mitis	DH	0.78	0.38	Payne2003-1.csv
Cervus elaphus	DH	0.17	0.19	Appleby1983-1.csv
Chlorocebus aethiops	DH	0.71	0.42	Isbell1998-A.csv
Chlorocebus aethiops	DH	0.75	0.38	Struhsaker1967-6.csv
Chlorocebus aethiops	DH	0.60	0.34	Struhsaker1967-7.csv
Colobus polykomos	DH	0.48	0.42	Korstjens2002-1.csv
Columba livia	UN	-0.05	0.15	Masure1934-3.csv
Corvus monedula	DH	0.50	0.34	Roell1978-11.csv
Corvus monedula	DH	0.27	0.23	Tamm1977-1b.csv
Crocuta crocuta	DH	0.56	0.45	Frank1986-1.csv
Crocuta crocuta	DH	0.72	0.37	Holekamp1991-1.csv
Crocuta crocuta	DH	0.68	0.50	Holekamp1993-1a.csv
Crocuta crocuta	DH	0.74	0.29	Jenks1995-3.csv
Crocuta crocuta	BL	0.73	0.56	Tilson1984-1.csv
Cyanocitta cristata	DH	0.61	0.36	Tarvin1997-5.csv
Dinoponera quadriceps	DH	0.18	0.24	Monnin1999-1.csv
Eledone moschata	DH	0.48	0.20	Mather1985.csv
Emys orbicularis	DH	0.25	0.26	Rovero1999-2a.csv

Continued on next page

Species	Pattern	Focus	Position	File name
Emys orbicularis	DH	0.12	0.18	Rovero1999-2b.csv
Equus caballus	CC	0.74	0.39	CluttonBrock1976-3.csv
Equus caballus	CC	0.78	0.35	Ellard1989-3.csv
Equus caballus	BL	0.71	0.44	Heitor2006-3.csv
Equus caballus	DH	0.49	0.29	Heitor2010-2.csv
Equus caballus	CC	0.65	0.38	Wells1979-3.csv
Erythrocebus patas	BL	0.46	0.58	Isbell1998-B.csv
Felis catus	DH	0.66	0.45	Bonanni2007-2.csv
Felis catus	BL	0.50	0.54	Natoli1991-2.csv
Fringilla coelebs	CC	0.86	0.37	Marler1955b.csv
Galaxias argenteus	DH	0.53	0.37	David2003-2a.csv
Galaxias argenteus	DH	0.53	0.44	David2003-2b.csv
Gorilla beringei	DH	0.37	0.24	Robbins2008-2.csv
Gorilla gorilla	DH	0.24	0.33	Scott1999-2b.csv
Gorilla gorilla	DH	0.13	0.19	Scott1999-2c.csv
Haemorhous mexicanus	DH	0.72	0.40	Thompson1960-JJ59A.csv
Haemorhous mexicanus	DH	0.71	0.41	Thompson1960-JJ59B.csv
Haemorhous mexicanus	DH	0.74	0.39	Thompson1960-JJ59C.csv
Haemorhous mexicanus	DH	0.52	0.31	Thompson1960-ND54.csv
Haemorhous mexicanus	DH	0.37	0.21	Thompson1960-ND57.csv
Haliaeetus albicilla	DH	0.36	0.29	Kolodziejczyk2005-1.csv
Junco hyemalis	DH	0.24	0.31	Yasukawa1983-1a.csv
Junco hyemalis	DH	0.27	0.30	Yasukawa1983-1b.csv
Junco hyemalis	DH	0.08	0.22	Yasukawa1983-2a.csv
Junco hyemalis	DH	-0.05	0.08	Yasukawa1983-2b.csv
Junco hyemalis	DH	0.64	0.37	Yasukawa1983-5a.csv
Junco hyemalis	DH	0.58	0.29	Yasukawa1983-5b.csv
Junco hyemalis	BL	0.53	0.43	Yasukawa1983-6a.csv
Junco hyemalis	DH	0.57	0.24	Yasukawa1983-6b.csv
Lampropholis guichenoti	DH	0.35	0.22	Torr1996-1.csv
Leptothorax sp.	DH	0.35	0.21	Ortius1995-2a.csv
Loxodonta africana	CC	0.77	0.32	Archie2006-A.csv
Loxodonta africana	DH	0.67	0.42	Archie2006-AA.csv
Loxodonta africana	DH	0.50	0.36	Archie2006-CB.csv
Loxodonta africana	DH	0.68	0.37	Archie2006-FB.csv
Loxodonta africana	DH	0.57	0.35	Archie2006-GB.csv
Loxodonta africana	DH	0.52	0.51	Archie2006-JA.csv
Loxodonta africana	DH	0.74	0.40	Archie2006-OA.csv
Loxodonta africana	BL	0.91	0.55	Archie2006-PC.csv
Loxodonta africana	DH	0.77	0.36	Archie2006-SI.csv
Loxodonta africana	DH	0.50	0.42	Wittemeyer2007-1.csv
Macaca arctoides	DH	0.40	0.41	Richter2009-1.csv
Macaca fascicularis	DH	0.36	0.33	deWaal1977-1.csv
Macaca fascicularis	BL	0.49	0.37	deWaal1977-2.csv
Macaca fascicularis	DH	0.65	0.42	Sterck1997-4KB.csv
Macaca mulatta	DH	0.71	0.44	deWaal1985-1.csv
Macaca mulatta	DH	0.13	0.18	Varley1966-1.csv
Macaca thibetana	DH	0.60	0.41	Berman2004-B.csv
Mandrillus sphinx	CC	0.64	0.30	Setchell2005-3.csv
Mandrillus sphinx	DH	0.58	0.36	Setchell2005-4.csv
Mareca penelope	DH	0.53	0.39	Poisbleau2005-1c.csv
Melanochromis auratus	DH	0.27	0.26	Nelissen1985-2.csv
Meles meles	DH	0.60	0.46	Hewitt2009-P2005.csv
Meles meles	DH	0.48	0.37	Hewitt2009-PO2004.csv
Meles meles	DH	0.50	0.42	Hewitt2009-SH1995.csv
Mustelus canis	DH	0.73	0.42	Allee1954-3.csv
Myiopsitta monachus	CC	0.65	0.30	monkparakeet.g1Q2to4.csv
Myiopsitta monachus	CC	0.65	0.27	monkparakeet.g2.Q2to4.noNBB.csv
Neoponera villosa	CC	0.75	0.36	Trunzer1999-1.csv
Notamacropus parryi	DH	0.54	0.30	Kaufmann1974-7.csv

Continued on next page

Species	Pattern	Focus	Position	File name
Nymphicus hollandicus	DH	0.31	0.29	Seibert2001-3.csv
Odocoileus hemionus	DH	0.69	0.37	Koutnik1981-3.csv
Odocoileus virginianus	DH	0.77	0.44	Collias1950-1.csv
Oncorhynchus masou	CC	0.83	0.39	Nakano1994-1.csv
Oreamnos americanus	DH	0.73	0.43	Cote2000-4.csv
Oreamnos americanus	DH	0.56	0.41	Fournier1995-3.csv
Osphranter rufus	DH	0.53	0.39	Russell1970-2a.csv
Osphranter rufus	DH	0.57	0.30	Russell1970-2b.csv
Ovis canadensis	DH	0.63	0.35	Hass1991-2.csv
Ovis canadensis	CC	0.48	0.17	Zine2000-1.csv
Pachycondyla sp.	DH	0.69	0.42	Ito1993-2a.csv
Pachycondyla sp.	DH	0.62	0.40	Ito1993-3a.csv
Pachycondyla sp.	DH	0.70	0.31	Ito1993-4a.csv
Pan paniscus	BL	0.27	0.42	Paoli2006-2.csv
Pan paniscus	BL	0.65	0.54	Vervaecke2000-2.csv
Pan troglodytes	DH	0.53	0.42	Murray2007-3.csv
Pan troglodytes	CC	0.70	0.36	Wittig2003-1.csv
Papio cynocephalus	DH	0.19	0.13	Cheney1977-B.csv
Papio cynocephalus	DH	0.85	0.32	Hausfater1975-10.csv
Papio cynocephalus	CC	0.81	0.23	Hausfater1975-11.csv
Papio cynocephalus	CC	0.88	0.29	Hausfater1975-6.csv
Papio cynocephalus	CC	0.80	0.38	Hausfater1982-1.csv
Papio cynocephalus	CC	0.73	0.38	Hausfater1982-2.csv
Papio cynocephalus	DH	0.61	0.40	Lee1979-1.csv
Papio cynocephalus	DH	0.52	0.41	Lee1979-2.csv
Papio cynocephalus	DH	0.64	0.31	McMahan1984-1.csv
Papio cynocephalus	CC	0.78	0.37	Samuels1987-2.csv
Parahyaena brunnea	DH	0.75	0.43	Owens1996-1.csv
Passer domesticus	BL	0.41	0.36	Moller1987-1.csv
Passer domesticus	BL	0.70	0.44	Moller1987-2.csv
Passer domesticus	DH	0.73	0.48	Moller1987-3.csv
Passer domesticus	DH	0.13	0.25	Solberg1997-1.csv
Passer domesticus	DH	0.34	0.21	Solberg1997-2.csv
Passer domesticus	CC	0.67	0.01	Solberg1997-3.csv
Phasianus colchicus	DH	0.88	0.42	Collias1951-3.csv
Phasianus colchicus	DH	0.74	0.26	Collias1951-4.csv
Phasianus colchicus	BL	0.53	0.46	Collias1951-5.csv
Phoca vitulina	DH	0.56	0.40	Sullivan1982-4.csv
Poecile atricapilla	DH	0.62	0.41	Hartzler1970-1.csv
Poecile atricapilla	CC	0.85	0.39	Smith1976-1.csv
Poecile montanus	DH	0.57	0.35	Lahti1994-A.csv
Poecile montanus	DH	0.42	0.34	Lahti1994-C.csv
Polistes canadensis	DH	0.37	0.26	West-Eberhard1986-5.csv
Potamonautes perlatus	DH	0.54	0.38	Somers1998-2a.csv
Potamonautes perlatus	DH	0.53	0.37	Somers1998-2b.csv
Potamonautes perlatus	DH	0.79	0.45	Somers1998-2c.csv
Protomognathus americanus	DH	0.80	0.28	Blatrix2004-2.csv
Protomognathus americanus	DH	0.72	0.39	Blatrix2004-3.csv
Protomognathus americanus	DH	0.74	0.38	Blatrix2004-5.csv
Protomognathus americanus	DH	0.72	0.34	Blatrix2004-6.csv
Quiscalus major	DH	0.65	0.40	Post1992-1.csv
Rangifer tarandus	DH	0.61	0.34	Barette1986.csv
Rangifer tarandus	DH	0.51	0.34	Hirotani1994-1.csv
Salvelinus leucomaenis	DH	0.60	0.36	Nakano1995-2.csv
Sapajus apella	DH	0.79	0.47	Izar2006-2.csv
Sauromalus ater	DH	0.30	0.20	Prieto1978-1.csv
Sciurus aberti	DH	0.73	0.34	Farentinos1972-D.csv
Semnopithecus entellus	DH	0.60	0.31	Lu2008-1c.2.csv
Serinus canaria	DH	0.04	0.10	Shoemaker1939-1.csv
Sphyrna tiburo	DH	0.44	0.27	Myrberg1974-17.csv

Continued on next page

Species	Pattern	Focus	Position	File name
Stegastes partitus	CC	0.82	0.30	Myrberg1972-3.csv
Streptopelia risoria	DH	0.11	0.15	Bennett1939-2.csv
Tadorna tadorna	DH	0.63	0.38	Patterson1977-6.csv
Trachypithecus phayrei	DH	0.24	0.29	Koenig2004-1b.csv
Tragelaphus angasii	DH	0.75	0.34	Collias1950-2.csv
Tropheus moorii	DH	0.63	0.39	Kohda1991-2.csv
Vollenhovia nipponica	DH	0.08	0.10	Satoh-C.csv
Vollenhovia nipponica	DH	0.01	0.11	Satoh-D.csv
Vollenhovia nipponica	DH	0.38	0.25	Satoh-E.csv
Xerus rutilus	DH	0.70	0.41	Oshea1976-1.csv
Zonotrichia leucophrys	DH	0.13	0.10	Parsons1980-2a.csv
Zonotrichia leucophrys	BL	0.60	0.41	Slotow1993-1.csv
Zonotrichia querula	BL	0.72	0.51	Watt1986-1a.csv
Zonotrichia querula	DH	0.72	0.41	Watt1986-1b.csv
Zonotrichia querula	DH	0.69	0.38	Watt1986-1c.csv
Zonotrichia querula	DH	0.37	0.38	Watt1986-1d.csv
Zonotrichia querula	DH	0.59	0.41	Watt1986-1e.csv
Zosterops lateralis	DH	0.32	0.18	Kikkawa1980-1.csv
Zosterops lateralis	DH	0.21	0.15	Williams1972-1.csv

SI 3. Characteristics of structured aggression

Most of the animal social groups in the empirical dataset had well-structured dominance hierarchies. Groups generally had real focus values consistent with low levels of randomly-directed aggression: 46% of groups ($N = 79$) were most similar to modelled data with 10% or less randomly-directed events (≤ 0.1); 63% of groups ($N = 109$) were most similar to modelled data with 20% or less randomly-directed events (≤ 0.2) and 90% of groups ($N = 155$) was most similar to modeled data with 60% or less randomly-directed events.

Only 14 groups (8%) had focus values most similar to modelled data with 80% or greater randomly-directed events (≥ 0.8); of these, none were categorized as a close competitor or bullying social dominance type.

Only 7 groups (4%) had focus values closest to modelled data with totally random aggression, which corresponds roughly to previous results with this dataset which found over-representation of transitive configurations, an indication of structured hierarchies, in all but 3% of groups (5).

SI 4. Ruling out factors affecting observed social dominance patterns

Factors other than phylogenetic relatedness could affect which social dominance pattern a social group used. Here we investigate whether the occurrence of social dominance patterns can be explained by either the conditions under which groups were observed (if data were collected from a natural population or one held in captivity) or the number of individuals in the group (for example, if only small groups showed evidence of a certain social dominance pattern).

We use the same randomized data as in the main text, where we tested whether randomizing social dominance types across orders would differ from observed occurrences. Here, we summarize by (1) observation conditions and (2) group size. If the observed pattern occurrences fall within the distributions of the randomized strategies, then we can conclude that we do not have any evidence that the factors explain the observed distributions.

We find that in almost all cases, the observed occurrences of social dominance occurrences fall within the distributions of the randomized strategies, demonstrating that neither observation conditions (Fig. S4.1) nor group size (Fig. S4.2) can be used to explain the observed social dominance patterns. It is important to note that these observed datasets are likely biased towards smaller group sizes and conditions where data collection is more manageable.

Fig. S4.1. The conditions under which groups were observed had little to no relationship with social dominance patterns. For groups observed in natural or captive conditions, the observed number of groups with each social dominance pattern (solid vertical lines) overlaps with the number of groups with each pattern when patterns are randomly allocated (shaded areas, density estimates) for each of the three main social dominance patterns (a) downward heuristic, (b) close competitors, and (c) bullying.

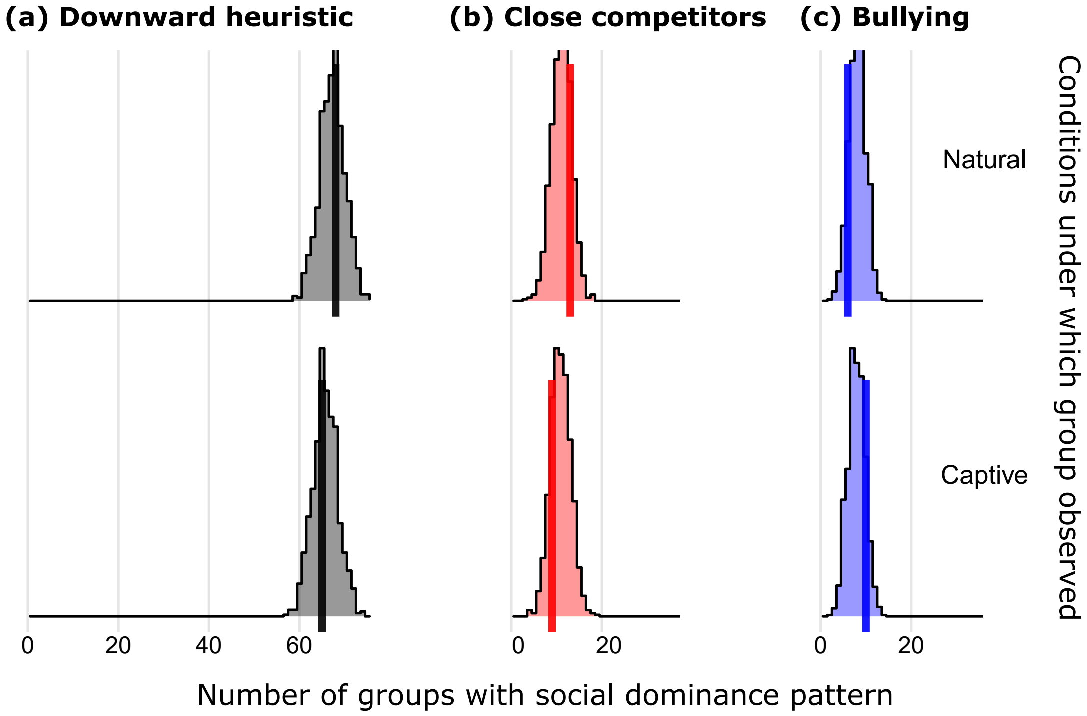
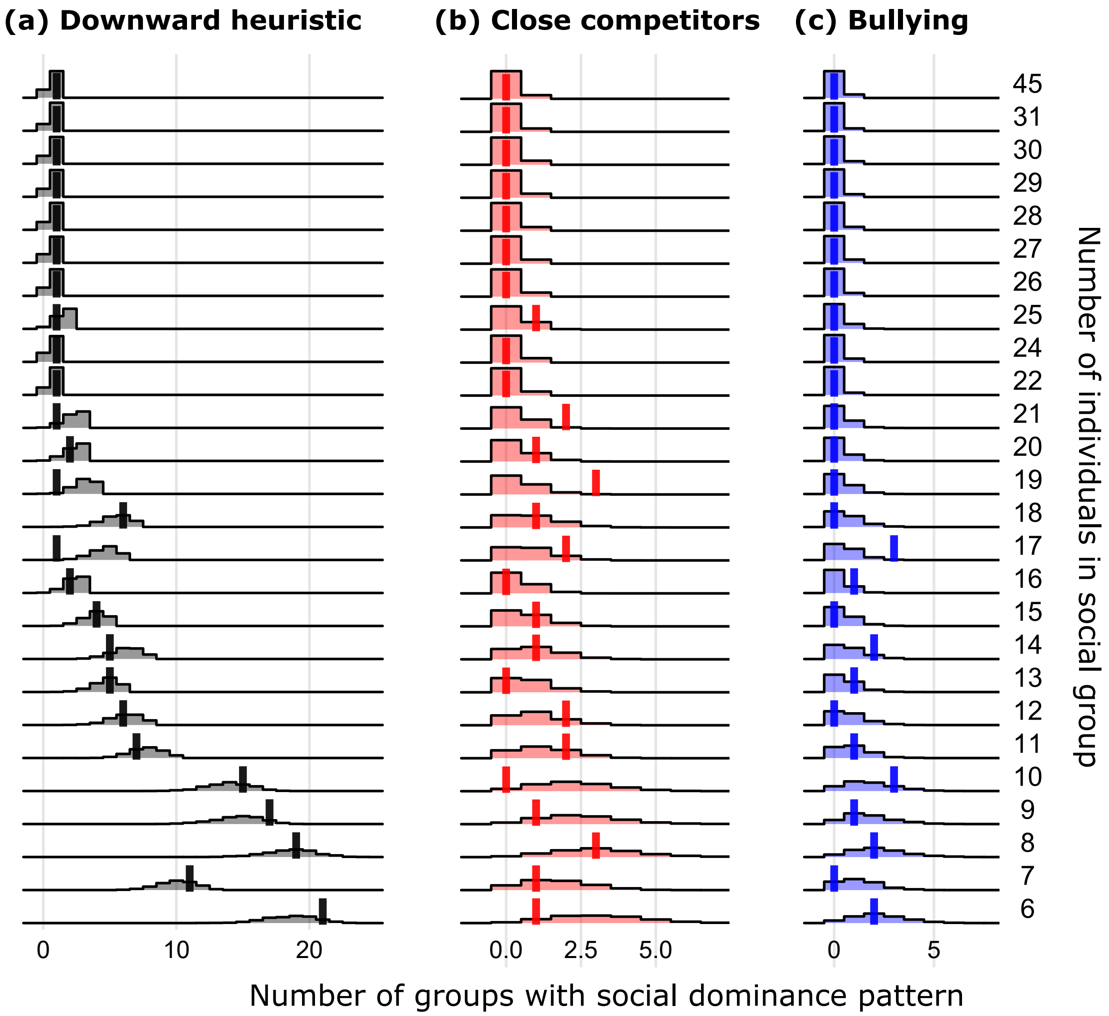


Fig. S4.2. The size of observed social groups had little to no relationship with social dominance patterns. Across different group sizes, the number of individuals in the group did not consistently explain the occurrence of each social dominance pattern (solid vertical lines): the observed number of groups overlaps with the number of groups with each pattern when patterns are randomly allocated (shaded areas, density estimates) for each of the three main social dominance patterns (a) downward heuristic, (b) close competitors, and (c) bullying.



SI 5. The emergence of social dominance patterns with winner and loser effects

In this section, we investigated whether social dominance patterns can emerge from systems in which individuals only have information about themselves and the outcomes of their past interactions, rather than any information about their own rank or the identities or ranks of their opponents. We use this approach as a test of our logic, that more information-rich social dominance strategies only reliably emerge when individuals have some information not just about themselves, but about others in their group. We constructed several variants of a winner/loser effects model and then tested how often different social dominance patterns emerged. In the model, each individual only has access to its own win/loss record, and can only adjust its behavior based on outcomes of these events — individuals do not have any information about which other individuals they interacted with or which individuals they have won or lost against. We draw the aggressiveness, i.e., the initial attack probability, of each individual from a uniform distribution on the interval $[0, 1]$.

We modeled nine variants of this model (see Table S5.1). We investigated a winner-effect only model, a loser-effect only model, a mixed winner and loser effect model, and a model with neither a winner or loser effect. Each of these models was further investigated by incorporating each winner/loser effect as either a transient effect or persistent effect on individual behavior. We used estimates of winner and loser effect strengths from the literature (6) plus a more extreme and more moderate value for comparison.

Table S5.1. The nine combinations of loser and winner effects with both transient and persistent effects. The NN combination has neither a loser nor a winner effect.

	Winner effect		
Loser effect	None	Transient	Persistent
None	NN	NT	NP
Transient	TN	TT	TP
Persistent	PN	PT	PP

Since individuals are not aware of their own and others' rank, the baseline "strategy" will be to select a random opponent and decide whether to attack based on an inherent level of aggressiveness corresponding to an initial attack probability $p_{i,0}$ for individual i. Computationally it is more convenient to work with odds than probabilities for the implementation of effects of past performance. So instead of the probability, we can use the initial odds of attacking, viz.

$$O_{i,0} = \frac{p_{i,0}}{1 - p_{i,0}}$$

To account for an individual's own past performance we implemented the following simple loser and winner effects, summarized in Table S5.2.

Table S5.2. Summary of effect types and the ways these effects are incorporated into the model behavior.

Effect type	**Model behavior**
Transient loser effect	Changes the odds of attacking to $O_i = O_{i,0} \cdot \alpha_{L,t}$, where $\alpha_{L,t} < 1$, if the latest encounter resulted in a loss
Transient winner effect	Changes the odds of attacking to $O_i = O_{i,0} \cdot \alpha_{W,t}$, where $\alpha_{W,t} > 1$, if the latest encounter resulted in a win
Persistent loser effect	Changes the odds of attacking by a factor $\alpha_{L,p} < 1$ every time an individual loses
Persistent winner effect	Changes the odds of attacking by a factor $\alpha_{W,p} > 1$ every time an individual wins

Note that the transient effects are not cumulative, *i.e.*, they change the odds relative to the initial odds for one encounter only. The persistent effects, on the other hand, change the odds relative to what they were in the previous encounter, *e.g.*, two consecutive losses will multiply (decrease) O_i by a factor $\alpha_{L,p}^2$. To avoid unrealistically low or high odds of attacking a lower limit of $\alpha_{\min} = 10^{-3}$ and an upper limit of $\alpha_{\max} = 10^3$ are imposed.

In the simulations we have attributed random initial aggression levels, $p_{i,0}$, but fixed the α factors controlling the winner and loser effects to the same value across all individuals.

We used three different sets of α values in our simulations: (a) **extreme values** ($\alpha_L = 0.05$ and $\alpha_W = 2.7$), (b) **realistic values** ($\alpha_L = 0.18$, $\alpha_W = 1.87$, which are the pooled estimates from the meta-analysis by Rutte et al. (6)), and (c) **moderate values** ($\alpha_L = 0.7$ and $\alpha_W = 1.3$). As a control we ran simulations with the same parameters but without any of the winner or loser effects. In all cases, we used a group size of $N = 10$, and let the simulation run for a total of 1000 attacks.

Some parts of the structure exhibited by our generative models, using only winner and/or loser effects, differed from the structure of the observed social groups in our hierarchy dataset (Section SI 3). For example, many more of the generated

Table S5.3. Frequency of the observed social dominance patterns for the each of the three parameter sets and for both transient and persistent winner and loser effects: (a) extreme values, $\alpha_L = 0.05$ and $\alpha_W = 2.7$, (b) realistic values, $\alpha_L = 0.18$, $\alpha_W = 1.87$ (which are the pooled estimates from the meta-analysis by Rutte et al. (6)), (c) moderate values, $\alpha_L = 0.7$ and $\alpha_W = 1.3$, and (d) the sum total of runs across all models a-c that exhibited each social dominance pattern.

	(a) Extreme α values		(b) Rutte et al. α values		(c) Moderate α values		(d) Summary (all α values)	
	Transient	Persistent	Transient	Persistent	Transient	Persistent	Transient	Persistent
Social dominance pattern	*n*	*n*	*n*	*n*	*n*	*n*	*n*	*n*
Bully	0	0	0	2	0	0	0	2
Close competitors	5	4	2	3	7	7	14	14
Downward heuristic	87	90	83	90	80	83	250	263
Undefined	8	6	15	5	13	10	36	21
All	100	100	100	100	100	100	300	300

groups show focus and/or position values less than zero (Fig. S5.1), which was not often observed in the empirical groups. To highlight this difference, we examined one of the parameter sets (the "realistic" values of $\alpha_L = 0.18$, $\alpha_W = 1.87$ from Rutte et al., for persistent effects). The plot of the full data set with all 100 groups is shown on the left in Figure S5.2 and the censored data including only non-negative values is shown on the right.

Using this censored dataset (which excludes artificial groups which which showed characteristics inconsistent with our empirical datasets), we found that the two more complex social dominance patterns were rarely produced by winner and/or loser effects. Transient effects resulted in 0% bully and $< 2\%$ close competitor patterns and persistent effects resulted in $< 3\%$ bully and just over 1% close competitor patterns (Table S5.4). For uncensored data, bully and close competitor patterns were still rare: across all combinations of α_L and α_W (including groups modeled with moderate, realistic, and extreme values) transient effects produced 0 groups with a bullying pattern and 4.7% of groups (14 groups total) with a close competitor pattern (Table S5.3d). Even considering the most extreme α_L and α_W alone (Table S5.3a), our models produced no groups with a bullying pattern and just 5% (transient effects) and 4%(persistent effects) with close competitors patterns.

These results show that bullying or close competitors social dominance patterns are rarely produced through winner and/or loser effects alone. These results provide additional justification for our argument that additional social information, beyond an individual merely remembering the outcomes of its own encounters, is generally needed to produce the more complex social dominance patterns.

Table S5.4. Prevalence of the different social dominance patterns in the censored data set for the Rutte et al. values of α_W and α_L (values summarized from Fig. S5.2, right (persistent effects), with additional summaries for transient effects). The equivalent numbers for the uncensored data are shown in Table S5.3b.

	Number of groups		Percent of groups	
Social dominance pattern	Transient	Persistent	Transient	Persistent
Bully	0	2	0.00	2.27
Close competitors	1	1	1.82	1.14
Downward heuristic	54	84	98.18	95.45
Undefined	0	1	0.00	1.14
All	55	88	100.00	100.00

Fig. S5.1. The distribution of focus and position for each of the 100 groups simulated for each combination of parameters (α_W, α_L) and each of the two types of winner and loser effects (transient and persistent). The α_W, α_L combinations used are consistent with Table S5.3: extreme values where $\alpha_L = 0.05$ and $\alpha_W = 2.7$, Rutte et al. indicates "realistic" values where $\alpha_L = 0.18$, $\alpha_W = 1.87$ (which are the pooled estimates from the meta-analysis by Rutte et al. (6)), and moderate values where $\alpha_L = 0.7$ and $\alpha_W = 1.3$). The group-level social dominance pattern is indicated by the color of the point representing each group.

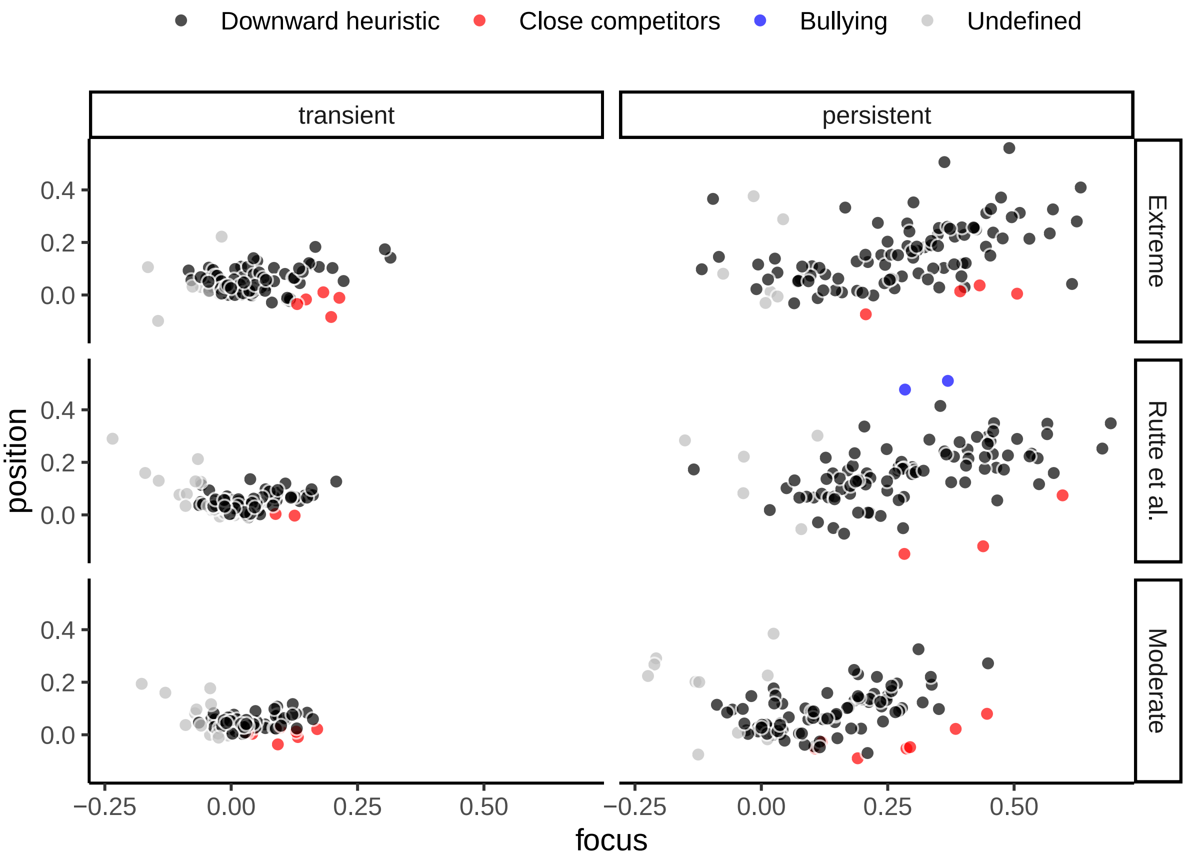
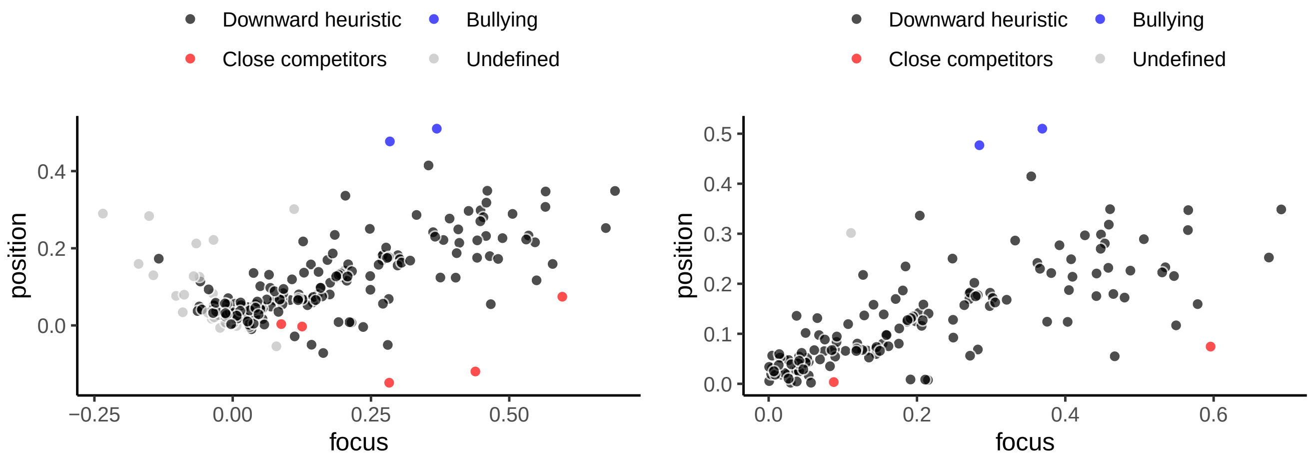


Fig. S5.2. Focus and position values for the simulated groups using the "realistic" values of $\alpha_L = 0.18$, $\alpha_W = 1.87$ (the pooled estimates from the meta-analysis by Rutte et al. (6)). *Left:* All 100 groups. *Right:* Same data, but subset to show only groups with focus ≥ 0 and position ≥ 0.



References

1. E Jaynes, G Bretthorst, *Probability Theory: The Logic of Science.* (Cambridge University Press), (2003).
2. DH Wolpert, DR Wolf, Estimating functions of probability distributions from a finite set of samples. *Phys. Rev. E* **52**, 6841 (1995).
3. DH Wolpert, S DeDeo, Estimating functions of distributions defined over spaces of unknown size. *Entropy* **15**, 4668–4699 (2013).
4. W Allee, J Dickinson Jr, Dominance and subordination in the smooth dogfish mustelus canis (mitchill). *Physiol. Zool.* **27**, 356–364 (1954).
5. D Shizuka, DB McDonald, The network motif architecture of dominance hierarchies. *J. The Royal Soc. Interface* **12** (2015).
6. C Rutte, M Taborsky, MW Brinkhof, What sets the odds of winning and losing? *Trends Ecol. & Evol.* **21**, 16–21 (2006).