

*Article*

## How to be a specialist? Quantifying specialisation in pollination networks

Carsten F. Dormann

Department Computational Landscape Ecology, Helmholtz Centre for Environmental Research-UFZ, 04318 Leipzig, Germany

E-mail: carsten.dormann@ufz.de

*Received 7 January 2011; Accepted 25 January 2011; Published online 7 May 2011*

IAEES

### Abstract

The analysis of ecological networks has gained a very prominent foothold in ecology over the last years. While many publications try to elucidate patterns about the networks, others are primarily concerned with the role of specific species in the network. The core challenge here is to tell specialists from generalists. While field data and observations can be used to directly assess specialisation levels, the indirect way through networks is burdened with problems.

Here, I review eight measures to quantify specialisation in pollination networks (degree, node specialisation, betweenness, closeness, strength, pollination support, Shannon's H and discrimination d'), the first four being based on binary, the others on weighted network data. All data and R-code are available as supplement and can be applied beyond pollination networks.

The indices convey different concepts of specialisation and hence quantify different aspects. Still, there is some redundancy, with node specialisation and closeness quantifying the same properties, as do degree, betweenness and Shannon's H.

Using artificial and real network data, I illustrate the interpretation of the different indices and the importance of using a null model to correct for expectations given the different observed frequencies of interactions. For a well-described network the distributions of specialisation values do not differ from null model expectations for most indices.

Finally, I investigate the effect of cattle grazing on the specialisation of an important pollinator in eight replicated pollination networks as an illustration of how to employ the specialisation indices, null models and permutation-based statistics in the analysis of specialisation in pollination networks.

**Keywords** bipartite network; degree; discrimination; node specialisation index; pollinator; pollination service index; strength; two-mode network.

### 1 Introduction

“All animals are equal, but some animals are more equal than others.” (Orwell, 1945). This certainly describes well the current paradigm of pollinator generalisation: Most pollinators seem to be generalists, with some spectacular, but rare, exceptions of high specialisation (as reviewed, e.g., in Waser, 2006). At the same time, recent re-analyses of pollination networks indicate that some degree of specialisation is common, both at the network level and for the pollinators themselves (Jordano, 1987; Vázquez and Simberloff, 2002; Vázquez and Aizen, 2004; Vázquez and Aizen, 2006; Bascompte et al., 2006; Blüthgen et al., 2007). However, there are

different ways to measure specialisation of pollinators, some more obvious than others. The reason is that there are also more than one way to think about specialisation. At the first level, a specialist pollinator can be defined as a pollinator that visits (and hence pollinates) only a limited number of plant species. However, one can also consider each pollinator in its community context, the second level. Then a specialist could be viewed as a species that has little overlap with the preferences of other pollinators in the community, i.e. one that has different flower preferences from the others. It will thus be less redundant than a generalist, and its loss from a community has potentially greater effect on the plant community. Finally, we may want to consider specialisation at the third level as the outcome of evolutionary diversification, from a hypothetical generalist pollinator community to an optimum distribution of specialisation. In this case, specialisation would give us a way to compare, across networks, the deviation from a network of only generalists.

Given this range of possible questions behind pollination specialisation studies, it is no surprise to find a range of specialisation indices with different intentions. Following the first definition, a specialist pollinator can be identified by the number of links (a species' *degree*). Specialists have lower *degree* than generalists. For the second definition, indices use information on the proportion of visits to other plant species (*strength*), on the position of a pollinator in the network (*node specialisation*, *betweenness*) or they are related to the pollination service a species actually provides. For the third definition, a network-independent measure is required, based for example on how much a pollinator discriminates against which plants are on offer. Below, I described in more detail eight specialisation indices that have been used to quantify specialisation in any of the three definitions.

Degree, node specialisation, betweenness and closeness centrality are binary indices, i.e. they make no use of the number of visits recorded for each interaction. In the extreme this means that even if a pollinator visits a single plant species in 90% of its visits, and distributes the other 10% over all plant species in the network, this species will have a high degree and hence count as a generalist. That is not intuitive, although it is consistent with a strict definition of specialisation. While degree counts the number of plants a pollinator interacts with, the other three binary indices are based on the position (as given e.g. by path length distances) of the pollinator in the network.

In a weighted network also information on which proportion of visits are paid to the different plant species are used. Strength, pollination service (*PSI*), Shannon's host diversity  $H$  and  $d'$  and use this quantitative information to calculate specialisation. They each address slightly different questions. A species' strength describes how much the plant community depends on the visits by this particular species (Jordano, 1987). The logic of *PSI* is that a pollinator that visits many different species will also deliver diluted pollen to any of the target species. Its value for this specific plant may thus be compromised. *PSI* attempts to quantify the service of a pollinator for all plants in the network. Shannon's host diversity is a weighted version of degree, down-weighting rare visitations.  $D'$ , on the other hand, corrects for different abundances of plants and pollinators. Its developers (Blüthgen et al., 2006) argue that a pollinator that makes use of the most common flower source should not be deemed a specialist. Rather, we should consider the discrimination between what is on offer and which plants the pollinator visits as important.

In the following, I first define the indices and discuss their properties. Then I look at the specialisation indices for the pollinator in an artificial network to illustrate how the indices reflect the actual degree of specialisation, also by contrasting observed visitations with expectations from a null model. Next, I analyse 21 pollination networks with quantitative information to explore correlations between the different indices. Thereafter I explore one network in more detail ("Safariland" recorded by Vázquez and Simberloff, 2003), which has been used previously for illustration purposes because it depicts clear examples of generalists and specialists. Finally, I examine how to statistically evaluate how differences between two land uses affect

pollinator specialisation, before reflecting on the usefulness of specialisation indices in pollination networks in the context of pollinator morphology and behaviour. Data and R-code for all analyses are provided in the supplementary material following the same sequence.

## 2 Specialisation Indices

Here I only present specialisation indices based on networks (Table 1), and none of those directly used on field data (see, for example, Fenster et al., 2004; Ollerton et al., 2007).

**Table 1** Overview of specialisation measures.  $N_I$  and  $N_J$  refer to the number of plant and pollinator species in the network, respectively.

Index	Network type	Min	Max	Value for specialist	Comments
Degree	Binary	1	$N_I$	low	Shannon's H and strength can be interpreted as quantitative versions of degree
Normalised degree	binary	0	1	low	Computed as degree/ $N_I$
Node specialisation index NSI	binary	1	$N_J/2$	high	Co-determined by specialisation of other pollinators in the network; based on path lengths
Betweenness centrality BC	binary	0	$N_J(N_J-1)/2^*$	low	Based on path lengths; similar to NSI, but more common
Closeness centrality CC	binary	0	$(N_J-1)/2^*$	low	Based on path lengths; see Butts (2009) for implementation details
Strength	weighted	0	$N_I^{**}$	unclear	Co-determined by specialisation of other pollinators in the network; computed as dependence-weighted degree
Pollination Support Index PSI	weighted	0	1	(high)	Specialists have a high PSI, but only when the plant is also specialised on them; common generalists can also have relatively high PSI values; extension of the idea of strength
Shannon's diversity H	weighted	0	$\ln N_I$	low	
Effective number of partners $d'$	weighted	0	$N_I$	low	Shannon's H converted into the degree scale
	weighted	0	1	high	Measures specialisation as discrimination from expectation based on how many interactions a plant has ***

\* Based on the non-normalised definition.

\*\* The maximum value for *strength* of pollinator  $i$  is the *degree* of  $i$ . Since the maximum *degree* is  $N_I$ , theoretically this is also the maximum for strength. For example in a "network" with only one pollinator and 10 plants, *degree* would be 9, as would be *strength*.

\*\*\* Number of interactions is taken as surrogate for their abundance or attractiveness to all pollinators. If true abundances are known, they can be used instead.

### 2.1 Degree (qualitative measure)

A pollinator's degree is simply the number of observed plant-links of that species: a higher degree value indicates a higher level of generalism. The intensity of interactions is irrelevant, i.e. degree is calculated based on a binary interaction matrix. Thus, degree describes specialisation in a qualitative way similar to describing diversity as number of species (Blüthgen et al., 2006). In the literature, it is the distribution of degrees within a network that has found much interest, more so than the degree of a pollinator itself (e.g. Bascompte et al., 2006,

Burns, 2007, Dunne et al., 2002, Jordano, 1987, Jordano et al., 2003). Normalised degree (Martín González et al., 2010) re-scales degree by dividing it by the number of plant species on offer, thus ranging it between 0 and 1.

## 2.2 Node specialisation index (NSI, qualitative measure)

*NSI* is calculated as the mean path length, also known as geodesic distance  $d_{ij}$ , in a one-mode network between each pollinator and every other pollinator (based on an idea presented by Dalsgaard et al., 2008):

$$NSI_i = \sum_{i < j}^n \frac{d_{ij}}{n(n-1)}. \text{ When two pollinators visit the same plant species, the path length between them is 1.}$$

Two pollinators not visiting the same plant species may be linked through a third pollinator, which has a visitation in common with either of them. Then, the path length between the two original pollinators is 2. Hence, a (minimum) *NSI* of 1 indicates that a species is linked to all other pollinators directly, while a *NSI* of 3 indicates that it is, on average, three links away from all other pollinators. The logic behind this approach is that a pollinator with a low *NSI* does not contribute much to the pollination of plant species, because it only pollinates plant species already visited by other pollinators. The appeal of the *NSI* is that it is directly based on network topology, but as such it also has some intrinsic problems: (1). When a network consists of two or more compartments, the path length to pollinators in a different compartment is infinity. Thus, the mean of all paths will also be infinity. This is a common and unresolved problem in network analysis. A common, but “non-canonical” solution (geodist help-page: Butts, 2007) is to give infinite path lengths the value of the longest observed path plus one. Another is to define such paths as “not available data” and hence omit them from calculations. *NSI*-values differ greatly between these two variations on how to handle compartments. (2). The *NSI* is not ranged between a minimum and a maximum, for example between 0 and 1. The lowest possible *NSI* is one, with all pollinator species being connected to all others (the self-loop of length 0 is omitted). The maximum, however, depends on the method of handling compartments. If no compartments are presents, the maximum value is  $N_j/2$  ( $N_j$  = number of pollinators). Thus, although *NSI* can be used to rank species by their degree of “node specialisation”, it does not allow for the quantification of an absolute degree of specialisation. In this respect the *NSI* is similar to degree and strength, but inferior to  $d'$  and *PSI* (see below). (3). *NSI* is defined through other pollinator species, and hence a measure of network position rather than species characteristics. If, for example, a pollinator was lost from the network, all *NSI* values would change because one potentially linking species is lost (independent of the way the pollinators themselves may respond to the changing community structure).

## 2.3 Betweenness centrality and closeness centrality (qualitative measures)

Both measures are similar to *NSI*, based on a qualitative one-mode representation (e.g. Borgatti and Everett, 2006). They are the two most common centrality measures employed in social network analysis to describe how pivotal a node is for the network (Freeman, 1979) and have been proposed as measure of generalisation in pollination networks by Martín González et al. (2010). Betweenness is the fraction of all shortest paths that pass through that node. Let  $n$  represent the number of nodes in a network,  $g_{ij}$  the number of shortest paths between them, while  $g_{ij}(k)$  is the number of shortest paths between  $i$  and  $j$  that go through  $k$ . Then the (normalised) betweenness centrality  $BC$  of  $k$  is defined as:

$$BC_k = 2 \sum_{i < j; k \neq i} \frac{g_{ij}(k) / g_{ij}}{(n-1)(n-2)}.$$

Closeness centrality  $CC$  of  $k$  is the inverse of the average distance  $d_{jk}$  to all nodes:

$$CC_k = \left( \sum_{j=1, j \neq k}^n \frac{d_{jk}}{n-1} \right)^{-1}.$$

Confusingly, two definitions of closeness exist, the other one being the mean distance, rather than the inverse of it. This alternative definition is thus extremely similar to the *NSI*. Low betweenness or closeness scores indicate specialisation. Their limitations are similar to that of *NSI* (see above), e.g. with respect to unconnected graphs and the fact that they are qualitative indices. Notice that *BC* and *CC* are often normalised by dividing each by the grand sum, thus making them sum to 1.

#### 2.4 Strength (quantitative measure)

This index aims at quantifying the dependence of the plant community on a given pollinator (Jordano, 1987). As an intermediate step, a matrix of dependencies is calculated by dividing the observed number of interactions by the total number of interactions for each plant. The derived values represent the dependence of each plant on each pollinator as the proportion of visits the plant receives from each pollinator. A pollinator's *strength* is simply the sum of dependencies for that pollinator. Formally, if  $a_{ij}$  is the number of visit pollinator  $j$

pays to plant  $i$ , then the dependence for this combination is given by  $p_{ij} = \frac{a_{ij}}{\sum_i a_{ij}}$ . The *strength* of species  $j$  is

then:  $s_j = \sum_i p_{ij}$ . High *strength* indicates a high relevance of this pollinator for the plants in the system,

which may, but need not, be a sign of specialisation.

#### 2.5 Pollination service index (*PSI*, quantitative measure)

With the pollination service index we follow the idea of *strength* one step further. A pollinator is more important for a plant species when it is a) common and b) specialised. A rare pollinator will also only rarely pollinate a flower, and a generalist may deliver a large proportion of non-target plant pollen (depending on the way pollen is deposited on the pollinator's body). *PSI* seeks to embrace both objectives by calculating the proportion of *conspecific* pollen delivered to the target plant. As such it is the product of dependencies of the pollinator (representing their specialisation) and dependencies of the plant (representing the importance of each plant species for each pollinator). For each pollinator, these values are summed. *PSI* has one main weakness (apart from the lack of evidence of its usefulness): A pollination event requires *two* visits (at least in non-autogamous plant species), one to pick up the pollen and one to deposit it, while the index assumes only one visit. This can easily be rectified by taking the dependence matrix of the plants to the power of two. This seems, however, too conservative, since pollen may hang on for several visits, thus reducing the exponent to an unknown value between 1 and 2. Formally, *PSI* is an extension of *strength*. Similar to  $p_{ij}$  we can define  $p_{ji}$

as the dependence of any pollinator  $j$  on visits to each plant species  $i$ :  $p_{ji} = \frac{a_{ij}}{\sum_j a_{ij}}$ . This represents the

proportion of visits a pollinator makes to every plant, and hence is a measure of how diluted the pollen it carries is. *PSI* is then defined as:  $PSI_j = \sum_i (p_{ij} \cdot p_{ji}^\beta)$ . The first factor describes how much a plant relies on a

pollinator and the second, how much the pollinator relies on the plant. Here, an exponent  $\beta$  is introduced, which adjusts how many visits a pollinator has to make to a plant in order to pollinate it. Since this type of data is usually not available, we set  $\beta=1$ .

## 2.6 Partner diversity (Shannon's H) (quantitative measure)

Shannon's diversity index ( $H = -\sum_{i=1}^n p_i \ln p_i$ , where  $p_i$  is the proportion of visits the focal pollinator pays to

species  $i$ ) can be used as a measure of specialisation, with high values indicating many plants being pollinated relatively evenly. By raising this value to the power of  $e$  we can compute the effective number of partners in units directly comparable to the number of plant species (Jost, 2006).

## 2.7 $d'$ -Index (quantitative measure)

The rationale for the  $d'$ -index was given by Blüthgen et al. (2006): "Hurlbert (1978) emphasized that not only proportional utilization, but also the proportional availability of each niche should be taken into account. A species that uses all niches in the same proportion as their availability in the environment should be considered more opportunistic than a species that uses rare resources disproportionately more." The  $d'$ -index calculates a Shannon entropy-like diversity index of each pollinator's visitation preferences, and then uses a heuristic search for the highest possible specialisation to determine the minimum specialisation possible under the constraints of observed plant and pollinator abundances. This, and the analytical solution for the minimum specialisation, are then used to re-scale the index to a range between 0 (perfect opportunist) and 1 (disproportionate specialist). While  $d'$  performed favourably on various data sets (Blüthgen et al., 2006), it is very sensitive for rare species, which either happen to visit a common plant (yielding a  $d'$ -value of 0) or a rare plant (yielding a  $d'$ -value near 1).

## 3 Computation

All indices can be calculated using the function "species level" in the R-package bipartite (Dormann et al., 2009, Dormann et al., 2008). There are two more indices returned which do not measure specialisation but which also quantify the relationship between pollinators and plants: Fisher's  $\alpha$  (a measure of partner diversity, representing the parameter of Fisher's logarithmic series fitted to the interactions of each species, see Fisher et al., 1943) and interaction balance (quantifying the asymmetry of interactions, i.e. if a pollinator is more specialised on the plants than the plants are, on average, specialised to this pollinator, see Vázquez et al., 2007). R-code for all analyses performed are available as supplementary material; data sets used below ship with the R-package itself and were taken from the NCEAS interaction web database maintained by Diego Vázquez (<http://www.nceas.ucsb.edu/interactionweb>).

What is clear from the index descriptions is that they describe, intentionally or unintentionally, different characteristics of the plant-pollinator system. Qualitative descriptors give little information about the ecological and evolutionary processes, since they do not describe how common interactions between species are. Quantitative descriptors on the other hand do quantify ecologically directly interpretable characteristics. The index  $d'$ , for example, specifies a pollinator's discriminatory behaviour: does it choose what is on offer or are there preferences beyond what it would encounter during random searches? Co-evolution between flowers and pollinators can only occur when the pollinator displays preferences, which the plant can then in turn try to amplify (through chemical attractors, flower structure or colour, and through "tailoring" reward types such as nectar or resin: e.g. Vega-Redondo, 1996, Fenster et al., 2006). The pollinator service index tries explicitly to quantify the benefit of a pollinator from the plant's perspective. For them, a sufficiently high provisioning of the right type (i.e. conspecific) of pollen is of crucial importance. Thus, many pollinators not necessarily deposit enough of the right pollen (although pollination success often goes hand in hand with pollinator diversity: Kremen et al., 2002).

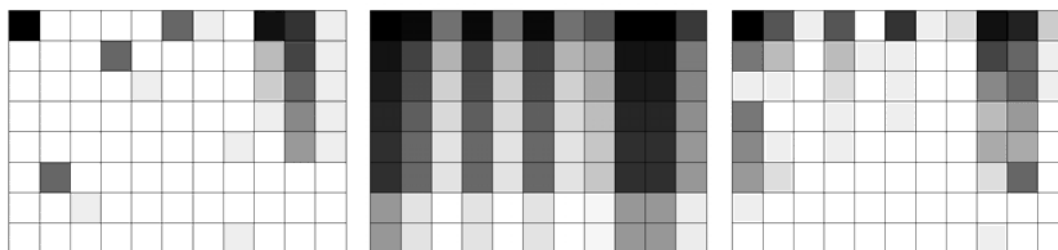
#### 4 An Artificial Example

In order to understand the different specialisation indices better, one can use an artificial network (Table 2), describing the visits of eleven pollinators to six plant species (a to f). They are arranged in a sequence from high to low specialisation. Within species with the same number of plant species visited (i.e. *degree*), the sequence is from those visiting rare plants to those visiting common plants. Pollinator S1c is clearly specialised, as are pollinators S4i and S3i. These latter two, however, have been observed only once, and the information on their preferences is hence uncertain: a single new visit can turn them from a specialist into a moderate generalist (compare S2r and S5r). Pollinators S2i and S4i are specialists, visiting only one plant species. Among those not visited by S2i is the plant species with the overall highest visitation rate (a). Thus, in the logic of the  $d'$ -index, we can consider S2i *more* specialised than S4i, because it avoids visiting a. Abundance effects can be investigated by species S7c and r. S7r has been observed fewer times than S7c, making its classification less certain.

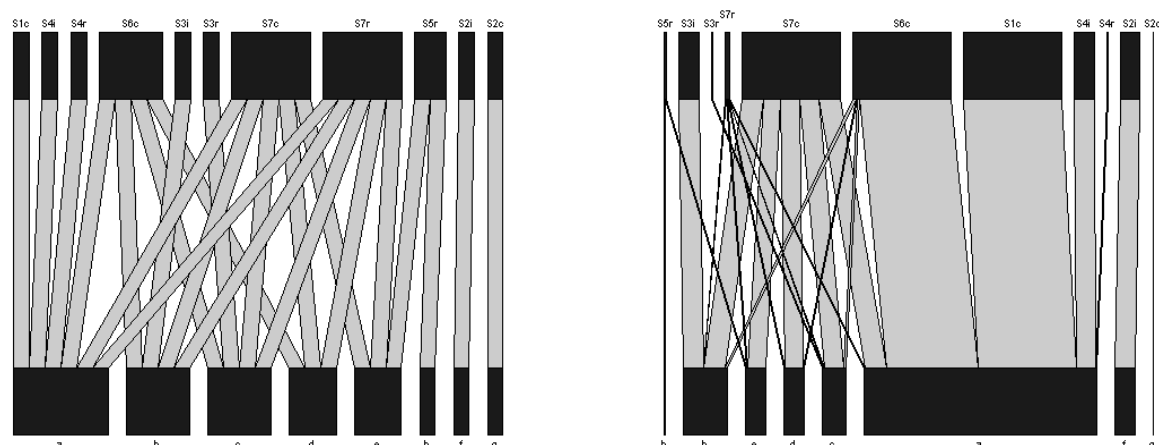
**Table 2** Example pollination network. Columns represent pollinators, rows (lower case letters) the plant species they visit. Numbers refer to observed visitations. For pollinators, species are labelled by their specialisation (S1 to S7 from highly specialised to highly generalised) and by their abundance (common, intermediate, rare). The species S1c, S6c and S7c are common pollinators, with S1c being a common specialist, S6c more generalist but still clearly specialised, and with S7c a common generalist. Species S2 through 4 are also highly specialised, but along a gradient of plant attractiveness (S2 on a specialised plant, S3 on a moderately attractive plant and S4 on a highly attractive plant). S5r is similar to S2r (i.e. rare and interacting with a specialised plant), but with an additional observation to investigate the effect of sampling. S7c and S7r are both generalists, but differently common. In real pollination networks, this set-up is not uncommon: some plants are visited very often (a), some intermediate (b to f), and some hardly ever (g, h). Also pollinators are often log-normally distributed in their abundance.

	S1c	S2i	S2r	S3i	S3r	S4i	S4r	S5r	S6c	S7c	S7r
a	100	0	0	0	0	20	1	0	94	22	1
b	0	0	0	20	0	0	0	0	3	21	1
c	0	0	0	0	1	0	0	0	2	20	1
d	0	0	0	0	0	0	0	0	1	19	1
e	0	0	0	0	0	0	0	1	0	18	1
f	0	20	0	0	0	0	0	0	0	0	0
g	0	0	1	0	0	0	0	0	0	0	0
h	0	0	0	0	0	0	0	1	0	0	0

Table 2 gives the values for all 11 indices discussed above, along with the ranking they imply. Two patterns are apparent: Firstly, some indices (*degree*, *BC*, *H*) do not differentiate in specialisation between species with only one link (S1 to S4). Only strength, *PSI* and  $d'$  are able to rank the majority of species unambiguously (bold printed values in Table 2 indicate non-ties). While this is to some extent an artefact of the evenly balanced number of interactions in this network, it is typical for networks with low sampling intensity and, of course, binary networks.



**Fig. 1** A matrix representation of the example network (left), the probability matrix based on marginal totals (centre) and a random matrix produced by the Patefield algorithm (right) (Patefield, 1981). Note that in either case column and row sums are the same as in the original (shading not comparable between panels). Dark shades represent many interactions (high probability in the case of the probability matrix).



**Fig. 2** A binary (left) and weighted (right) depiction of the example network. Pollinators are given on top, plant species at the bottom. Species are sorted to minimum overlap of lines, leading to a centralisation of common species in each trophic level. Notice particularly the shift in pollination relevance of species S1c (gain) and S7r (loss).

Second, the different indices do not identify the same species as most specialised. The setup of the network was in line with degree only. S4r was identified as comparatively generalised by three of the five weighted indices, while S7c was “promoted” to specialisation status similar to single-degree species. These results and Fig. 1 suggests that species S2i and S2r (both with only one link) are most extraordinary because their flowers are not visited by any other pollinator. Indeed, these two species receive the highest possible score by *PSI* and  $d'$  and the lowest by closeness, thus indicating a high specialisation (Table 2). The most generalist species (S7c and S7r) have ties for all but three indices: strength, *PSI* and  $d'$ , again indicating limited sensitivity of most indices to quantitative differences between them.

Species S2r and S5r were set up to illustrate the problem of singleton observations. S5r has a unique interaction (as has S2r), but with an additional non-unique one. As a consequence, it drops dramatically in virtually every index, indicating that all indices are liable to sampling intensity artefacts. The other construction in the example network is that of S2i, S3i and S4i compared to S2r, S3r and S4r. They are similarly specialised, but the first set are intermediately common species, the second rare. The only indices that actually report a difference between S2/3/4i and S2/3/4r are strength, *PSI* and  $d'$  (strength and *PSI* yielding identical values in this case).

## 5 Correlations Between Indices

It is evident from the last section that the eight different specialisation indices partly quantify the same type of specialisation. To find out how much redundancy is present in the set of eight indices, I calculated for all pollinators the key eight specialisation indices and quantified their correlation in each of 21 pollination networks. Table 5 displays the correlation between indices and Fig. 3 ordines their absolute values by similarity. *NSI* and closeness (*CC*) are grouped together, as are degree, betweenness (*BC*) and partner diversity (*H*). This graph suggests that two different properties of pollinators are measured, but these are not the two levels of specialisation alluded to in the introduction, i.e. number of plants visited (represented by degree, *H*, strength) and network position (*NSI*, *BC*, *CC*). Also the distinction between binary and weighted indices is not perfect, although likely to be responsible for the first axis.



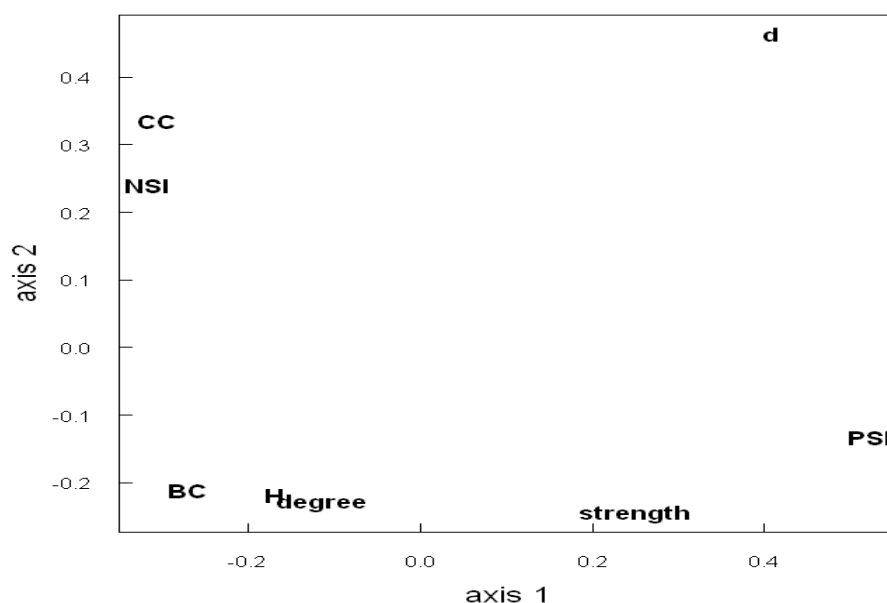
**Table 3** Indices of specialisation for the “pollinators” of the example matrix (Table 1). *Degree*, normalised degree, *NSI*, *BC* and *CC* are binary indices, while *strength*, *PSI*, partner diversity *H*, effective number of partners and *d'* make use of the weighted information provided. Values printed in bold are unique within indices (all others are ties). “—” indicates that no value could be calculated (because these species form their own compartment). Superscripts indicate ranks with 1 for highest specialisation down to 11 for lowest.

	Binary indices					Weighted indices				
	Degree	Normalised degree	Node specialisation index (NSI)	Betweenness centrality (BC)	Closeness centrality (CC)	strength	Pollination Service Index (PSI)	Partner diversity (H)	Effective number of partners	<i>d'</i>
S1c <sup>4.9</sup>	1 <sup>4</sup>	0.125 <sup>4</sup>	1.38 <sup>4</sup>	0.000 <sup>4.5</sup>	0.110 <sup>6</sup>	<b>0.42</b> <sup>7</sup>	<b>0.42</b> <sup>6</sup>	0.00 <sup>4</sup>	1 <sup>4</sup>	0.33 <sup>6</sup>
S2i <sup>3.1</sup>	1 <sup>4</sup>	0.125 <sup>4</sup>	—	0.000 <sup>4.5</sup>	0.000 <sup>1.5</sup>	1.00 <sup>3.5</sup>	1.00 <sup>1.5</sup>	0.00 <sup>4</sup>	1 <sup>4</sup>	1.00 <sup>1.5</sup>
S2r <sup>3.1</sup>	1 <sup>4</sup>	0.125 <sup>4</sup>	—	0.000 <sup>4.5</sup>	0.000 <sup>1.5</sup>	1.00 <sup>3.5</sup>	1.00 <sup>1.5</sup>	0.00 <sup>4</sup>	1 <sup>4</sup>	1.00 <sup>1.5</sup>
S3i <sup>3.8</sup>	1 <sup>4</sup>	0.125 <sup>4</sup>	1.63 <sup>2</sup>	0.000 <sup>4.5</sup>	0.093 <sup>3.5</sup>	<b>0.44</b> <sup>6</sup>	<b>0.44</b> <sup>5</sup>	0.00 <sup>4</sup>	1 <sup>4</sup>	<b>0.72</b> <sup>2</sup>
S3r <sup>3.9</sup>	1 <sup>4</sup>	0.125 <sup>4</sup>	1.63 <sup>2</sup>	0.000 <sup>4.5</sup>	0.093 <sup>3.5</sup>	<b>0.04</b> <sup>10</sup>	<b>0.04</b> <sup>9</sup>	0.00 <sup>4</sup>	1 <sup>4</sup>	<b>0.42</b> <sup>4.5</sup>
S4i <sup>5.5</sup>	1 <sup>4</sup>	0.125 <sup>4</sup>	1.38 <sup>4</sup>	0.000 <sup>4.5</sup>	0.110 <sup>6</sup>	<b>0.08</b> <sup>9</sup>	<b>0.08</b> <sup>8</sup>	0.00 <sup>4</sup>	1 <sup>4</sup>	<b>0.15</b> <sup>8</sup>
S4r <sup>6.2</sup>	1 <sup>4</sup>	0.125 <sup>4</sup>	1.38 <sup>4</sup>	0.000 <sup>4.5</sup>	0.110 <sup>6</sup>	<b>0.01</b> <sup>11</sup>	<b>0.01</b> <sup>11</sup>	0.00 <sup>4</sup>	1 <sup>4</sup>	<b>0.00</b> <sup>10</sup>
S5i <sup>4.9</sup>	2 <sup>5</sup>	<b>0.250</b> <sup>5</sup>	1.63 <sup>2</sup>	0.000 <sup>4.5</sup>	<b>0.084</b> <sup>7</sup>	<b>1.05</b> <sup>2</sup>	<b>0.53</b> <sup>4</sup>	<b>0.69</b> <sup>6</sup>	2 <sup>6</sup>	0.69 <sup>3</sup>
S6c <sup>5.9</sup>	4 <sup>6</sup>	0.500 <sup>6</sup>	<b>1.12</b> <sup>5</sup>	0.179 <sup>9</sup>	<b>0.127</b> <sup>8</sup>	<b>0.59</b> <sup>5</sup>	<b>0.38</b> <sup>7</sup>	0.29 <sup>5</sup>	1.3 <sup>5</sup>	<b>0.21</b> <sup>7</sup>
S7c <sup>5.5</sup>	5 <sup>7.5</sup>	0.625 <sup>7.5</sup>	1.00 <sup>6.5</sup>	0.410 <sup>10.5</sup>	0.136 <sup>9.5</sup>	<b>3.20</b> <sup>1</sup>	<b>0.62</b> <sup>3</sup>	1.61 <sup>7.5</sup>	4.99 <sup>7</sup>	<b>0.42</b> <sup>4.5</sup>
S7i <sup>8.1</sup>	5 <sup>7.5</sup>	0.625 <sup>7.5</sup>	1.00 <sup>6.5</sup>	0.410 <sup>10.5</sup>	0.136 <sup>9.5</sup>	<b>0.17</b> <sup>8</sup>	<b>0.03</b> <sup>10</sup>	1.61 <sup>7.5</sup>	5 <sup>8</sup>	<b>0.08</b> <sup>9</sup>

**Table 4** Z-scores for the eight specialisation indices and the example network, based on 1000 replicate null model runs. Z-scores significantly different from null model values are printed in bold. (The critical value of  $|z| < 2$  is only valid for normally distributed data, hence we counted the number of instances a null model had values larger or equal to the observed.)

Pollinator	Degree	NSI*	BC*	CC	Strength*	PSI	H	<i>d'</i>
S1c	<b>-8.57</b>	<b>0.43</b>	<b>-5.85</b>	<b>5.33</b>	<b>-2.73</b>	<b>20.58</b>	<b>-12.57</b>	<b>35.39</b>
S2i	<b>-4.19</b>	NA	-1.97	<b>-33.34</b>	1.70	<b>54.53</b>	<b>-4.49</b>	<b>35.76</b>
S2r	0.00	NA	0.00	<b>-8.59</b>	<b>9.78</b>	<b>9.78</b>	0.00	<b>4.01</b>
S3i	<b>-3.99</b>	<b>11.25</b>	-1.93	-0.13	0.08	<b>23.69</b>	<b>-4.25</b>	<b>25.70</b>
S3r	0.00	<b>2.53</b>	0.00	<b>0.99</b>	0.34	0.34	0.00	1.44
S4i	<b>-4.02</b>	<b>7.23</b>	-2.06	<b>5.88</b>	<b>-1.06</b>	0.79	<b>-4.36</b>	<b>3.81</b>
S4r	0.00	1.06	0.00	<b>2.52</b>	-0.23	-0.23	0.00	-0.69
S5r	0.82	<b>5.48</b>	0.39	-1.03	<b>9.94</b>	<b>9.91</b>	0.82	<b>5.15</b>
S6c	<b>-4.03</b>	<b>39.50</b>	-0.06	<b>11.21</b>	<b>-2.49</b>	<b>13.46</b>	<b>-9.81</b>	<b>19.48</b>
S7c	<b>-2.37</b>	-0.04		<b>14.16</b>	1.64	<b>49.19</b>	<b>4.27</b>	<b>41.08</b>
<b>S7r</b>	<b>2.91</b>	-0.96	<b>5.27</b>	<b>15.87</b>	<b>0.33</b>	0.02	<b>2.29</b>	0.55

\* Z-scores cannot be meaningfully used to derive statistical significances due to substantial deviation from normal distribution of values (see Fig. 4).



**Fig. 3** Non-parametric multidimensional scaling (nMDS) of the Pearson correlation matrix of the eight indices based on the analysis of 21 pollination networks. Axis 1 (largely dominated by the separation of binary and weighted indices) explains 54% of the variation, axis 2 (no interpretation) an additional 40%. Absolute correlation values were used since indices may be highly but negatively correlated.

**Table 5** Median correlation between indices across the 21 pollination networks. Upper triangle gives Pearson's  $r$  (and interquartile range), lower triangle gives Kendall's  $\tau$ . Bold printed values are consistently significant (19 or more of the 21 networks with significant correlations).

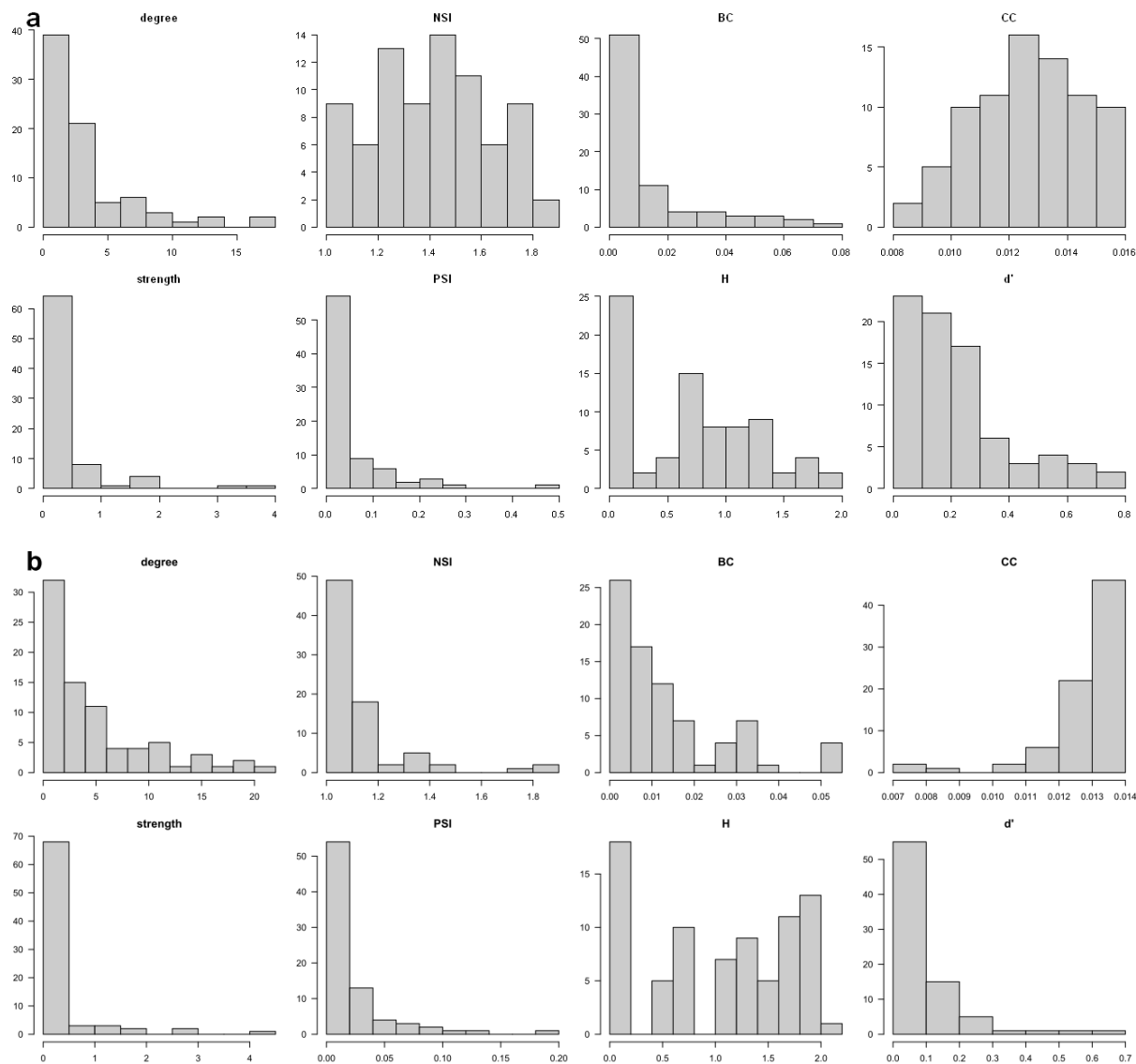
	Species degree	Strength	Pollination Service Index PSI	Node specialisation index	Betweenness	Closeness	Partner diversity	d'
Degree	1	<b>0.780</b> (0.123)	<b>0.500</b> (0.234)	<b>-0.523</b> (0.283)	<b>0.843</b> (0.114)	0.580 (0.309)	<b>0.903</b> (0.096)	0.147 (0.274)
Strength	<b>0.615</b> (0.073)	1	<b>0.816</b> (0.227)	-0.260 (0.384)	<b>0.505</b> (0.326)	0.214 (0.321)	<b>0.641</b> (0.179)	0.385 (0.318)
PSI	0.480 (0.200)	<b>0.872</b> (0.115)	1	-0.094 (0.398)	0.257 (0.161)	-0.049 (0.396)	0.406 (0.301)	<b>0.699</b> (0.200)
NSI	<b>-0.536</b> (0.176)	-0.168 (0.235)	-0.043 (0.247)	1	<b>-0.531</b> (0.252)	<b>-0.998</b> (0.453)	<b>-0.524</b> (0.296)	0.310 (0.321)
BC	<b>0.857</b> (0.198)	<b>0.499</b> (0.176)	0.320 (0.245)	<b>-0.629</b> (0.151)	1	<b>-0.540</b> (0.209)	<b>0.699</b> (0.176)	-0.057 (0.297)
CC	<b>0.583</b> (0.217)	0.142 (0.268)	-0.054 (0.260)	<b>-0.993</b> (0.113)	<b>0.647</b> (0.182)	1	0.551 (0.290)	-0.515 (0.405)
H	<b>0.928</b> (0.057)	<b>0.548</b> (0.093)	<b>0.437</b> (0.168)	<b>-0.498</b> (0.175)	<b>0.794</b> (0.215)	<b>0.555</b> (0.231)	1	0.102 (0.335)
d'	0.140 (0.229)	<b>0.596</b> (0.344)	<b>0.706</b> (0.259)	0.221 (0.187)	0.003 (0.262)	-0.278 (0.317)	0.112 (0.233)	1

Normalised degree is only a rescaled version of degree and hence has an  $r$  (and  $\tau$ ) of 1 with degree. Similarly, effective number of partners is only a non-linearly rescaled version of partner diversity and hence has a  $\tau$  of 1 (but an  $r$  of 0.976).

## 6 A null model for Specialisation

Currently our expectations about which proportion of pollinators in a network should be specialised are very uninformed. Studies such as those by Oleson et al. (2007) or Martín Gonzáles et al. (2010) provide some ideas by classifying pollinators as well-connected (and hence generalists) or peripheral (and hence specialists). Fig.

4a depicts an attempt to quantify the distribution of specialisation according to the eight indices for an intensely sampled but still large pollination network (Memmott, 1999). It shows, again, that results are very different for the indices. Overall, all indices point towards a large proportion of specialists (i.e. many low values: see Table 1 for which values indicate specialisation), while the distribution of  $d'$  (and  $PSI$ ) -values indicates a very small proportion.



**Fig. 4** Distribution of index values across the 79 pollinators in the network of a) Memmott (1999) and b) a null model. Notice that most of the eight indices exhibit non-normal distributions, making the application of  $z$ -scores inappropriate.

Such histograms cannot really identify specialists by themselves, because we do not know what a generalised pollination network with the same number of observations per species would look like. Furthermore, network dimensions (ratio of number of plants and pollinators) as well as sampling intensity (mean number of interactions per cell) have been shown to greatly affect network indices (Dormann et al., 2009). A null model approach allows us to correct for such possible artefacts.

In a network, observed patterns of specialisation can have three causes: 1. “true” specialisation (Vázquez & Aizen, 2006); 2. competitive displacement (e.g. Aizen et al., 2008); and 3. chance. While 1 and 2 are impossible to disentangle without more data (e.g. from systems where the likely dominant competitor is absent), the null model approach seeks to avoid interpreting chance as specialisation (Blüthgen, 2010). Chance, in turn, can have various causes, and this null model approach focuses on two: a) artefacts due to low sampling intensity and b) intrinsic differences in flower attractiveness or abundance.

For each pollinator, the null model scatters the observed number of interactions over all flowers. Across all pollinators, however, the number of visits to each flower is also kept at the observed value. Thus, column and row totals are kept constant. This null model is referred to as the Patefield algorithm (Patefield, 1981; Blüthgen et al., 2006). Note that this is a null model for quantitative (i.e. weighted) networks, not for qualitative (binary) ones. For binary networks, the approaches proposed by Miklós & Podani (2004) could be used (see also below for evaluation of binary indices, suggesting that binary matrices do not contain sufficient information to derive specialisation).

One criticism of this null model is its conservatism. If flower abundance is driven by pollinator abundance, then the null model deletes the outcome of ecological interactions. How likely this is to be relevant we do not know. While pollen limitation as such has frequently been reported (reviewed in Knight et al., 2005, 2006), there is, as yet, little evidence that the abundance of flowers of a *specific* plant is affected by the number of visits from a *specific* pollinator (but see Kunin, 1993). If, however, another pollinator is able to replace and complement the target species, no abundance consequences are to be expected.

While the evidence for specialisation-driven abundances is accumulating (Aguilar et al., 2006), the problem of over-interpretation is omnipresent without null model corrections (Vázquez and Aizen, 2006). Plant abundances are more likely to be limited by nutrients and water than by pollinators (Harper, 1977; Ghazoul, 2005; Bos et al., 2007). Not subtracting mere sampling effects would lead us to interpret the fact that some plants are locally more abundant than others to be an indication of their specialisation. This whole dilemma can be summarised in one sentence: It is mere speculation to attribute abundance patterns to the structure of ecological networks, when it can be shown that a large proportion of these structures are sampling artefacts.

How can we use null models to investigate whether an index mis-interprets random variation in an unspecialised pollinator as specialisation? An example for the network of Memmott (1999) is given in Fig. 4b. It shows that apart from *NSI*, its highly correlated counterpart *CC* and  $d'$ , all index distributions can be remarkably similar for real and null model networks. Thus, in order to formally test whether a given specialisation index value indicates specialisation or not, I generated 1000 null models for the example network (Figs. 1 and 2, Table 2) using the Patefield algorithm. These null models have the same number of observations per plant and pollinator species, but the interactions are spread randomly (obeying the marginal total constraints: Fig. 1). As a consequence, null model pollinators are extreme generalists. *Z*-scores (i.e. observed value – mean null model value, divided by the standard deviation of the null model values) indicate how clearly an index differentiates between observed values and null model values (see Table 4).

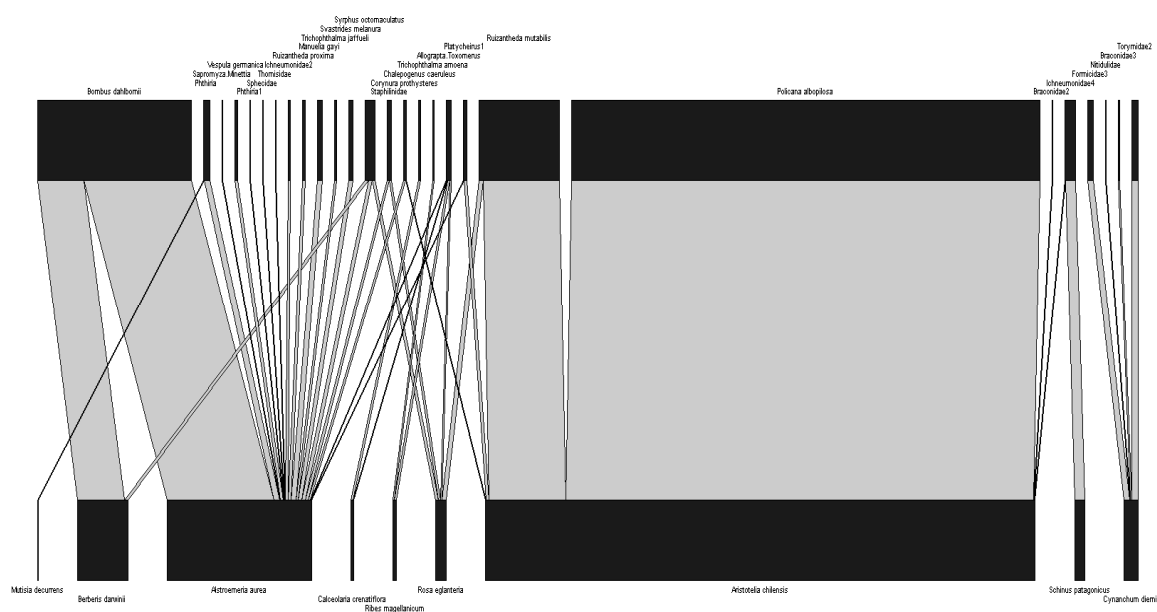
In our example network (Table 1), species *S7c* and *S7r* are supposed to be generalists and were also classified as such by most indices (Table 3). The quantitative indices (notably *PSI* and  $d'$ ) yielded relatively high values. When comparing the observed values to those of null models using *z*-scores (Table 4), only *NSI*, *BC* and strength are indistinguishable from the null model generalist, while all other indicate a significant amount of specialisation. Most indices quantified species *S1c*, *S2i*, *S3i*, *S4i* and *S6c* as specialists. *PSI* and  $d'$  additionally show the intended sequence of discrimination from *S2i*, *S3i* to *S4i* by decreasing *z*-scores (because their recorded visits are less and less likely to occur by chance). Surprisingly, however, several more indices

picked up the increasing specialisation of the singletons S2r, S3r and S4r:  $CC$ , strength,  $PSI$  and  $d'$ . Here  $degree$  and  $H$  displayed their inability to extract meaningful information from singletons.

Finally, the substantial change in  $z$ -scores of all binary indices (and  $H$ ) from species S2r to S5r is also noteworthy. S5r was supposed to represent the same species as S2r, just with an additional, random interaction. The data do not allow much inference on species S2r, so the  $z$ -scores should not have changed much. Strength,  $PSI$  and  $d'$  were indeed robust to this sampling effect, in stark contrast to the binary indices.

## 7 Specialisation Calculations in a Real Network

To illustrate which values we can expect in real networks, and how these relate to the observed interactions, I calculated the difference specialisation indices for the pollination network “Safariland” (Fig. 5).



**Fig. 5** Bipartite graph of the pollination network Safariland (Vázquez & Simberloff 2003).

Different indices identify different species as specialists (Table 6). Species degrees are low (i.e. 1) for 18 out of 27 pollinator species, thus not allowing the identification of the most specialised pollinator. *Chalepogenus caeruleus* was most specialised according to both  $d'$  and  $NSI$ , and *Ichneumonidae4* according to  $PSI$ . *Bombus dahlbomii* had the highest strength value (but also high values for  $PSI$  and  $d'$ ).

Clearly, the Safariland pollinator network does not offer a single species as the obvious choice for the most specialised. Most of the above named species are plausible candidates. I regard *Bombus dahlbomii* as a particularly good candidate, simply because its commonness leaves little room for statistical artefacts (in contrast to the singleton *Trichophthalma amoena*, for example). This subjective judgement is, however, only supported when using strength as the relevant measure. Species observed during very few flower visitations (e.g. less than 4) may suffer more from “incidental” identification as specialist.

**Table 6** Specialisation index values for the pollinators of the Safariland pollination network (Vázquez & Simberloff 2003; Fig. 3). Values significantly different from null model values are printed in bold. Species are sorted by abundance from common to rare.

Pollinator	Degree	NSI	BC	CC	Strength	PSI	H	d'
<i>Policana albopilosa</i>	1	0.000	0.031	1.87	0.852	0.852	0.00	0.691
<i>Bombus dahlbomii</i>	2	0.000	0.045	1.27	1.671	0.798	0.61	0.858
<i>Ruizantheda mutabilis</i>	2	0.048	0.036	1.64	0.539	0.153	0.20	0.155
<i>Trichophthalma amoena</i>	1	0.000	0.026	2.09	0.400	0.400	0.00	0.847
<i>Syrphus octomaculatus</i>	3	0.023	0.047	1.23	0.360	0.110	1.09	0.386
<i>Manuelia gayi</i>	1	0.000	0.045	1.27	0.034	0.034	0.00	0.320
<i>Allograpta.Toxomerus</i>	4	0.417	0.049	1.14	0.988	0.332	1.28	0.648
<i>Trichophthalma jaffueli</i>	1	0.000	0.045	1.27	0.014	0.014	0.00	0.265
<i>Phthiria</i>	2	0.000	0.045	1.27	1.038	0.145	0.35	0.392
<i>Platycheirus</i> 1	2	0.244	0.050	1.09	0.010	0.005	0.50	0.000
<i>Sapromyza.Minettia</i>	1	0.000	0.045	1.27	0.005	0.005	0.00	0.200
Formicidae3	1	0.000	0.007	1.00	0.400	0.400	0.00	0.812
Nitidulidae	1	0.000	0.007	1.00	0.050	0.050	0.00	0.551
Staphilinidae	2	0.023	0.047	1.23	0.219	0.097	0.68	0.409
Ichneumonidae4	2	0.000	0.031	1.86	1.001	0.938	0.23	0.901
Braconidae3	1	0.000	0.007	1.00	0.100	0.100	0.00	0.617
<i>Chalepogenus caeruleus</i>	1	0.000	0.026	2.09	0.750	0.750	0.00	0.950
<i>Vespula germanica</i>	1	0.000	0.045	1.27	0.019	0.019	0.00	0.283
Torymidae2	1	0.000	0.007	1.00	0.450	0.450	0.00	0.832
Phthiria1	1	0.000	0.045	1.27	0.005	0.005	0.00	0.200
<i>Svastrides melanura</i>	1	0.000	0.045	1.27	0.029	0.029	0.00	0.308
Sphecidae	1	0.000	0.045	1.27	0.005	0.005	0.00	0.200
Thomisidae	1	0.000	0.045	1.27	0.005	0.005	0.00	0.200
<i>Corynura prothysteres</i>	2	0.244	0.050	1.09	0.016	0.011	0.56	0.121
Ichneumonidae2	1	0.000	0.045	1.27	0.019	0.019	0.00	0.283
<i>Ruizantheda proxima</i>	1	0.000	0.045	1.27	0.019	0.019	0.00	0.283
Braconidae2	1	0.000	0.031	1.86	0.001	0.001	0.00	0.000

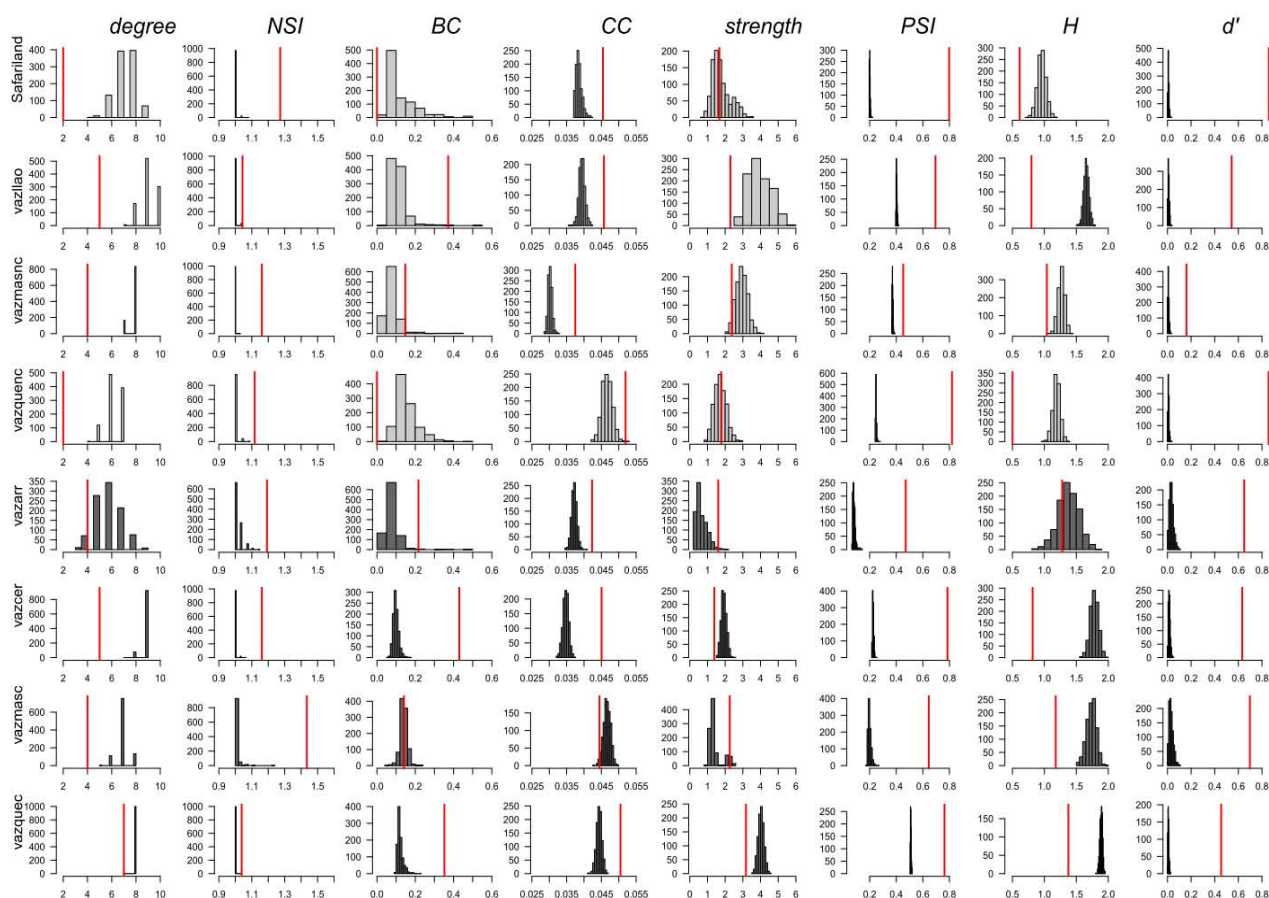
## 8 An Argentinian Case Study: Effects of Cattle Grazing on Specialisation of Pollinators

The indices and null model correction introduced in the previous sections can be used to investigate shifts in pollinator specialisation. For example, Vázquez and Simberloff (2003) report on the effect of cattle grazing on pollination network structure. We can use their data to quantify the specialisation of the common bumble bee *Bombus dahlbomii* across eight networks, four grazed and four ungrazed. This example shall demonstrate the approach one can take to correct the “raw” specialisation index for what a generalist pollinator of the same abundance would yield in the same network (the null model introduced earlier).

The challenge is twofold: first, for each network, the raw specialisation index needs to be corrected relative to the null model generalist; second, using these corrected values, a test statistic must be computed to allow an inferential assessment of the difference between the two treatments.

Fig. 6 shows, for all eight networks and all eight indices, the position of the observed value relative to the null model values. These represent random realisations of a perfect generalist. Hence, when the observed value is within the histogram of null model values, *Bombus dahlbomii* is classified as generalist (e.g. Fig. 6 Safariland and strength). For most plots, the observed value is consistently on one side of the histogram, indicating consistent specialisation, but not so for BC, CC and strength. However, comparing these plots with those for a more obvious generalist, *Vespula germanica* (Fig. 7), we find few clear differences. From these data, we may thus want to generate different types of summary plots to elucidate further on grazing-induced

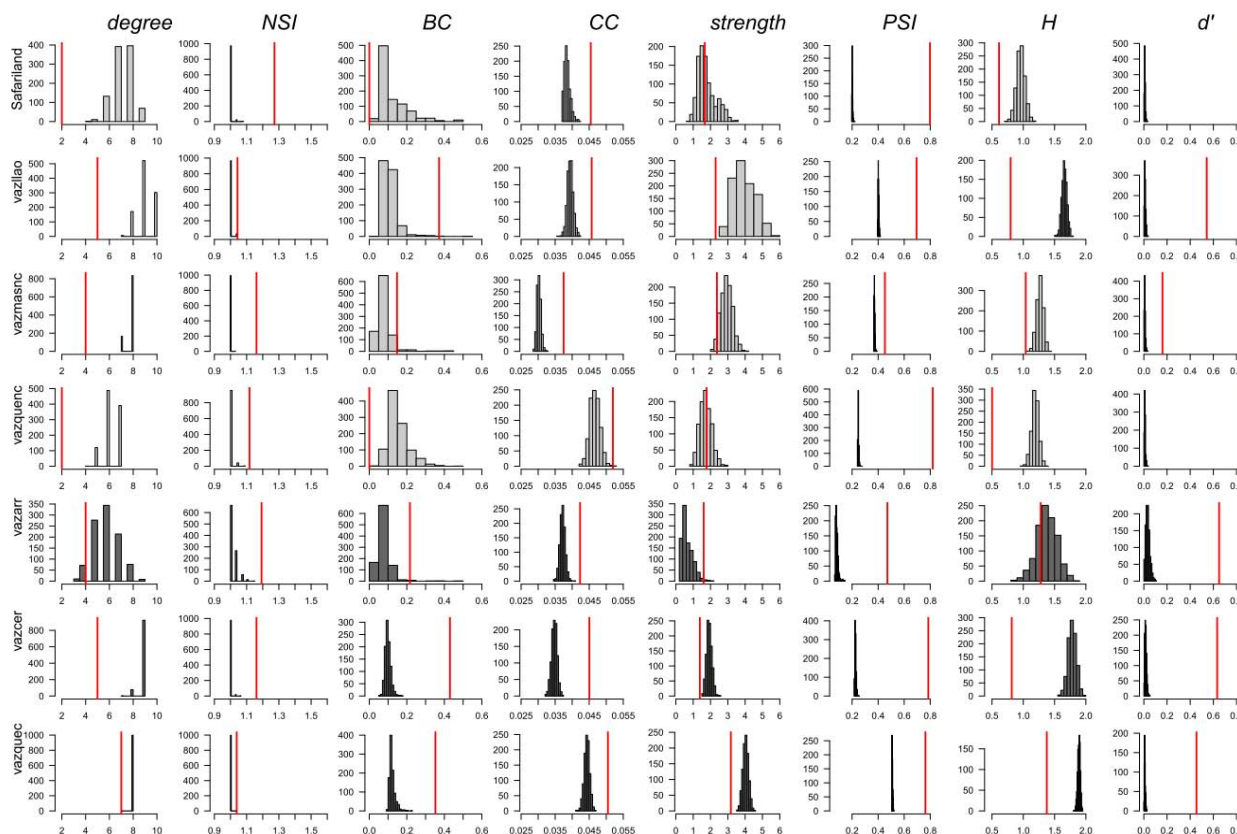
differences (Fig. 8): those of the raw data, those of difference between observed and mean null model values and  $z$ -scores.



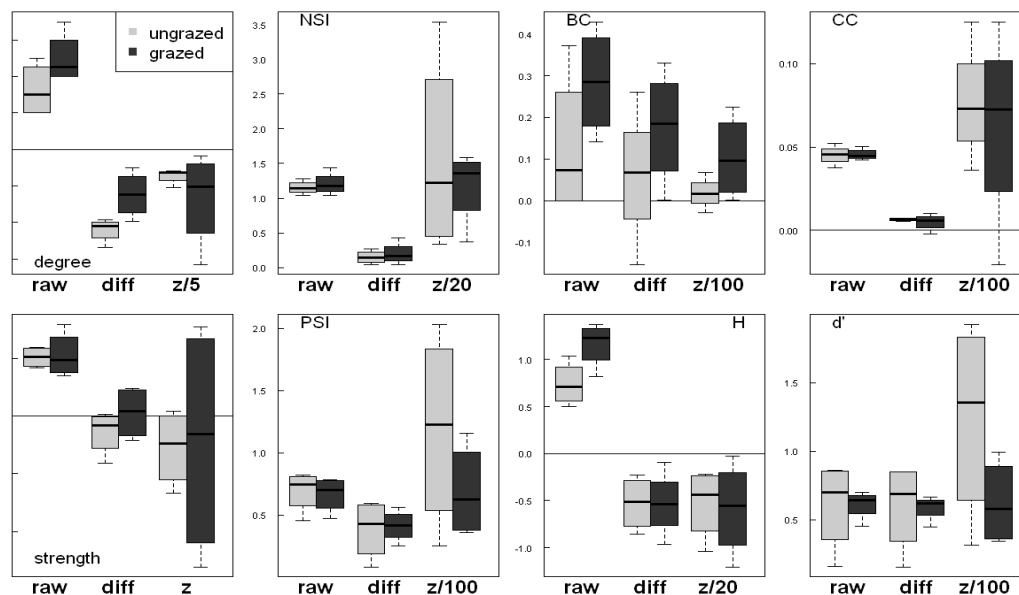
**Fig. 6** Observed and null model specialisation values of *Bombus dahlbomii* for the analysis of specialisation shift. Red lines indicate observed value, histograms distribution of 1000 null models. These represent the position of a perfect generalist. Light grey indicates no grazing, dark grey grazed sites. Names of the data sets (on the left side of the panels) is according to their name in the bipartite package.

These graphs (Fig. 8) can be read, for example for degree, like this: on grazed sites, *B. dahlbomii* seems to have a slightly higher number of links than on ungrazed sites (“raw”). This difference is amplified when correcting for the position of a perfect generalist (“diff”). Now it also becomes apparent that *B. dahlbomii* has fewer links than a generalist (and can hence be judged to be a specialist). However, when additionally taking into account the spread of values for a generalist (see also Fig. 6) the observed degrees cannot be distinguished between grazed and ungrazed sites (although they are still significantly specialised, set  $z/5$ ).

In particular the step from differences to  $z$ -scores usually has very strong effects because it incorporates the spread of null model values. In this example, only the *BC* index yielded a significant difference in  $z$ -scores ( $P = 0.026$ , based on an ANOVA with  $F$ -value distribution drawn from the analyses of the 1000 null models; see supplementary material for details and code). Here, *Bombus dahlbomii* in grazed sites had higher betweenness than those in ungrazed sites, indicating a decrease in specialisation with grazing. (A Bonferroni adjustment for the eight comparisons would render this finding insignificant, however.)



**Fig. 7** Same as Fig. 6, but for *Vespula germanica*. This species is more general, being indistinguishable from a null model generalist in most networks and for most indices. Network “vazmasnc” did not comprise this species and is hence omitted here.



**Fig. 8** Specialisation index values of *Bombus dahlbomii* for grazed (dark grey) and ungrazed (light grey) sites. First two boxes represent the raw index values, as computed from the networks. The second pair represents corrected values, i.e. differences between raw values and the mean of the null models. They position the boxes relative to a perfect generalist (which would have a value of 0). The third pair is the z-scores (divided by a constant for more convenient comparison in the plots).



## 9 Specialisation Indices vs. Pollination Ecology: A Cautionary Remark

This paper mainly deals with technical issues around the calculation of specialisation. Ecologically more important are probably behavioural differences within species. For example, individuals within a species may display very high levels of flower constancy, although the species as such is a generalist (e.g. Cakmak and Wells, 1994, Waser, 1986). This would yield much higher conspecific pollen deposition, making the species act as a specialist for many plant species. None of the above indices is able to capture such behaviour, nor are any of the many other indices proposed for ecological network analysis (e.g. Bersier et al., 2002, Dormann et al., 2009). Note, for example, that *Bombus dahlbomii* is actually reported in the literature as a generalist (Cooley et al., 2008, Abrahamovich et al., 2001), although in our analysis it receives very high specialist scores. Furthermore, virtually all analyses of pollination networks are based on observations of visitation, not of pollination events, and we should not over-interpret results based on visitation networks: Smooth skinned hoverflies may well be common flower visitors, but they are not on a par with bees in terms of pollination efficiency (Schittenhelm et al., 1997, but see Alarcón, 2010). It is largely unknown how much pollen is transferred, how much of it is conspecific pollen, and how much of that is viable.

We should also not forget that pollination networks are not static, but rather highly dynamic (Petanidou et al., 2008). Inferences about a species based on only a single (or few) pollination networks will thus not represent the species' adaptability. Finally, the interaction between pollinators and plants may appear to be strong, but rarely is (Waser, 2006). Apart from some well-documented examples of co-evolution (Lunau, 2004), pollinators shift readily between plants when their abundance changes (e.g. Brown and Mitchell, 2001, Lopezariza-Mikel et al., 2007, Fründ et al., 2010), and most plant species have alternative, if usually less efficient, ways of sustaining a population (self-pollination and clonal growth, see, e.g., Kron et al., 1993). Thus, while specialisation indices may help us to get a better understanding of the role different species play within a network, such analyses alone will not be sufficient to deduce population-level consequences, both for plants and pollinators. To measure specialisation in an ecological meaningful way is not trivial, and the traditional plant-centred approach of visitation webs would profit from being complimented by pollinator-centred approach, such as analysis of pollen carried by the pollinator or provided to the brood.

## 10 Conclusions

Measuring specialisation in pollinators requires careful definition of what defines a specialist. As pollinator visiting only few plant species, degree and  $BC$  qualify as suitable indices. Strength adds the qualitative aspect to this question and is also less sensitive to singletons. For specialism viewed in a community context, as discrimination or minimum similarity to other pollinators,  $CC$  and  $d'$  can be recommended, with the latter making use of quantitative data. When looking for a measure of specialisation that can be compared across networks,  $BC$ ,  $PSI$  and  $d'$  yield consistent trends when compared to null models. The discrimination index  $d'$  is the only one where null models are not required, since it corrects for availability by its definition, making it a particularly suitable candidate for cross-network comparisons. Otherwise null models allow the positioning of the observed relative to a perfect generalist, and hence to statistically assess the significance of a specialisation value. In order to be able to select the "best" specialisation index for the question at hand, a clear definition of the required type of specialisation is indispensable. Ideally, this could be formulated as a model that generates networks with specialists. The index that picks up the intended specialist signal best would also be the best choice.

## Acknowledgements

I'd like to thank Diego Vázquez for making data on pollination networks publicly available (through the NCEAS database: <http://www.nceas.ucsb.edu/interactionweb>), Nico Blüthgen for stimulation and encouragement and Jeroen Everaars, Andrea Holzschuh, Bo Dalsgaard, Bernd Gruber, Diego Vázquez and Stella Watts for constructive comments on a previous version. This work was funded by the Helmholtz Association (VH-NG 247).

## References

- Abrahamovich AH, Telleria MC, Diaz NB. 2001. Bombus species and their associated flora in Argentina. *Bee World*, 82, 76-87
- Aguilar R, Ashworth L, Galetto L, Aizen MA. 2006. Plant reproductive susceptibility to habitat fragmentation: review and synthesis through a meta-analysis. *Ecology Letters*, 9: 968-980
- Aizen MA, Morales CL, Morales JM. 2008. Invasive mutualists erode native pollination webs. *PLoS Biology*, 6: e31
- Alarcón R. 2010. Congruence between visitation and pollen-transport networks in a California plant-pollinator community. *Oikos*, 119: 35-44
- Bascompte J, Jordano P, Olesen JM. 2006. Asymmetric coevolutionary networks facilitate biodiversity maintenance. *Science*, 312: 431-433
- Bersier LF, Banasek-Richter C, Cattin MF. 2002. Quantitative descriptors of food-web matrices. *Ecology*, 83: 2394-2407
- Blüthgen N. 2010. Why network analysis is often disconnected from community ecology: A critique and an ecologist's guide. *Basic and Applied Ecology*, 11: 185-195
- Blüthgen N, Menzel F, Blüthgen N. 2006. Measuring specialization in species interaction networks. *BMC Ecology*, 6: 12
- Blüthgen N, Menzel F, Hovestadt T, et al. 2007. Specialization, constraints and conflicting interests in mutualistic networks. *Current Biology*, 17: 1-6
- Borgatti, SP, Everett MG. 2006. A graph-theoretic perspective on centrality. *Social Networks*, 28: 466-484
- Bos MM, Veddeler D, Bogdanski AK, et al. 2007. Caveats to quantifying ecosystem services: Fruit abortion blurs benefits from crop pollination. *Ecological Applications*, 17: 1841-1849
- Brown BJ, Mitchell RJ. 2001. Competition for pollination: effects of pollen of an invasive plant on seed set of a native congener. *Oecologia*. 129: 43-49
- Burns KC. 2007. Network properties of an epiphyte metacommunity. *Journal of Ecology*, 95: 1142-1151
- Butts CT. 2007. sna: Tools for Social Network Analysis. R package version 1.5. <http://erzuli.ss.uci.edu/R.stuff>
- Cakmak I, Wells H. 1994. Honey bee forager individual constancy: Innate or learned. *BeeScience*, 3: 161-169
- Cooley AM, Carvallo G, Willis JH. 2008. Is floral diversification associated with pollinator divergence? Flower shape, flower colour and pollinator preference in Chilean *Mimulus*. *Annals of Botany*, 101: 641-650
- Dalsgaard B, Martín González AM, Olesen JM, et al. 2008. Pollination networks and functional specialization: a test using Lesser Antillean plant-hummingbird assemblages. *Oikos*, 117: 789-793
- Dormann CF, Fründ J, Blüthgen N, et al. 2009. Indices, graphs and null models: analysing bipartite ecological networks. *The Open Ecology Journal*, 2: 7-24
- Dormann CF, Gruber B, Fründ J. 2008. Introducing the bipartite package: analysing ecological networks. *Rnews*, 8: 8-11
- Dunne JA, Williams RJ, Martinez ND. 2002. Food-web structure and network theory: the role of connectance and size. *Proceedings of the National Academy of Science USA*, 99: 12917-12922

- Fenster CB, Armbruster WS, Wilson P, et al. 2004. Pollination syndromes and floral specialization. *Annual Review of Ecology, Evolution and Systematics* 35: 375-403
- Fenster CB, Cheely G, Dudash MR, et al. 2006. Nectar reward and advertisement in hummingbird-pollinated *Silene virginica* (Caryophyllaceae). *American Journal of Botany*, 93: 1800-1807
- Fisher RA, Corbet AS, Williams CB. 1943. The relation between the number of species and the number of individuals in a random sample of animal population. *Journal of Animal Ecology*, 12: 42-58
- Freeman LC. 1979. Centrality in social networks I: Conceptual clarification. *Social Networks*, 1: 215-239
- Fründ J, Linsenmair KE, Blüthgen N. 2010. Pollinator diversity and specialization in relation to flower diversity. *Oikos*, 119: 1581-1590
- Ghazoul J. 2005. Buzziness as usual? Questioning the global pollination crisis. *Trends in Ecology and Evolution*, 20: 367-373.
- Harper JL. 1977. *Population Biology of Plants*. London: Academic Press, USA
- Hurlber, SH. 1978. Measurement of niche overlap and some relatives. *Ecology*, 59: 67-77
- Jordano P. 1987. Patterns of mutualistic interactions in pollination and seed dispersal—connectance, dependence asymmetries, and coevolution. *American Naturalist*, 129: 657-677
- Jordano P, Bascompte J, Olesen JM. 2003. Invariant properties in coevolutionary networks of plant-animal interactions. *Ecology Letters*, 6: 69-81
- Jost L. 2006. Entropy and diversity. *Oikos*, 113: 363-375
- Kremen C, Williams NM, Thorp RW. 2002. Crop pollination from native bees at risk from agricultural intensification. *Proceedings of the National Academy of Sciences of the United States of America*, 99: 16812-16818
- Knight TM, Steets JA, Vamosi JC, et al. 2005. Pollen limitation of plant reproduction: pattern and process. *Annual Review of Ecology, Evolution and Systematics*, 36: 467-497
- Knight TM, Steets JA, Ashman TL. 2006. A quantitative synthesis of pollen supplementation experiments highlights the contribution of resource reallocation to estimates of pollen limitation. *American Journal of Botany*, 93: 271-277
- Kron P, Stewart SC, Back A. 1993. Self-compatibility, autonomous self-pollination, and insect-mediated pollination in the clonal species *Iris versicolor*. *Canadian Journal of Botany-Revue Canadienne de Botanique*, 71: 1503-1509
- Kunin WE. 1993. Sex and the Single Mustard: Population density and pollination behavior. *Ecology*, 74: 2145-2160
- Lopezaraiza-Mikel ME, Hayes RB, Whalley MR, et al. 2007. The impact of an alien plant on a native plant-pollinator network: an experimental approach. *Ecology Letters*, 10: 539-50
- Lunau K. 2004. Adaptive radiation and coevolution - pollination biology case studies. *Organisms Diversity & Evolution*, 4: 207-224
- Martín Gonzáles AM, Dalsgaard B, Olesen JM. 2010. Centrality measures and the importance of generalist species in pollination networks. *Ecological Complexity*, in press, doi:10.1016/j.ecocom.2009.03.008.
- Memmott J. 1999. The structure of a plant-pollinator food web. *Ecology Letters*, 2: 276-280
- Olesen JM, Bascompte J, Dupont YL, et al. 2007. The modularity of pollination networks. *Proceedings of the National Academy of Science USA*, 104: 19891-19896
- Miklós I, Podani J. 2004. Randomization of presence-absence matrices: comments and new algorithms. *Ecology*, 85: 86-92
- Ollerton J, Killick A, Lamborn E, et al. 2007. Multiple meanings and modes: on the many ways to be a generalist flower. *Taxon*, 56: 717-728

- Orwell G. 1945. *Animal Farm*. Secker and Warburg, London, UK
- Patefield WM. 1981. Algorithm AS159. An efficient method of generating  $r \times c$  tables with given row and column totals. *Applied Statistics*, 30: 91-97
- Petanidou T, Kallimanis AS, Tzanopoulos J, et al. 2008. Long-term observation of a pollination network: fluctuation in species and interactions, relative invariance of network structure and implications for estimates of specialization. *Ecology Letters*, 11: 564-575
- Schittenhelm S, Gladis T, Rao VR. 1997. Efficiency of various insects in germplasm regeneration of carrot, onion and turnip rape accessions. *Plant Breeding*, 116: 369-375
- Vázquez DP, Aizen MA. 2004. Asymmetric specialization: A pervasive feature of plant-pollinator interactions. *Ecology*, 85: 1251-1257
- Vázquez DP, Aizen MA. 2006. Community-wide patterns of specialization in plant-pollinator interactions revealed by null models. *Plant-Pollinator Interactions: From Specialization to Generalization* (eds NM Waser & J Ollerton). University of Chicago Press, Chicago, 200-219
- Vázquez DP, Melián CJ, Williams NM, et al. (2007) Species abundance and asymmetric interaction strength in ecological networks. *Oikos*, 116: 1120-1127
- Vázquez DP, Simberloff D. 2002. Ecological specialization and susceptibility to disturbance: Conjectures and refutations. *American Naturalist*, 159: 606-623
- Vázquez DP, Simberloff D. 2003. Changes in interaction biodiversity induced by an introduced ungulate. *Ecology Letters*, 6: 1077-1083
- Vega-Redondo F. 1996. Pollination and reward: A game-theoretic approach. *Games and Economic Behavior*, 12: 127
- Waser NM. 1986. Flower constancy: definition, cause, and measurement. *American Naturalist*, 127: 593-603
- Wase, NM. 2006. Specialization and generalization in plant-pollinator interactions: a historical perspective. *Plant-Pollinator Interactions: From Specialization to Generalization* (eds N. M. Waser & J. Ollerton). University of Chicago Press, Chicago, 3-18

```

setwd("D:/Data/aktuell/BESS/R_FoodWeb/specialisation")
require(bipartite)
# REQUIRES BIPARTITE VERSION 1.11 (or higher)!! (before the species names are
inconsistent and BC/CC incorrect)

#----- example network -----
exa <- matrix(0, 8,11)
exa[1,] <- c(100, 20, 20, 0,0,0,0,0,0,0)
exa[1:4,11] <- c(94,3,2,1) # as contrast to A
exa[1:5,2] <- c(22,21,20,19,18)
exa[2,4] <- 20
exa[6,5] <- 20
exa[1,6] <- 1
exa[3,7] <- 1
exa[7,8] <- 1
exa[c(5,8),9] <- 1
exa[1:5,10] <- 1
exa <- exa[,c(1,5,8,4,7,3,6,9,11,2,10)]
rownames(exa) <- letters[1:8]
colnames(exa) <- paste("S", c("1c","2i","2r","3i","3r","4i","4r","5r","6c","7c","7r"), sep="")

exa
# plot network:
windows(height=6, width=12)
quartz(height=3, width=6)
par(mfrow=c(1,2), mar=c(0,0,0,0))
plotweb(exa>0)
plotweb(exa)
require(sna)
gplot(exa>0, gmode="twomode", vertex.cex=2) # plot as binary graph!
gplot(as.one.mode(exa, weighted=F, project="higher")) # plot only the pollinator-projection

# calculate indices for this network
indices <- specieslevel(exa, index="ALLBUTD")$"higher trophic level"[-4]
write.table(round(indices[,c(1,2,5,6,7,3,4,9,10,11)], 3), file="exampleweb_indices.txt", sep="\t")
indices.red <- indices[-c(2,8,10)]

#----- run null models for example web -----
# visualise the p-mat of the null model:
pweb <- tcrossprod(rowSums(exa), colSums(exa)/sum(exa))
dimnames(pweb) <- dimnames(exa)
visweb(exa, type="none")
visweb(round(pweb*10000), type="none")

```

```

# generate null models and calculate their index values:
exanulls <- nullmodel(exa)
ex1 <- exanulls[[1]];dimnames(ex1) <- dimnames(pweb)
visweb(ex1, type="none")
# this will take some minutes:
exanull.res <- lapply(exanulls, function(x) specieslevel(x, index="ALLBUTD")$"higher trophic
level"[-c(2,4,9,11)])
save(exanull.res, file="exanulls.res.Rdata")

```

```

# calculate z-scores:
z.mat <- matrix(0, 11, 8)
colnames(z.mat) <- colnames(indices)[-c(2,8,10)]
rownames(z.mat) <- colnames(exa)
for (i in 1:8){
  mean.n <- apply(sapply(exanull.res, function(x) x[,i]), 1, mean, na.rm=T)
  sd.n <- apply(sapply(exanull.res, function(x) x[,i]), 1, sd, na.rm=T)
  z <- (indices.red[,i] - mean.n)/ifelse(sd.n==0, 1, sd.n) # sd is 0 when null model values are
constant (i.e. mean.n==index);then the z-score should be 0
  z.mat[,i] <- z
}
z.mat
write.table(round(z.mat[,c(1,4,5,6,2,3,7,8)], 2), file="exampleweb_z.scores.txt", sep="\t")

```

```

# calculate ROBUST z-scores:
z.rob.mat <- matrix(0, 11, 8)
colnames(z.rob.mat) <- colnames(indices)[-c(2,8,10)]
rownames(z.rob.mat) <- colnames(exa)
for (i in 1:8){
  median.n <- apply(sapply(exanull.res, function(x) x[,i]), 1, median, na.rm=T)
  IQR.n <- apply(sapply(exanull.res, function(x) x[,i]), 1, IQR, na.rm=T)
  z <- (indices.red[,i] - median.n)/ifelse(IQR.n==0, 1, IQR.n) # sd is 0 when null model values are
constant (i.e. median.n==index);then the z-score should be 0
  z.rob.mat[,i] <- z
}
z.rob.mat
write.table(round(z.rob.mat[,c(1,4,5,6,2,3,7,8)], 2), file="exampleweb_z.robust.scores.txt",
sep="\t")

```

```

# calculate p-values:
# note that this is not entirely correct, because NAs will also lead to a p-value of 0!
p.mat <- matrix(0, 11, 8)
colnames(p.mat) <- colnames(indices)[-c(2,8,10)]
rownames(p.mat) <- colnames(exa)
for (i in 1:8){

```

```

null.vals <- sapply(exanull.res, function(x) x[,i])
p2 <- 1:3
for (j in 1:11){ # for each species, calculate the number of times the null model value was
higher than the observed:
  p.lower <- null.vals[j,] < indices.red[j,i]
  p.higher <- null.vals[j,] > indices.red[j,i]
  p.equal <- null.vals[j,] == indices.red[j,i]
  sample.size <- NCOL(null.vals) - max(sum(is.na(null.vals[j,])), sum(is.na(p.lower)),
sum(is.na(p.higher)), sum(is.na(p.equal)))
  #p2[j] <- sum(p.higher)/sample.size
  p2[j] <- (min(sum(p.lower, na.rm=T), sum(p.higher, na.rm=T))+sum(p.equal, na.rm=T)) /
sample.size #turn this into probabilities, acknowledging possible NAs (in NSI!)
}
# p2 <- colSums(p, na.rm=T)/colSums(!is.na(p))
p.mat[,i] <- p2 #ifelse(p2>0.5, 1-p2, p2)*2 #two-tailed test
}
p.mat
round(p.mat[,c(1,4,5,6,2,3,7,8)], 3)

```

```

#----- calculate these indices for all webs -----

```

```

webs <- data(package="bipartite")$results[,3]
data(list=webs)

```

```

res <- lapply(webs, function(x) specieslevel(get(x), index="ALLBUTD")$"higher trophic
level"[-4])

```

```

res

```

```

save(res, file="allwebs_special_indices_obs.Rdata")

```

```

load(file="allwebs_special_indices_obs.Rdata")

```

```

#----- calculate index correlations across webs -----

```

```

index.cor.pearson <- index.cor.spear <- index.cor.ken <- array(NA, dim=c(11, 11, 21),
dimnames=list(colnames(res[[1]]), colnames(res[[1]]), webs))

```

```

for (i in seq_along(res)){

```

```

  r <- res[[i]]

```

```

  index.cor.pearson[,i] <- cor(r, use="complete.obs", method="pearson")

```

```

  index.cor.spear[,i] <- cor(r, use="complete.obs", method="spear")

```

```

  index.cor.ken[,i] <- cor(r, use="complete.obs", method="ken")

```

```

}

```

```

index.cor.ken[,5]

```

```

hist(index.cor.pearson[7,5,]) # not always normally distributed

```

```

# summarise in means and median

```

```

#apply(index.cor.pearson, c(1,2), mean, na.rm=T)
#apply(index.cor.pearson, c(1,2), sd, na.rm=T)
apply(index.cor.pearson, c(1,2), median, na.rm=T)
apply(index.cor.pearson, c(1,2), IQR, na.rm=T)

outmat <- round(apply(index.cor.pearson, c(1,2), median, na.rm=T),3)
outmat[lower.tri(outmat)] <- round(apply(index.cor.ken, c(1,2), median,
na.rm=T),3)[lower.tri(outmat)]
write.table(outmat, file="outmat.txt", sep="\t")
outmatqqr <- round(apply(index.cor.pearson, c(1,2), IQR, na.rm=T),3)
outmatqqr[lower.tri(outmat)] <- round(apply(index.cor.ken, c(1,2), IQR,
na.rm=T),3)[lower.tri(outmat)]
write.table(outmatqqr, file="outmatqqr.txt", sep="\t")

cc <- apply(index.cor.pearson, c(1,2), median, na.rm=T)
cc.red <- cc[-c(2,8, 10),-c(2,8,10)]
par(mar=c(5,5,1,1))
plot(cmd <- cmdscale(1-abs(cc.red), k=2), type="n", xlab="axis 1", ylab="axis 2", tcl=0.5,
cex.lab=1.5, las=1)
text(cmd, c("degree", "strength", "PSI", "NSI", "BC", "CC", "H", "d"), cex=1.25, font=2)
cmdscale(1-abs(cc.red), k=4, eig=T)

.8120/sum(cmdscale(1-abs(cc.red), k=4, eig=T)$eig)
.6073/sum(cmdscale(1-abs(cc.red), k=4, eig=T)$eig)

#----- null models -----
# Safariland
saf <- specieslevel(Safariland, index="ALLBUTD")$"higher trophic level"
write.table(round(saf[,c(1,7,8,6,3,5,10,12)],3), file="Safariland_indices.txt", sep="\t", quote=F)

#----- null models -----
nulls <- list()
for (i in seq_along(webs)){
  nulls[[i]] <- nullmodel(get(webs[i]))
}
save(nulls, file="nulls.Rdata")

## let's look at the specialisation values of an extreme specialist, Chalepogenus.caeruleus in
Safariland (web 1):
null.res.1 <- array(NA, dim=c(8, 27, 1000))
for (i in 1:1000){
  null.res.1[,i] <- t(specieslevel(nulls[[1]][[i]], index="ALLBUTD")$"higher trophic
level"[-c(2,4,9,11)])

```



```

    print(i)
  }
  save(null.res.1, file="null.res.1.Rdata")

  null1.ind <- null.res.1[, 17, ]
  (res[[1]][17, -c(2,8,10)] - apply(null1.ind, 1, mean))/apply(null1.ind, 1, sd)
  #
  # degree strength PSI node.specialisation.index
  # betweenness closeness partner.diversity d
  #Chalepogenus.caeruleus -1.372702 9.918387 41.17186 19.00112
  -0.750377 -5.935758 -1.481948 15.7503

  praw <- rowSums(apply(null1.ind, 2, function(x) x > res[[1]][17, -c(2,8,10)])) / NCOL(null1.ind)
  (p.value <- ifelse(praw > 0.5, 1-praw, praw))
  #[1] 0.243 0.003 0.000 0.000 0.243 0.000 0.243 0.000

  ## let's look at the specialisation values of an extreme generalist, Corynura.prothysteres in
  Safariland (web 1):
  null1.ind <- null.res.1[, 24, ]
  (res[[1]][24, -c(2,8,10)] - apply(null1.ind, 1, mean))/apply(null1.ind, 1, sd)
  #
  # degree strength PSI NSI
  # betweenness closeness partner.diversity d
  #Corynura.prothysteres 0.004130498 -0.2206802 0.03651698 0.7250512 12.73813
  8.95974 0.03052268 1.203328
  praw <- rowSums(apply(null1.ind, 2, function(x) x > res[[1]][24, -c(2,8,10)])) / NCOL(null1.ind)
  (p.value <- ifelse(praw > 0.5, 1-praw, praw))
  #[1] 0.218 0.389 0.209 0.166 0.000 0.000 0.322 0.109

  # the 8 indices for all pollinators in Safariland:
  write.table(signif(res[[1]][,-c(2,8,10)], 3), file="obs.vals.Safariland.txt", sep="\t")

  # now z-scores for all pollinators:
  z.scores <- matrix(NA, 8, 27)
  p.values <- matrix(NA, 8, 27)
  for (i in 1:27){
    null1.ind <- null.res.1[, i, ]
    z.scores[,i] <- as.numeric(res[[1]][i, -c(2,8,10)] - apply(null1.ind, 1, mean,
na.rm=T))/apply(null1.ind, 1, sd, na.rm=T)

    praw <- rowSums(apply(null1.ind, 2, function(x) x > res[[1]][i, -c(2,8,10)])) / NCOL(null1.ind)
    p.values[,i] <- ifelse(praw > 0.5, 1-praw, praw)

    print(i)
  }

```

```

#-----
# d' distribution:
data(memmott1999)
ind.names <- c("degree", "NSI", "BC", "CC", "strength", "PSI", "H", "d'")
ds <- dfun(t(memmott1999))
hist(ds$dprime)
plot(density(ds$dprime, from=0, to=0.8), main="distribution of d' in Memmott 1999", las=1,
tcl=-0.5, log="", xlim=c(0,1))
rug(ds$dprime)

#require(MASS)
#fitdistr(unname(ds$dprime), dbeta, list(shape1=2, shape2=2))

mem <- res[[7]][,c(1,5,6,7,3,4,9,11)]
par(mfrow=c(2,4), mar=c(3,2.8,2,.2))
for (i in 1:8) hist(mem[,i], main=ind.names[i], col="grey80", las=1)
mtext("a", side=3, line=-2, outer=TRUE, adj=0.02, cex=2, font=2)

# same plot for one null model run:
set.seed(42)
mem.null <- nullmodel(memmott1999, 1)[[1]]
mem.null.res <- specieslevel(mem.null, index="ALLBUTD")$"higher trophic
level"[-c(2,4,9,11)][,c(1,4:6,2,3,7:8)]
par(mfrow=c(2,4), mar=c(3,2.8,2,.2))
for (i in 1:8) hist(mem.null.res[,i], main=ind.names[i], col="grey80", las=1)
mtext("b", side=3, line=-2, outer=TRUE, adj=0.02, cex=2, font=2)

#-----
# specialisation shift:
webs <- data(package="bipartite")$results[,3]
webs <- webs[c(1,15:21)]
data(list=webs)

specs <- NA
for (i in 1:length(webs)){
  specs <- c(specs, colnames(get(webs[i])))
}
sort(table(specs)) # two species in all 8 networks: Bombus dahlbomii and Syrphus octomaculatus

# for Bombus dahlbomii calculate the 8 specialisation indices in all 8 networks:
calc.ind <- function(web, focal.poll="Bombus dahlbomii") {
  # helper function to calculate the specialisation indices only for Bombus dahlbomii
  out <- specieslevel(web, index=c("degree", "NS", "BC", "CC", "strength", "PSI", "diversity",

```

```

"d"))$"higher trophic level"[focal.poll,]
  unlist(out)
}
ind.names <- c("degree", "NSI", "BC", "CC", "strength", "PSI", "H", "d")
vaz.obs <- t(sapply(webs, function(x) calc.ind(get(x))))
cattle <- c("grazed", "ungrazed")[c(2,1,1,2,2,1,2,1)] #taken from V 路 zquez & Simberloff
(2003), Fig. 1)
par(mfrow=c(2,4), mar=c(3,3,3,1))
for (i in c(1,4:6,2,3,7,8)) boxplot(unlist(vaz.obs[,i]) ~ cattle, main=ind.names[i], las=1, tcl=0.5,
notch=FALSE, col="grey80")

# now run the null models for these webs, adding another dimension of complexity:

>null.array <- array(NA, dim=c(8,8,1000), dimnames=list(rownames(vaz.obs), colnames(vaz.obs),
as.character(1:1000)))
#for (i in webs){
# web <- get(i)
# spec.index <- which(colnames(web) == "Bombus dahlbomii")
# null <- nullmodel(web, dim(null.array)[3])
# null.vals <- sapply(null, calc.ind, focal.poll=spec.index)
# null.array[i,,] <- unname(null.vals)
#}
#save(null.array, file="vazquez_null.array.Rdata")
load("vazquez_null.array.Rdata")

z.mat <- vaz.obs - apply(null.array, c(1,2), mean) /apply(null.array, c(1,2), sd)

matrix(rowSums(apply(null.array, 3, function(x) x > vaz.obs)), 8,8, byrow=F)

## make a panel plot depicting the 64 combinations of networks and indices, nullmodels vs.
observed
# make the plot window LARGE (at least A4)!
windows(width=15, height=10)
par(mfrow=c(8,8), mar=c(2,2.8,1,0.5), oma=c(0,2.5,2.8,0))

for (i in c(1,4,6,8,2,3,5,7)){#loop through the 8 sites, ungrazed first
  for (j in c(1,4:6,2:3,7:8)){# loop through the 8 indices
    xlims <- range(c(null.array[,j,], vaz.obs[,j]), na.rm=T)*c(0.9, 1.1)
#   xlims <- range(c(null.array[i,j,], vaz.obs[i,j]), na.rm=T)*c(0.9, 1.1)
    hist(null.array[i,j,], col={if(i %in%c(1,4,6,8)) "grey80" else "grey40"}, las=1, main="",
xlim=xlims)#paste(ind.names[j], rownames(z.mat)[i]))
    abline(v=vaz.obs[i,j], col="red", lwd=2)
  } #grazed sites are dark grey, ungrazed light grey
}
}

```

```

mtext("degree", side=3, line=0, outer=T, cex=1.5, font=3, at=0.12, adj=1)
mtext("NSI", side=3, line=0, outer=T, cex=1.5, font=3, at=0.11*2, adj=1)
mtext("BC", side=3, line=0, outer=T, cex=1.5, font=3, at=0.115*3, adj=1)
mtext("CC", side=3, line=0, outer=T, cex=1.5, font=3, at=0.12*4, adj=1)
mtext("strength", side=3, line=0, outer=T, cex=1.5, font=3, at=0.123*5, adj=1)
mtext("PSI", side=3, line=0, outer=T, cex=1.5, font=3, at=0.12*6, adj=1)
mtext("H", side=3, line=0, outer=T, cex=1.5, font=3, at=0.119*7, adj=1)
mtext("d", side=3, line=0, outer=T, cex=1.5, font=3, at=0.12*8, adj=1)

mtext("Safariland", side=2, line=0, outer=T, cex=1, font=1, at=8*0.125, adj=1)
mtext("vazllao", side=2, line=0, outer=T, cex=1, font=1, at=7*0.121, adj=1)
mtext("vazmasnc", side=2, line=0, outer=T, cex=1, font=1, at=6*0.124, adj=1)
mtext("vazquenc", side=2, line=0, outer=T, cex=1, font=1, at=5*0.123, adj=1)
mtext("vazarr", side=2, line=0, outer=T, cex=1, font=1, at=4*0.12, adj=1)
mtext("vazcer", side=2, line=0, outer=T, cex=1, font=1, at=3*0.122, adj=1)
mtext("vazmasc", side=2, line=0, outer=T, cex=1, font=1, at=2*0.124, adj=1)
mtext("vazquec", side=2, line=0, outer=T, cex=1, font=1, at=0.122, adj=1)

# now we examine the effect of grazing on the raw specialisation indices:
obs.data <- cbind.data.frame(cattle, vaz.obs)
colnames(obs.data)

# Next, we calculate the difference between the observed and the perfect generalist (i.e. mean of
null models):
windows(width=12.5, height=8)
par(tcl=0, mar=c(3,3,1,1), mfrow=c(2,4))

# degree:-----
index = "species.degree"
diffs <- obs.data[,index] - rowMeans(null.array[,index,])
zs <- (obs.data[,index] - rowMeans(null.array[,index,])) / apply(null.array[,index,], 1, sd)
zs <- ifelse(is.infinite(zs), NA, zs)/5
#plot all 3:
data2plot <- cbind(cbind(raw=obs.data[,index], diff=diffs, z=zs)[c(1,4,6,8),],
cbind(raw=obs.data[,index], diff=diffs, z=zs)[c(2,3,5,7),])
boxplot(data2plot, at=c(1,3,5,1.8,3.8,5.8), col=rep(c("grey80", "grey20"), each=3), las=1,
names=rep("", 6))
abline(h=0)
axis(side=2, at=seq(-6,6,by=2), tcl=0.5, labels=FALSE)
mtext("raw", side=1, line=0.75, outer=F, at=1.4, font=2, cex=1.5)
mtext("diff", side=1, line=0.75, outer=F, at=3.4, font=2, cex=1.5)
mtext("z/5", side=1, line=0.75, outer=F, at=5.4, font=2, cex=1.5)
legend("bottomleft", "degree", bty="n", cex=2, inset=c(-0.05,-0.025))
legend("topright", c("ungrazed", "grazed"), col=c("grey80", "grey20"), pch=15, bty="o", cex=2)

```

```

# NSI:-----
index = "node.specialisation.index"
diffs <- obs.data[,index] - rowMeans(null.array[,index,])
zs <- (obs.data[,index] - rowMeans(null.array[,index,])) / apply(null.array[,index,], 1, sd)
zs <- ifelse(is.infinite(zs), NA, zs)/20
#plot all 3:
data2plot <- cbind(cbind(raw=obs.data[,index], diff=diffs, z=zs)[c(1,4,6,8),],
cbind(raw=obs.data[,index], diff=diffs, z=zs)[c(2,3,5,7),])
boxplot(data2plot, at=c(1,3,5,1.8,3.8,5.8), col=rep(c("grey80", "grey20"), each=3), las=1,
names=rep("", 6))
axis(side=2, at=c(0,0.5,1,1.5,2,2.5), tcl=0.5, labels=FALSE)
mtext("raw", side=1, line=0.75, outer=F, at=1.4, font=2, cex=1.5)
mtext("diff", side=1, line=0.75, outer=F, at=3.4, font=2, cex=1.5)
mtext("z/20", side=1, line=0.75, outer=F, at=5.4, font=2, cex=1.5)
legend("topleft", "NSI", bty="n", cex=2, inset=c(-0.05,-0.05))

# BC:-----
index = "betweenness"
diffs <- obs.data[,index] - rowMeans(null.array[,index,])
zs <- (obs.data[,index] - rowMeans(null.array[,index,])) / apply(null.array[,index,], 1, sd)
zs <- ifelse(is.infinite(zs), NA, zs)/100
#plot all 3:
data2plot <- cbind(cbind(raw=obs.data[,index], diff=diffs, z=zs)[c(1,4,6,8),],
cbind(raw=obs.data[,index], diff=diffs, z=zs)[c(2,3,5,7),])
boxplot(data2plot, at=c(1,3,5,1.8,3.8,5.8), col=rep(c("grey80", "grey20"), each=3), las=1,
names=rep("", 6))
abline(h=0)
boxplot(data2plot, at=c(1,3,5,1.8,3.8,5.8), col=rep(c("grey80", "grey20"), each=3), las=1,
names=rep("", 6), add=T)
axis(side=2, at=seq(0,0.3,by=0.1), tcl=0.5, labels=FALSE)
mtext("raw", side=1, line=0.75, outer=F, at=1.4, font=2, cex=1.5)
mtext("diff", side=1, line=0.75, outer=F, at=3.4, font=2, cex=1.5)
mtext("z/100", side=1, line=0.75, outer=F, at=5.4, font=2, cex=1.5)
legend("topleft", "BC", bty="n", cex=2, inset=c(-0.05,-0.05))

# CC:-----
index = "closeness"
diffs <- obs.data[,index] - rowMeans(null.array[,index,])
zs <- (obs.data[,index] - rowMeans(null.array[,index,])) / apply(null.array[,index,], 1, sd)
zs <- ifelse(is.infinite(zs), NA, zs)/100
#plot all 3:
data2plot <- cbind(cbind(raw=obs.data[,index], diff=diffs, z=zs)[c(1,4,6,8),],
cbind(raw=obs.data[,index], diff=diffs, z=zs)[c(2,3,5,7),])

```

```

boxplot(data2plot, at=c(1,3,5,1.8,3.8,5.8), col=rep(c("grey80", "grey20"), each=3), las=1,
names=rep("", 6))
abline(h=0)
axis(side=2, at=seq(0,0.15,by=0.05), tcl=0.5, labels=FALSE)
mtext("raw", side=1, line=0.75, outer=F, at=1.4, font=2, cex=1.5)
mtext("diff", side=1, line=0.75, outer=F, at=3.4, font=2, cex=1.5)
mtext("z/100", side=1, line=0.75, outer=F, at=5.4, font=2, cex=1.5)
legend("topleft", "CC", bty="n", cex=2, inset=c(-0.05,-0.05))

```

```

# strength:-----
#par(tcl=0, mar=c(3,3,1,1))#, mfrow=c(2,4)
index = "strength"
diffs <- obs.data[,index] - rowMeans(null.array[,index,])
zs <- (obs.data[,index] - rowMeans(null.array[,index,])) / apply(null.array[,index,], 1, sd)
zs <- ifelse(is.infinite(zs), NA, zs)
#plot all 3:
data2plot <- cbind(cbind(raw=obs.data[,index], diff=diffs, z=zs)[c(1,4,6,8),],
cbind(raw=obs.data[,index], diff=diffs, z=zs)[c(2,3,5,7),])
boxplot(data2plot, at=c(1,3,5,1.8,3.8,5.8), col=rep(c("grey80", "grey20"), each=3), las=1,
names=rep("", 6))
abline(h=0)
boxplot(data2plot, at=c(1,3,5,1.8,3.8,5.8), col=rep(c("grey80", "grey20"), each=3), las=1,
names=rep("", 6), add=TRUE)
axis(side=2, at=c(-4,-2, 0,2), tcl=0.5, labels=FALSE)
mtext("raw", side=1, line=0.75, outer=F, at=1.4, font=2, cex=1.5)
mtext("diff", side=1, line=0.75, outer=F, at=3.4, font=2, cex=1.5)
mtext("z", side=1, line=0.75, outer=F, at=5.4, font=2, cex=1.5)
legend("bottomleft", "strength", bty="n", cex=2, inset=c(-0.05,0))

```

```

# PSI:-----
#par(tcl=0, mar=c(3,3,1,1))#, mfrow=c(2,4)
index = "Pollination.Service.Index.PSI"
diffs <- obs.data[,index] - rowMeans(null.array[,index,])
zs <- (obs.data[,index] - rowMeans(null.array[,index,])) / apply(null.array[,index,], 1, sd)
zs <- ifelse(is.infinite(zs), NA, zs)/100
#plot all 3:
data2plot <- cbind(cbind(raw=obs.data[,index], diff=diffs, z=zs)[c(1,4,6,8),],
cbind(raw=obs.data[,index], diff=diffs, z=zs)[c(2,3,5,7),])
boxplot(data2plot, at=c(1,3,5,1.8,3.8,5.8), col=rep(c("grey80", "grey20"), each=3), las=1,
names=rep("", 6))
axis(side=2, at=c(0,0.5,1,1.5,2,2.5), tcl=0.5, labels=FALSE)
mtext("raw", side=1, line=0.75, outer=F, at=1.4, font=2, cex=1.5)
mtext("diff", side=1, line=0.75, outer=F, at=3.4, font=2, cex=1.5)
mtext("z/100", side=1, line=0.75, outer=F, at=5.4, font=2, cex=1.5)

```

```
legend("topleft", "PSI", bty="n", cex=2, inset=c(-0.05,-0.05))
```

```
# partner.diversity:-----  
#par(tcl=0, mar=c(3,3,1,1))#, mfrow=c(2,4)  
index = "partner.diversity"  
diffs <- obs.data[,index] - rowMeans(null.array[,index,])  
zs <- (obs.data[,index] - rowMeans(null.array[,index,])) / apply(null.array[,index,], 1, sd)  
zs <- ifelse(is.infinite(zs), NA, zs)/20  
#plot all 3:  
data2plot <- cbind(cbind(raw=obs.data[,index], diff=diffs, z=zs)[c(1,4,6,8),],  
cbind(raw=obs.data[,index], diff=diffs, z=zs)[c(2,3,5,7),])  
boxplot(data2plot, at=c(1,3,5,1.8,3.8,5.8), col=rep(c("grey80", "grey20"), each=3), las=1,  
names=rep("", 6))  
abline(h=0)  
axis(side=2, at=seq(-1,1,by=0.5), tcl=0.5, labels=FALSE)  
mtext("raw", side=1, line=0.75, outer=F, at=1.4, font=2, cex=1.5)  
mtext("diff", side=1, line=0.75, outer=F, at=3.4, font=2, cex=1.5)  
mtext("z/20", side=1, line=0.75, outer=F, at=5.4, font=2, cex=1.5)  
legend("topright", "H", bty="n", cex=2, inset=c(0.05,-0.05))
```

```
# d':-----  
#par(tcl=0, mar=c(3,3,1,1))#, mfrow=c(2,4)  
index = "d"  
diffs <- obs.data[,index] - rowMeans(null.array[,index,])  
zs <- (obs.data[,index] - rowMeans(null.array[,index,])) / apply(null.array[,index,], 1, sd)  
zs <- ifelse(is.infinite(zs), NA, zs)/100  
#plot all 3:  
data2plot <- cbind(cbind(raw=obs.data[,index], diff=diffs, z=zs)[c(1,4,6,8),],  
cbind(raw=obs.data[,index], diff=diffs, z=zs)[c(2,3,5,7),])  
boxplot(data2plot, at=c(1,3,5,1.8,3.8,5.8), col=rep(c("grey80", "grey20"), each=3), las=1,  
names=rep("", 6))  
axis(side=2, at=c(0,0.5,1,1.5,2,2.5), tcl=0.5, labels=FALSE)  
mtext("raw", side=1, line=0.75, outer=F, at=1.4, font=2, cex=1.5)  
mtext("diff", side=1, line=0.75, outer=F, at=3.4, font=2, cex=1.5)  
mtext("z/100", side=1, line=0.75, outer=F, at=5.4, font=2, cex=1.5)  
legend("topleft", "d'", bty="n", cex=2, inset=c(-0.05,-0.05))
```

# The problem is that the specialisation index values are affected by the network properties (such as number of species, sampling intensity, etc.). Thus, we need to calculate a corrected value (e.g. the z-score). Or, we explore the t-statistic between the two grazing types but compute the significance level from the null model. This latter approach is also usable for nonnormally distributed data, using Fisher's signed rank test, rather than the t-test.

```

obs.t.PSI <- t.test(obs.data$Pollination.Service.Index.PSI[obs.data$cattle=="grazed"],
obs.data$Pollination.Service.Index.PSI[obs.data$cattle=="ungrazed"], var.equal=F)$statistic
null.t.PSI <- apply(null.array[, 3, ], 2, function(x) t.test(x[obs.data$cattle=="grazed"],
x[obs.data$cattle=="ungrazed"]))$statistic)
plot(density(null.t.PSI)); abline(v=obs.t.PSI, col="red")
sum(null.t.PSI > obs.t.PSI) # 25
# this means that the observed difference statistics is significantly higher than would be expected
for a generalist

```

```

obs.degree.dev <- anova(glm(species.degree ~ cattle, family=poisson, data=obs.data),
test="Chisq")$Deviance[2]
obs.NSI.F <- anova(lm(node.specialisation.index ~ cattle, data=obs.data))$"F value"[1]
obs.BC.F <- anova(lm(betweenness ~ cattle, data=obs.data))$"F value"[1]
obs.CC.F <- anova(lm(closeness ~ cattle, data=obs.data))$"F value"[1]
obs.strength.F <- anova(lm(strength ~ cattle, data=obs.data))$"F value"[1]
obs.PSI.F <- anova(lm(Pollination.Service.Index.PSI ~ cattle, data=obs.data))$"F value"[1]
obs.H.F <- anova(lm(partner.diversity ~ cattle, data=obs.data))$"F value"[1]
obs.d.F <- anova(lm(d ~ cattle, data=obs.data))$"F value"[1]
# none of the indices hints at a change in specialisation of Bombus dahlbomii under change in
grazing

```

```

# use nullmodels to get a F-value distribution:
dimnames(null.array)[1:2]

```

```

null.degree.dev <- apply(null.array[, 1, ], 2, function(x) anova(glm(x ~ cattle, family=poisson,
data=obs.data))$Deviance[2])
sum(null.degree.dev > obs.degree.dev) #150
null.NSI.F <- apply(null.array[, 4, ], 2, function(x) anova(lm(x ~ cattle, data=obs.data))$"F
value"[1])
sum(null.NSI.F > obs.NSI.F) #964
null.BC.F <- apply(null.array[, "betweenness", ], 2, function(x) anova(lm(x ~ cattle,
data=obs.data))$"F value"[1])
sum(null.BC.F > obs.BC.F) #987
null.CC.F <- apply(null.array[, "closeness", ], 2, function(x) anova(lm(x ~ cattle,
data=obs.data))$"F value"[1])
sum(null.CC.F > obs.CC.F) #943
null.strength.F <- apply(null.array[, "strength", ], 2, function(x) anova(lm(x ~ cattle,
data=obs.data))$"F value"[1])
sum(null.strength.F > obs.strength.F) #131
null.PSI.F <- apply(null.array[, 3, ], 2, function(x) anova(lm(x ~ cattle, data=obs.data))$"F
value"[1])
sum(null.PSI.F > obs.PSI.F) #975
null.H.F <- apply(null.array[, 7, ], 2, function(x) anova(lm(x ~ cattle, data=obs.data))$"F
value"[1])

```



```

sum(null.H.F > obs.H.F) #893
null.d.F <- apply(null.array[, 8, ], 2, function(x) anova(lm(x ~ cattle, data=obs.data))$F
value"[1])
sum(null.d.F > obs.d.F) #686

#####
#####
#####
# this is nothing new: it repeats the analysis of Bombus for a different species:
#####
#####
webs <- data(package="bipartite")$results[,3]
webs <- webs[c(1,15:21)]
data(list=webs)

specs <- NA
for (i in 1:length(webs)){
  specs <- c(specs, colnames(get(webs[i])))
}
sort(table(specs)) # two species in all 8 networks: Bombus dahlbomii and Syrphus octomaculatus
# We pick a third one here, Vespula germanica, which is present in 7 of the 8 networks.

# for Vespula germanica calculate the 8 specialisation indices in all 8 networks:
calc.ind <- function(web, focal.poll="Vespula germanica") {
  # helper function to calculate the specialisation indices only for Vespula germanica
  out <- specieslevel(web, index=c("degree", "NS", "BC", "CC", "strength", "PSI", "diversity",
"d"))$"higher trophic level"[focal.poll,]
  unlist(out)
}
ind.names <- c("degree", "NSI", "BC", "CC", "strength", "PSI", "H", "d")
vaz.obs <- t(sapply(webs, function(x) calc.ind(get(x))))
cattle <- c("grazed", "ungrazed")[c(2,1,1,2,2,1,2,1)] #taken from V 路 zquez & Simberloff
(2003), Fig. 1)
par(mfrow=c(2,4), mar=c(3,3,3,1))
for (i in c(1,4:6,2,3,7,8)) boxplot(unlist(vaz.obs[,i]) ~ cattle, main=ind.names[i], las=1, tcl=0.5,
notch=FALSE, col="grey80")

# now run the null models for these webs, adding another dimension of complexity:

#null.array <- array(NA, dim=c(8,8,1000), dimnames=list(rownames(vaz.obs), colnames(vaz.obs),
as.character(1:1000)))
#for (i in webs){
# web <- get(i)

```

```

# spec.index <- which(colnames(web) == "Vespula germanica")
# if (length(spec.index) == 0){
#   null.array[i,,] <- matrix(NA, 8, 1000)
# } else {
#   null <- nullmodel(web, dim(null.array)[3])
#   null.vals <- sapply(null, calc.ind, focal.poll=spec.index)
#   null.array[i,,] <- unname(null.vals)
# }
# print(i)
#}
#save(null.array, file="vazquez_Vespula.germanica_null.array.Rdata")
load("vazquez_Vespula.germanica_null.array.Rdata")

z.mat <- vaz.obs - apply(null.array, c(1,2), mean) /apply(null.array, c(1,2), sd)

matrix(rowSums(apply(null.array, 3, function(x) x > vaz.obs)), 8,8, byrow=F)

## make a panel plot depicting the 64 combinations of networks and indices, nullmodels vs.
observed
# make the plot window LARGE (at least A4!)
windows(width=15, height=10)
par(mfrow=c(7,8), mar=c(2,2.8,1,0.5), oma=c(0,2.5,2.8,0))
for (i in c(1,4,6,8,2,3,7)) {#loop through the 8 sites, ungrazed first
  for (j in c(1,4:6,2:3,7:8)) {# loop through the 8 indices
    xlims <- range(c(null.array[,j,], vaz.obs[,j]), na.rm=T)*c(0.9, 1.1)
    hist(null.array[i,j,], col={if(i %in% c(1,4,6,8)) "grey80" else "grey40"}, las=1, main="",
    xlim=xlims)#paste(ind.names[j], rownames(z.mat)[i]))
    abline(v=vaz.obs[i,j], col="red", lwd=2)
  } #grazed sites are dark grey, ungrazed light grey
}
mtext("degree", side=3, line=0, outer=T, cex=1.5, font=3, at=0.12, adj=1)
mtext("NSI", side=3, line=0, outer=T, cex=1.5, font=3, at=0.11*2, adj=1)
mtext("BC", side=3, line=0, outer=T, cex=1.5, font=3, at=0.115*3, adj=1)
mtext("CC", side=3, line=0, outer=T, cex=1.5, font=3, at=0.12*4, adj=1)
mtext("strength", side=3, line=0, outer=T, cex=1.5, font=3, at=0.123*5, adj=1)
mtext("PSI", side=3, line=0, outer=T, cex=1.5, font=3, at=0.12*6, adj=1)
mtext("H", side=3, line=0, outer=T, cex=1.5, font=3, at=0.119*7, adj=1)
mtext("d", side=3, line=0, outer=T, cex=1.5, font=3, at=0.12*8, adj=1)

mtext("Safariland", side=2, line=0, outer=T, cex=1, font=1, at=7*0.14, adj=1)
mtext("vazllao", side=2, line=0, outer=T, cex=1, font=1, at=6*0.14, adj=1)
mtext("vazmasnc", side=2, line=0, outer=T, cex=1, font=1, at=5*0.14, adj=1)
mtext("vazquenc", side=2, line=0, outer=T, cex=1, font=1, at=4*0.14, adj=1)
mtext("vazarr", side=2, line=0, outer=T, cex=1, font=1, at=3*0.14, adj=1)

```

```
mtext("vazcer", side=2, line=0, outer=T, cex=1, font=1, at=2*0.14, adj=1)
#mtext("vazmasc", side=2, line=0, outer=T, cex=1, font=1, at=2*0.124, adj=1)
mtext("vazquec", side=2, line=0, outer=T, cex=1, font=1, at=0.14, adj=1)
```