ORIGINAL ARTICLE

Sources of prey availability data alter interpretation of outputs from prey choice null networks

Jordan P. Cuff[1,2,3] | Maximillian P. T. G. Tercel[1,4] | Fredric M. Windsor[1,3] | Ben S. J. Hawthorne[3] | Peter A. Hambäck[5] | James R. Bell[2] | William O. C. Symondson[1] | Ian P. Vaughan[1]

[1]School of Biosciences, Cardiff University, Cardiff, UK

[2]Rothamsted Insect Survey, Rothamsted Research, West Common, Hertfordshire, UK

[3]School of Natural and Environmental Sciences, Newcastle University, Newcastle upon Tyne, UK

[4]Durrell Wildlife Conservation Trust, Les Augrès Manor, Trinity, Jersey

[5]Department of Ecology, Environment and Plant Sciences, Stockholm University, Stockholm, Sweden

Correspondence
Jordan P. Cuff, School of Natural and Environmental Sciences, Newcastle University, Newcastle-upon-Tyne, NE1 7RU, UK.
Email: jordancuff@gmail.com

Abstract

1. Null models provide a valuable baseline against which fundamental ecological hypotheses can be tested and foraging choices that cannot be explained by neutral processes or sampling biases can be highlighted. In this way, null models can advance our understanding beyond simplistic dietary descriptions to identify drivers of interactions. This method, however, requires estimates of resource availability, which are generally imperfect representations of highly dynamic systems. Optimising method selection is crucial for study design, but the precise effects of different resource availability data on the efficacy of null models are poorly understood.

2. Using spider–prey networks as a model, we used prey abundance (suction sample) and activity density (sticky trap) data, and combinations of the two, to simulate null networks. We compared null diet composition, network properties (e.g., connectance and nestedness) and deviations of simulations from metabarcoding-based spider dietary data to ascertain how different prey availability data alter ecological interpretation.

3. Different sampling methods produced different null networks and inferred distinct prey selectivity. Null models based on prey abundance and combined frequency-of-occurrence data generated null diet compositions, which more closely resembled the diet composition determined by metabarcoding. Null models based on prey abundance, activity density and proportionally combined data generated null network properties most like the networks constructed via dietary metabarcoding.

4. We show that survey method choice impacts all aspects of null network analyses, the precise effects varying between methods but ultimately altering ecological interpretation by increasing disparity in network properties or trophic niches between null and directly constructed networks. Merging datasets can generate

more complete prey availability data but is not a panacea because it introduces different biases. The choice of method should reflect the research hypotheses and study system being investigated. Ultimately, survey methods should emulate the foraging mode of the focal predator as closely as possible, informed by the known ecology, natural history and behaviour of the predator.

INTRODUCTION

Trophic interactions are fundamental to evolutionary and ecological processes (Vázquez & Aizen, 2003). The identity and frequency of interactions are determined by resource preferences and choices, and assessing resource choice is crucial in predicting and understanding trophic dynamics, behaviour and ecology more broadly (Cuff, Tercel, et al., 2022). For example, assessing the structure of these interactions can provide insight into network assembly, functioning and response to perturbation (Allesina et al., 2008). It is, however, difficult to contextualise independent observations of resource choice, given the highly system-dependent nature of such patterns (Vázquez & Aizen, 2003).

Null models facilitate testing of ecological hypotheses by comparing observations with null expectations that represent specific mechanisms (e.g., trait-dependent interactions) and generate random data, with various successful applications across ecology and biogeography (Gotelli, 2001; Gotelli & Graves, 1996). Null modelling can reveal when trophic interactions deviate from random by providing a baseline representation of both the frequency and identity of interactions as would be generated by random foraging (Vaughan et al., 2018; Vázquez & Aizen, 2003). Most simply, this approach can assess how trophic interactions relate to prey abundances in which the most abundant prey are likely to be the most commonly consumed (Agustí et al., 2003; Cuff, Tercel, et al., 2022; Vaughan et al., 2018). This approach can also provide valuable information on predators (e.g., behaviour, preferences, nutritional requirements), prey (e.g., palatability, detectability, defences, escape ability) and the trophic network in which they exist (e.g., network metrics like nestedness and linkage density), provided that input data represent the relative availability of different prey types to the focal predator. Individual-based approaches to null modelling (i.e., generating null data for each individual), by not fixing network properties a priori, reduce constraints on null network generation for comparison against observed networks, leading to more realistic and stochastic null networks (Grimm & Berger, 2016). These approaches can highlight network structures that are not generated by neutral mechanisms, or arise as an artefact of sampling methods, by maintaining the characteristics of the observed data (e.g., the degree of each consumer; Blüthgen et al., 2008; Vaughan et al., 2018). The interaction identities and frequencies, and the concomitant network structures, can nevertheless differ greatly depending on the input data used.

If accurately constructed, null models can elucidate the fundamental mechanisms underpinning species interactions. Null model approaches have therefore been used to explore a range of research questions including prey selectivity changes in response to perturbations (Cuff et al., 2021), seasonal variations in prey availability (Gajski et al., 2023; Verschut et al., 2019), host–parasite–parasitoid specialisation (Ramirez et al., 2022), pollinator preferences across different landscapes (Gómez-Martínez et al., 2022), changes in foraging ecology corresponding with weather conditions (Cuff, Windsor, et al., 2023) and plant–invertebrate commensalisms (Cuff, Evans, et al., 2022). Alongside taxonomic units (e.g., species), the nodes in these networks can represent data such as consumer age class (Davies et al., 2022), functional groups (Méndez-Castro et al., 2020) and environmental context (Cuff, Windsor, et al., 2023), increasing the value and applicability of these models. Many of these examples, particularly those concerning plant resources, assess preferences of active consumers for static resources, for which resource availability is relatively straightforward to estimate, but the interpretation is confounded when both the consumer and the resource are mobile (e.g., predator–prey systems), and little guidance exists regarding resource availability estimation. To avoid biases caused by improper resource availability estimates, it is paramount that the choice of sampling method aims to closely match the prey available to a predator (Table 1).

In this article, we show how survey method choice affects null-model-based resource choice analyses, with significant implications for broader studies relating resource and interaction data to understand drivers of interactions such as predator foraging ecology. Given that spiders can employ both active hunting and sit-and-wait predation, two data types representing prey availability were collected. Prey activity density and abundance samples were collected at each sampling location using sticky traps and suction sampling, respectively. These sampling methods were then used to generate null networks based on dietary data, but we also used two different methods to combine these two estimates into one prey availability index. Using these different prey availability estimates, we tested the following hypotheses: (i) survey method choice affects the results of null model analysis by altering the identity and frequency of simulated trophic interactions and ultimately network properties; (ii) different measures of prey availability (i.e., abundance and activity density) differ in their relationship to observed interactions, reflecting their emulation of the foraging behaviour of the consumer and (iii) the type of observed interaction data used alters inferred foraging choices and the

TABLE 1 Methods used to generate invertebrate prey availability data, example studies they might be used for and some specific considerations.

Data type	Example method	Example study	Specific considerations
Abundance	Vacuum sampling	Predator–prey networks in short vegetation	Duration, area and time of day, suction power, nozzle circumference, prevalence of litter and other physical obstructions, substrate
	Sweep netting	Predator–prey networks in long grass	Duration, area, time of day, effort applied, vegetation type/composition
	Branch beating	Predator–prey networks associated with trees	Duration, effort and number of beats, time of day, beating tray position and area, branch complexity and foliage
Activity density	Sticky trapping	Web-building spider–prey networks	Trapping duration, trap area and colour, position, height and orientation, adhesive quality
	Pitfall trapping	Ground-active predator–prey networks	Circumference and depth of trap, colour, trap volume, position, lures/baits, trapping fluid
	Malaise trapping	Insect–pitcher plant networks	Height, colour and size of trap, lures/baits, trapping fluid

Note: Methods should be selected to best reflect the experience of focal predators and to reduce eventual biases. Note that all methods may be sensitive to weather conditions.

structure of null networks. Through discussion of these hypotheses, we also provide guidance for researchers embarking on relevant studies and discuss how to overcome inaccuracies introduced by survey method biases.

MATERIALS AND METHODS

Fieldwork

Data collection was described previously by Cuff, Tercel, et al. (2022). This study pertains to a subset of those data, collected between 1st May and 9th July 2018 at 19 separate locations, for which paired sticky trap and vacuum sample data were collected (described below). Briefly, money spiders (Araneae: Linyphiidae) and wolf spiders (Araneae: Lycosidae) were visually located along transects in two adjacent barley fields at Burdons Farm, Wenvoe in South Wales (51°26′24.8″ N, 3°16′17.9″ W) and collected from webs and the ground. Transects were randomly distributed across the entire field. Along these transects, separate 4-m^2 quadrats, at least 10 m apart, were searched, and all observed linyphiids and lycosids were collected. Spiders were placed in 100% ethanol using an aspirator, regularly changing the meshing to limit potential cross-contamination. Linyphiids occupying webs were prioritised for collection, but ground-active spiders were also collected. Spiders were taken to Cardiff University, transferred to fresh ethanol and stored at −80°C in 100% ethanol until DNA extraction. Extraction, amplification and sequencing of DNA, and bioinformatic analysis is described by Cuff, Tercel, et al. (2022) and Drake et al. (2022) and is also detailed in Supplementary Information 1. The resultant sequencing read counts were converted into relative proportions (all values made to sum to one within each sample), and a mean value across the two primer pairs was retained for each taxon within each sample. Relative read

abundances were converted to presence–absence data of each detected prey taxon in each individual spider, but relative read abundance data were also retained for separate analyses to compare experimental outcomes between data types.

To estimate prey availability using sticky traps, we placed one white dry 100 mm × 125 mm trap (Oecos) in the 4-m^2 quadrat centred at the position where the spider was collected. The trap was suspended with wire approximately 25 mm above the ground to catch falling, crawling and flying invertebrates and was left in place for 72 h. Collected invertebrates were identified on the traps under a stereomicroscope. To estimate prey availability using suction sampling, ground and crop stems were sampled using a 'G-vac' for approximately 30 s at each location. Suction samples were emptied into a bag and any organisms within them immediately killed with ethyl acetate and frozen for storage. Samples were subsequently sorted and placed into 70% ethanol. All invertebrates were identified to family level to match the resolution of the least resolved of the metabarcoding-derived trophic interaction data and due to difficulties associated with identification to finer taxonomic resolution for many taxa. Exceptions included springtails of the superfamily Sminthuroidea (Sminthuridae and Bourletiellidae were often indistinguishable following sampling and preservation due to the fine features necessary to distinguish them), which were left at superfamily, mites (many of which were immature or in poor condition), which were identified to order level and wasps of the superfamily Ichneumonoidea, which were identified no further due to obscurity of wing venation due to damage following sampling.

Statistical analysis

All analyses were conducted in R v4.0.3 (R Core Team, 2021) and carried out on invertebrate data at the family or superfamily level. Alongside the dietary data derived from metabarcoding, and prey
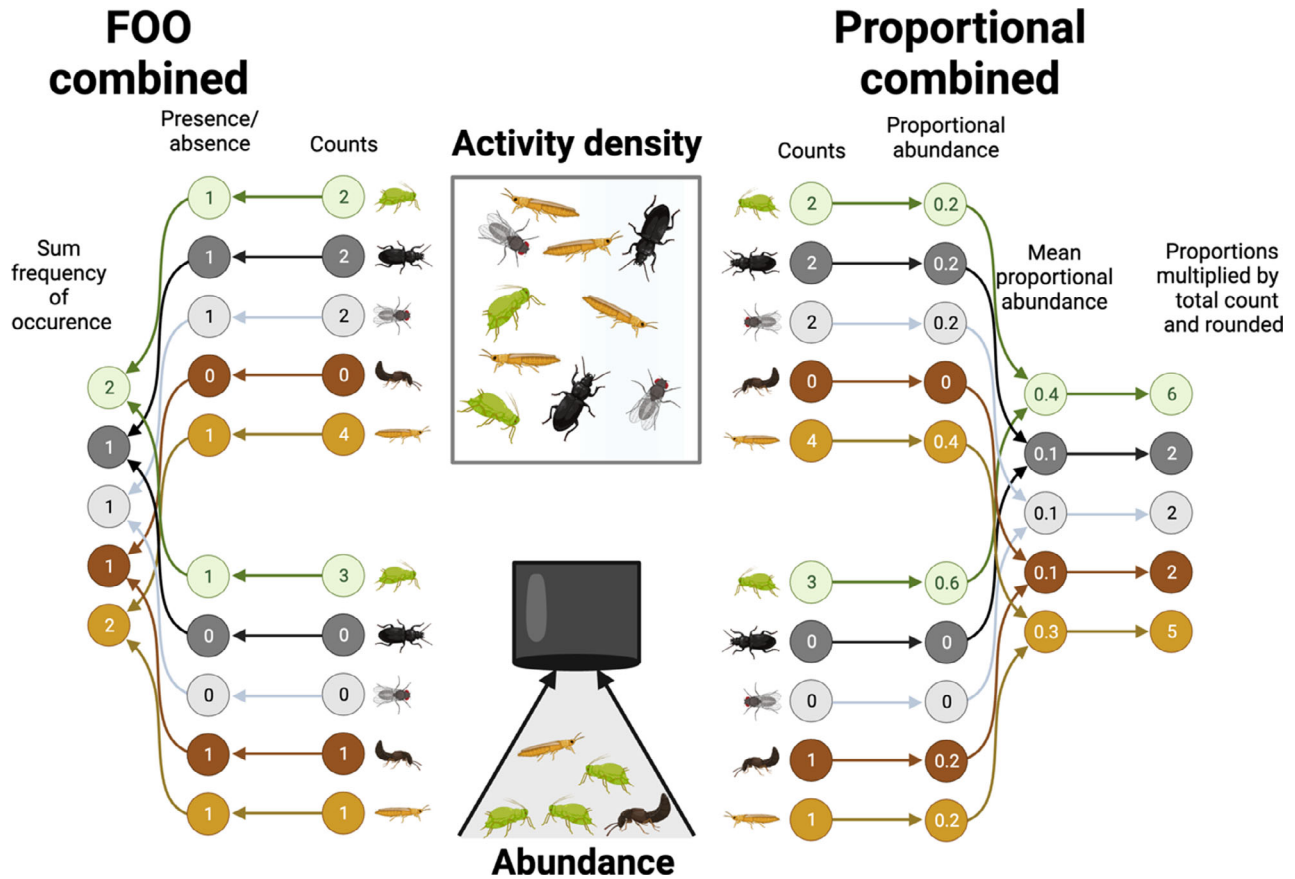
FIGURE 1 Data processing for the two combined method datasets from the original sticky trap and suction sample data. Figure created with Biorender. FOO, frequency of occurrence.

availability as determined directly by suction sampling (abundance) and sticky trapping (activity density), three additional datasets were generated where two were designed to combine data from the two trapping methods (Figure 1). The first approach simply set all invertebrate taxa detected in the field to have equal abundance, to provide a baseline against which to assess the effects of different prey availability estimates. When generating the two combined datasets (which combined the abundance and activity density data), it was apparent that simply adding them together would underrepresent one of the datasets as abundance and activity density are measured in different units. Therefore, a 'proportional combined' dataset was generated by converting counts to relative proportions of each sample (to equally weight the two methods), which were then combined by averaging proportions between the two methods for each sample, multiplied by the total count of individuals across both methods for each sample (to create realistic abundance values), and then rounded to the nearest integer (to return count data). In addition, a 'frequency of occurrence (FOO) combined' dataset was generated by converting counts to binary presence–absence values of each sample, which were then summed between the two methods for each sample, amounting to 0, 1 or 2. To assess the diversity represented by the two sampling methods and their combinations, and the completeness of those datasets, coverage-based rarefaction and extrapolation were carried out,

and Hill diversity was calculated (Chao et al., 2014; Roswell et al., 2021) using the 'iNEXT' package with families represented by frequency-of-occurrence across samples (Chao et al., 2014; Hsieh et al., 2016; Figures S1–S3). Completeness is determined by comparison of the observed diversity to the extrapolated diversity. The differences in the communities represented by these different datasets (including the dietary data) were visualised via non-metric multidimensional scaling (NMDS) using the 'metaMDS' function in the 'vegan' package (Oksanen et al., 2016) in two dimensions and 9999 simulations, with Bray–Curtis distance. Centroid coordinates for each dataset were extracted, and pairwise distances were calculated between model centroids:

$$\text{distance} = \sqrt{\left((x2 - x1)^2 + (y2 - y1)^2\right)}.$$

The remaining analyses were performed using both presence–absence and relative read abundance dietary data separately to show how differences in the treatment of the observed data are reflected in the outcomes of the analyses. Figures and outputs given in the main text relate to the presence–absence data, while relative read abundance figures and outputs are presented in the Supplementary Information. Prey preferences of spiders were analysed using network-based null models in the 'econullnetr' package (Vaughan et al., 2018) with the

PREY CHOICE NULL NETWORK DATA SOURCES

Ecological
Entomology

Royal
Entomological
Society

5

'generate_null_net' function. Econullnetr generates null models based on prey availability to predict how consumers would forage if based on the availability of resources alone. These null models are then compared against the observed interactions of consumers (in our case, interactions of spiders with their prey based on dietary metabarcoding) to ascertain the extent to which resource consumption deviates from random. In five separate null models, prey availability was represented separately by the datasets described above: abundance (suction sampling), activity density (sticky trapping), proportional combined, FOO combined and equal prey abundance.

Alongside generating null networks from which interaction structures can be investigated, null model-predicted trophic interactions, or null diet compositions, were generated via an econullnetr null model with 999 simulations and outputs extended to allow comparison of the null interactions for individual consumers (generate_null_net_indiv; Cuff, Windsor, et al., 2023). This meant that the null models produced two data types which, for clarity, we refer to as null networks (i.e., the structural properties of the interactions generated by the null models) and null diet compositions (i.e., the specific interaction partners of the spiders generated by the null models). To compare effect sizes in the comparison between null networks for each resource taxon, mean prey preference standardised effect size (SES) values were calculated from the individual spiders per model. The SES values were plotted and joined between taxa to visualise paired differences using 'ggplot' (Wickham, 2016). A visualisation of the per-individual differences in null diet composition and observed data was generated via NMDS using the 'metaMDS' function in the 'vegan' package (Oksanen et al., 2016) in two dimensions and 9999 simulations, with Euclidean distance. Centroid coordinates for each prey availability data type and the observed data were extracted, and pairwise distances were calculated between centroids using the equation presented above for the community data NMDS. The 'observed' network (i.e., the network constructed solely from dietary metabarcoding data, not necessarily the objectively 'true' network) and each null network were visualised with the associated prey choice effect sizes as a bipartite network using 'ggnetwork' (Briatte, 2021; Wickham, 2016) via an 'igraph' object (Csardi & Nepusz, 2006). The degree of each prey node, weighted nestedness and linkage density were generated using the 'bipartite' package (Dormann et al., 2008) for each network and compared visually via ggplot2.

RESULTS

Dataset description

The dietary dataset used in this study contains data from 70 individual spiders that cumulatively interacted with 25 prey families, totalling 142 individual detected interactions. Sampling datasets contain data from 19 locations, with prey abundance data, determined via suction sampling, including 4766 individual invertebrates across 61 families (93.9% complete; Figure S1) and prey activity density data,

determined via sticky trapping, including 3513 individual invertebrates across 55 families (91.5% complete; Figure S2). The prey availability data gained by combining the two measures of prey availability includes 85 families (95.1% complete; Figure S3). The five prey availability datasets were relatively distinct and varied in their compositional similarity to the directly detected dietary data, with FOO combined being the most similar (Bray–Curtis distance = 0.545), followed by abundance (0.847), proportional combined (1.136), equal prey abundance (1.245) and activity density (1.362), respectively (Figure 2).

Differences in inferred foraging ecology

Predator selectivity differed substantially between datasets (Figure 3; Figure S6). The equal prey abundance, proportional combined and FOO combined datasets generated no significantly negative effect sizes (i.e., avoidances). The two combined datasets generated selectivity results largely consistent with the abundance data, but the FOO combined data generated null networks with the fewest significant deviations from the observed data. The activity density and abundance data showed some consistency, but sometimes generated opposite patterns, and they tended to determine more significant negative and positive deviations from observed data, respectively. The effect sizes inconsistently differed between datasets (Figures S4 and S7).

Differences in null network dietary compositions

The different prey availability datasets produced compositionally distinct null diet compositions (i.e., the interaction partners of the spiders generated by null models). The mean Euclidean coordinates of the null diet compositions generated via NMDS differed in their distance from the mean observed dietary composition (Figure 4; Figure S8), with FOO combined (Euclidean distance = 0.062), abundance (0.080) equal prey abundance (0.082), proportional combined (0.143) and activity density (0.269) being progressively further from the observed diets, respectively. All but abundance followed a relatively linear progression of difference from the observed diets with respect to NMDS axes.

Differences in null network structure

The properties of null networks generated using the different prey availability datasets differed substantially (Figures 5 and 6; Figures S9 and S10). Nestedness and linkage density of the abundance, activity density and proportional combined null networks more closely resembled that of the network directly generated from dietary data (Figure 6; Figure S10), as did the degree of prey nodes in many instances (Figures S5 and S11).
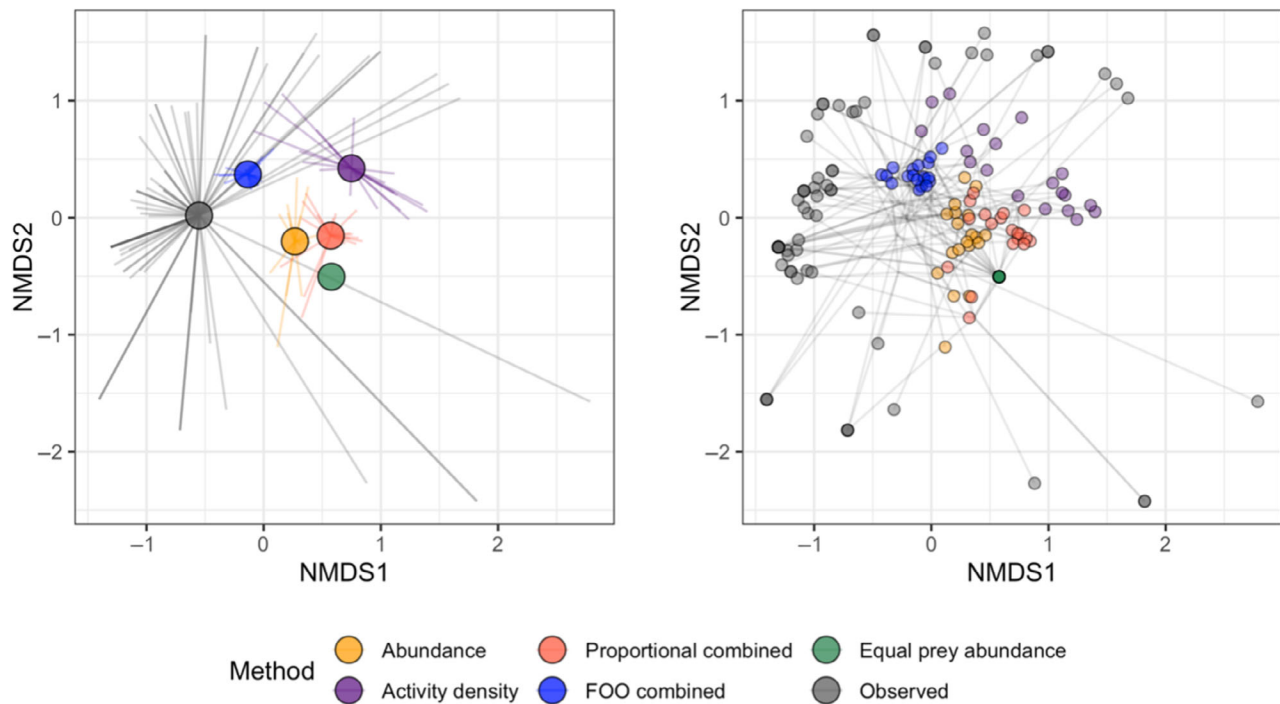
FIGURE 2 Spider plot derived from non-metric multi-dimensional scaling (NMDS) of spider diets ('observed') and prey communities derived from different sampling methods. In the left plot, the terminal end of each line represents a prey community or diet of a spider, joined by the centroids of diets from each data source (larger nodes; mean coordinates in that group), with distance between them indicating their dissimilarity (i.e., proximate points are similar, distant points are dissimilar). In the right plot, each point represents a prey community or diet of a spider, with lines joining points representing the same sample and those lines meeting at the mean coordinates for that sample. FOO, frequency of occurrence.

DISCUSSION

The choice of prey sampling method has implications for the interpretation of predator selectivity, diet composition and network structure when using null network models, but the precise nature of these effects is complex. The implications of prey sampling method are dependent on the specific questions being addressed, with the methods for estimating prey availability having different effects upon the frequency of interactions, diet composition and network structures of null networks and, consequently, the inferred selectivity of predators. Increasing the completeness of datasets by merging sampling methods could theoretically better simulate trophic interactions by increasing the accuracy of prey availability data. We show, however, that combining methods is not straightforward and may generate null diet compositions or network structures that differ more from those generated from directly detected interactions, depending on the method of merging. Even though relying on single biased methods can misrepresent true prey availability, we lack methods for reliably improving these estimates. Because all methods for estimating prey availability have their biases, we must always treat estimated predator selectivity and corresponding network metrics with caution. The optimal representation of prey availability in these contexts depends on the hypotheses being tested and, most crucially, the ecology of the system being studied.

Considerations when choosing sampling methods for null network choice analysis

We have shown that different sampling methods affect the outcomes of predator selectivity analyses. This conclusion is intuitive since altering the data from which null models are generated will naturally change the structure of the null networks and the identities of the resource nodes within. The effects are, however, far more nuanced and differ depending on the characteristics being assessed. Some sampling methods generated null networks with structural properties more like the network constructed using interactions directly detected by metabarcoding, and other sampling methods generated null diet compositions with prey identities and frequencies more like the directly detected network. This finding predicates that optimal sampling method choice is not only system-dependent but also contingent upon the specific hypotheses being tested.

Sampling methods have well-documented and characterised taxonomic and functional biases which can vary greatly, as exemplified by the methods used in this study. Although suction sampling can capture many of the smaller near-ground prey commonly exploited by spiders (Cooper & Whitmore, 1990; Harper & Guynn, 1998), it can disproportionately represent thrips, spiders, true bugs, flies and wasps (Doxon et al., 2011; Zentane et al., 2016). The 'peripheral suction effect', whereby some taxa are predisposed to being collected from beyond the sampling area, exacerbates these taxonomic biases,
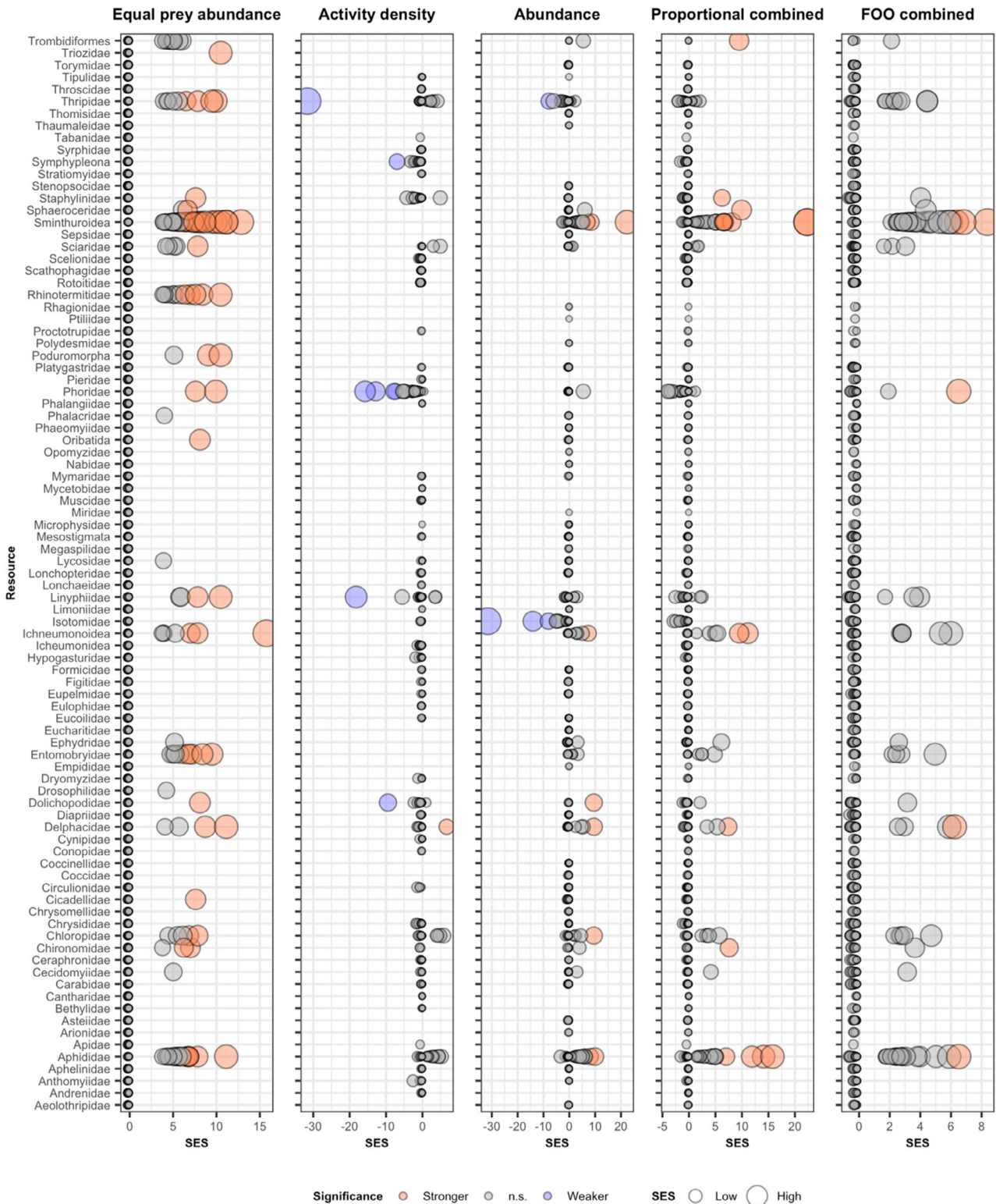
FIGURE 3 Prey choice standardised effect sizes (SESs) for each prey taxon and individual consumers for the five different null models. Larger points reflect larger deviations of SESs from zero (i.e., stronger or weaker selection). Red, grey and blue points denote significantly more frequent (stronger selection), non-significant (n.s.) and significantly less frequent (weaker selection) interactions compared with the null model ($p < 0.05$). Absent points are those for which data were not available. FOO, frequency of occurrence.

although it can be mitigated by enclosing the sampling area (Cherrill, 2015). Other factors, such as time of day, weather conditions and time spent suction sampling, may also influence the outcomes (Bell et al., 2000; Brook et al., 2008). Similarly, sticky traps elicit different biases based on the colour used, which determines attraction of different taxa (Böckmann & Meyhöfer, 2017; Chittka & Menzel, 1992;
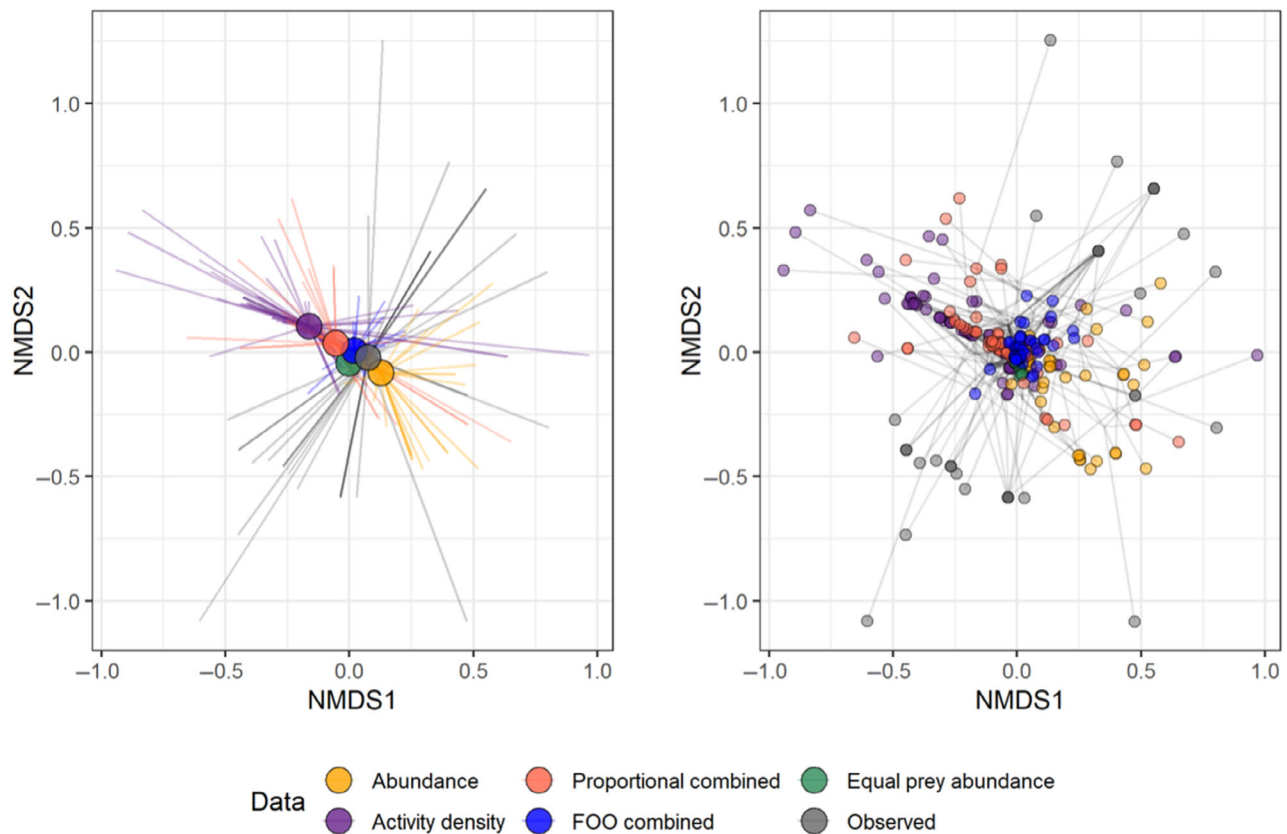
FIGURE 4 Spider plot derived from non-metric multi-dimensional scaling (NMDS) of observed and null diet compositions of spiders. In the left plot, the terminal end of each line represents the diet composition of a spider (whether observed or null model generated), joined by the centroids of diets from each data source (larger nodes; mean coordinates in that group), with distance between them indicating their dissimilarity (i.e., proximate points are similar, distant points are dissimilar). In the right plot, each point represents the diet composition of a spider, with lines joining points representing the same sample and those lines meeting at the mean coordinates for that sample. FOO, frequency of occurrence.

Döring et al., 2012; Hoback et al., 1999; Sétamou et al., 2014). Ultimately, all sampling methods are biased, but when representing the availability of prey to a predator, the sampling method should ideally emulate, as closely as possible, the foraging behaviour and prey capture opportunities of the predator (Table 1).

In a hypothetical study concerning the prey preferences of trapdoor spiders (Araneae: Halonoproctidae), the researchers involved should consider the foraging mode of the predator to select an appropriate prey survey method. Given that trapdoor spiders are sit-and-wait predators, it is likely that activity density would better reflect the prey available to the spiders. Since the trapdoors of these spiders, from which they ambush prey, resemble pits in the ground, pitfall traps may be an intuitive choice of sampling method. In fact, one study highlights that pitfall trap abundances correlate with trapdoor spider activity (Bradley, 1996), suggesting that this would be an appropriate choice. Through reference to existing literature and natural history records (e.g., Bradley, 1996; Coyle & Icenogle, 1994; Gupta et al., 2015), the researchers could identify the likely prey of trapdoor spiders and whether this method will collect them. Further considerations like the diameter of the trap opening (Table 1) could also be adjusted according to existing data within the system to design the most realistic representation of prey capture. In the case of the

trapdoor spider, prey vibrational cues instigate predation (Nakamura et al., 2022); thus, the prey collected during surveys must also be considered from this physiological perspective (i.e., whether they would produce a viable vibrational cue to trigger their predation) unless this is the mechanism of choice being explored via null modelling.

Datasets generated by combining data from both sampling methods were more complete and included a broader spectrum of available prey, leading to fewer false negatives in null models (i.e., prey that were detected in the guts of spiders, but not found in the prey availability data). Due to the greater imbalance in prey counts in these datasets, however, interaction frequencies were distributed across a greater number of prey, leading to less realistic network topologies and interaction weights. These 'false positive' null interactions which did not occur in the directly detected interaction data could be mitigated by restricting resource availability data to only those taxa with which consumers were found to interact, but this would obviate any investigation of the mechanisms potentially driving the exclusion of these taxa from consumer interactions. Merging sampling data from different survey methods is thus a complicated solution that may introduce additional biases; although it overcomes individual methodological biases and increases sampling completeness, it may inflate deviation of directly detected network properties
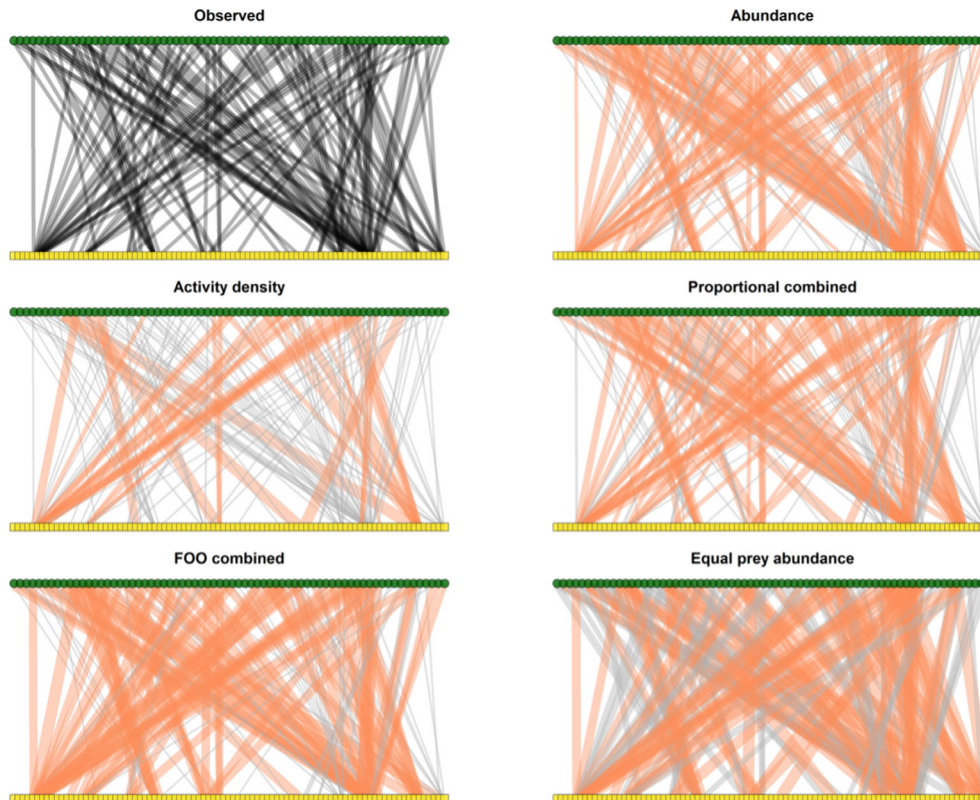
FIGURE 5 Network generated solely from dietary data ('observed') and null networks produced from the different prey availability datasets. Link weights represent the number of observed interactions for the 'observed' network, and otherwise the frequency of interactions expected based on prey availability according to null models. Red links represent those for which observed interaction frequencies significantly exceeded those expected from null models. No interactions are plotted which were significantly less frequent than expected. FOO, frequency of occurrence.
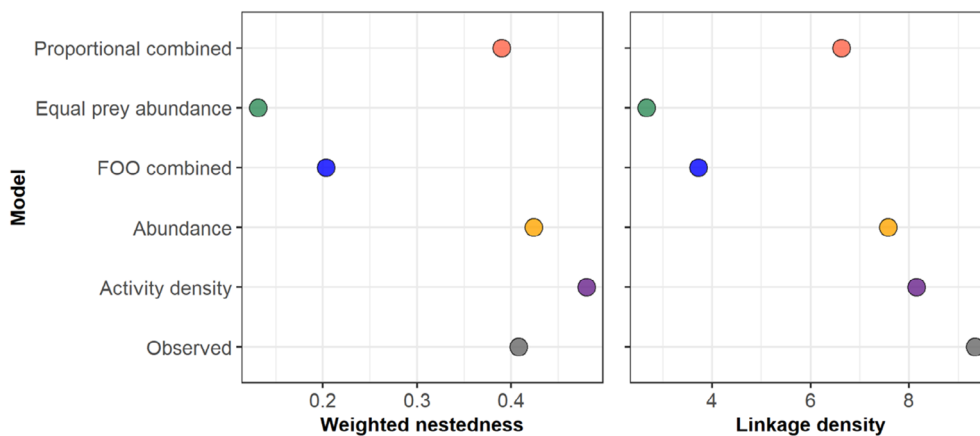

FIGURE 6 Weighted nestedness and linkage density of the 'observed' network and the null networks generated from the different prey availability datasets. FOO, frequency of occurrence.

from null models. This could equally be due to an unaccounted-for ecological phenomenon though. Inundating null models with available prey will nevertheless distribute interactions across a broader range of resources, potentially reducing linkage density, modularity and nestedness. Poorly informed data-saturated null networks in which any consumer–resource interaction is permitted will thus poorly represent a baseline against which to compare real-world trophic

interactions in which consumers are time, energy and resource limited and will thus be more selective. Care must be taken to ensure that only plausible interactions are represented by the resultant null models, or those pertaining to the hypothesis being tested.

It is important to consider how survey data align with the observed data with which they are being compared. Directly detected or observed network data are often subject to biases and limitations

with similar implications for differences between observed and null networks. For example, the time scales over which observations and prey survey data are collected may mismatch if the prey with which predators interact are no longer present at the time of surveying (e.g., the long period of detection of DNA data compared with active prey abundance surveys). This problem is particularly noteworthy for prey present at different times in the diel cycle, which may be under-represented by abundance data. It is thus crucial to consider how measures of observed interactions and prey availability data align during experimental design.

Dietary metabarcoding data, increasingly used in null models assessing resource choice (Cuff et al., 2021; Cuff, Tercel, et al., 2022; Davies et al., 2022; Evens et al., 2020; Gajski et al., 2023; Moorhouse-Gann et al., 2022; Verschut et al., 2019; Villsen et al., 2022), also present several distinct considerations. Alongside lacking the context of prey life stage, sex and other contextual information that such null models could otherwise include to assess how prey traits affect foraging, quantification of metabarcoding data is a vital concern. Trophic interactions represented as binary presence–absence data lack realistic interaction weights, whereas relative read abundances generated by metabarcoding present quantities, but they are often taxonomically biased (Deagle et al., 2019; Lamb et al., 2019). Neglecting quantitative data will, at least at the individual level, mismatch the weighted interactions of null models, whereas introducing biased data could poorly reflect realistic foraging. By running these analyses with both data types (Figures S6–S11), we have shown that the overall differences between observed data and corresponding null networks were relatively consistent regardless of the observed data used. The key differences were that relative read abundances generated more significant deviations between null and observed data but that these were largely just stronger effect sizes for the same taxa highlighted by presence–absence-based analyses. This implies that the core findings apply regardless of the observed data used, although additional investigation across different contexts (i.e., data types, study organisms and systems) is required to further demonstrate this. Another important consideration for the use of relative read abundances is how these might relate to biomass rather than straightforward abundance of resources. Although the link between biomass and read counts is inconsistent (Lamb et al., 2019), future studies could explore the potential of null networks based on biomass as a more accurate reflection of read counts.

Dietary metabarcoding data are also subject to false positives (Drake et al., 2022) and false negatives (Littleford-Colquhoun et al., 2022). This is particularly insidious for omnivorous consumers (Tercel et al., 2021) but can be exacerbated by amplification of the DNA of the focal predator itself (Cuff, Kitson, et al., 2023). The presence of false positives/negatives undoubtedly alters the congruence of observed data with null networks and the inferred strength of resource preferences; thus, a careful approach that limits data loss whilst preserving data integrity is required (Littleford-Colquhoun et al., 2022). By merging networks constructed with DNA and direct observation data (i.e., visually recorded interaction data), some of the biases and limitations of both methods can be overcome (Cuff, Windsor, et al., 2022), but how this affects congruence of dietary and prey availability data may vary greatly. Ultimately, the methodological biases imposed on the prey availability and observed interaction data should align as much as possible to limit the impact of dataset mismatches on ecological outcomes.

Example of the ecological information gained from null model comparisons

To assess the relevance of different measures of prey availability, it is important to consider what these measures should achieve. The aim of null modelling is to investigate specific mechanisms as drivers of ecological patterns through comparison of observed data with data generated according to a specific null hypothesis (Gotelli, 2001; Gotelli & Graves, 1996). In the context of prey choice, the models used in this study are purposed to identify interactions that occur more or less frequently than would be expected if predators randomly sampled from the community of prey available to them (Vaughan et al., 2018). The objective of sampling is therefore to represent the prey available to each predator, for which the ecology of each forager is a vital consideration.

The predators in this study have variable foraging behaviour and ecology, best represented by the two Linyphiidae subfamilies, Linyphiinae and Erigoninae. One *Pardosa* (Lycosidae) spider was also included in the analyses which, as an active ground-hunting spider, might forage in a manner more closely resembling the prey collected by pitfall trapping, as in the trapdoor spider example above, but the results from this single spider are unlikely to influence the overall results of the study. The Linyphiidae spiders, however, use webs to forage, the position and size of which differs even within families (Harwood et al., 2001). Linyphiinae spiders build larger sheet webs a few centimetres above the ground, whereas Erigoninae spiders build smaller webs closer to the ground, which they leave regularly to forage (Sunderland et al., 1986), lending to separation of their trophic niches (Harwood et al., 2003). Given the difference between active and passive foraging, it might have been predicted that the interactions of these spider groups would differ in their similarity to null networks generated using abundance and activity density data, with Erigoninae resembling abundance and Linyphiinae resembling activity density.

The interactions of all spider groups more closely resembled abundance than activity density, and the interactions of Erigoninae, although more like the abundance-based null network, were more similar to the activity density null diet compositions than Linyphiinae (Figure S12). The prey availability data appeared to influence the results much more than the subfamily of spider though. This difference may nonetheless indicate that neither method nor their combinations perfectly reflected the availability of prey to the spiders (particularly Linyphiinae). Other methods were trialled within this study, including longer ground-pinned sticky traps and acetate sheets coated with ecological glue (Oecotak adhesive) which approximately matched the webs of each spider in size and position, both of which

captured too few intact prey to generate null networks and were thus abandoned. Alternatively, the low congruence of Linyphiinae interactions with null networks may suggest that Linyphiinae simply have stronger density-independent preferences than Erigoninae (see Cuff, Tercel, et al., 2022 for an in-depth taxonomic comparison of preferences), as these methods are intended to determine. The latter case would highlight that, although null networks need to represent prey availability accurately, similarity between null networks and networks derived from directly detected or observed interactions does not absolutely equate to the suitability of the survey method given that density-independent foraging is commonplace. Perfectly simulating the diet of the consumer, although useful for predictive applications (Cuff, Windsor, et al., 2023), is not usually the aim of null networks used in this context. Instead, null networks should represent interactions that are physiologically, spatially and temporally accessible to the consumer. As such, observed interactions can inform which sampling method to deploy in specific contexts, but they cannot govern these decisions without introducing dogma.

The equal prey abundance null diet compositions were more similar to the observed diet than those based on activity density or proportionally combined data, indicating that spiders achieved relatively even interactions across the diversity of prey available, arguably irrespective of prey activity density. This pattern suggests that prey community diversity is a greater driver of interaction diversity than prey activity densities for these spiders, which has important implications not only for optimal foraging theory, since it suggests foraging within optimal patches rather than for optimal prey (MacArthur & Pianka, 1966), but also for the 'iDNA' (invertebrate-derived DNA) monitoring concept (Cutajar & Rowley, 2020; Drinkwater et al., 2021) since it highlights the potential validity of using invertebrates as samplers of DNA for diversity assessment.

The results of this study are ultimately limited by sample size (70 spiders across 19 locations), division of samples across different spiders with potentially different foraging behaviours, and taxonomic resolution (only family level identification of prey availability data and consequent restriction of dietary metabarcoding data to the same level). Finer resolution of identification may elucidate more intricate patterns of prey choice, which may be further influenced by the biases of specific sampling methods. The approaches used here, in which individual consumer interaction data are compared against hundreds of simulations (Vaughan et al., 2018), can generate results from relatively few individuals, but the accuracy of those results will always be greater with more data. Although we contend that the sample size here has allowed us to draw sound conclusions, larger studies of this kind are needed to validate the findings, ideally across a greater range of taxa to ensure the findings are generalisable. In doing so, this can provide explicit guidance for the design of robust prey choice analyses.

Conclusions

We have demonstrated the substantial influence of different measures of prey availability on ecological interpretation of null network-based prey choice analyses. We also have shown that the choice of survey method should be dependent on the hypotheses, study system and predator foraging mode being investigated. Merging different data types increases data completeness and can produce null networks similar to those constructed from directly detected interactions, but the networks, interactions and null diet compositions generated were also similar to those generated with data from a single measure of prey availability. During experimental design, researchers must carefully consider not only the foraging mode and ecology of the focal predator to represent their available prey but also the breadth of prey that they are likely to interact with and the concomitant network structures. Increasingly data-rich approaches to null network analysis have further implications for methodological choices. Integrating data such as environmental context (Cuff, Windsor, et al., 2023) and consumer functional trait data (Ibanez, 2012) can refine or test predictions regarding consumer choices and could greatly expand the utility and benefit of null modelling approaches, but robust resource availability data must be sought in all such applications.

AUTHOR CONTRIBUTIONS

Jordan P. Cuff: Conceptualization; data curation; formal analysis; visualization; writing – original draft; methodology; investigation; project administration; writing – review and editing; funding acquisition. **Maximillian P. T. G. Tercel:** Conceptualization; investigation; writing – review and editing; formal analysis. **Fredric M. Windsor:** Formal analysis; investigation; writing – review and editing. **Ben S. J. Hawthorne:** Investigation; writing – review and editing. **Peter A. Hambäck:** Investigation; writing – review and editing. **James R. Bell:** Writing – review and editing; project administration; supervision; investigation; funding acquisition. **William O. C. Symondson:** Funding acquisition; investigation; supervision; project administration; writing – review and editing. **Ian P. Vaughan:** Funding acquisition; investigation; supervision; project administration; writing – review and editing; conceptualization; formal analysis.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

All data and code are openly available via Zenodo: https://doi.org/10.5281/ZENODO.7908186 (Cuff, Tercel, et al., 2023).

ORCID

Jordan P. Cuff https://orcid.org/0000-0002-0198-4940

REFERENCES

Agustí, N., Shayler, S.P., Harwood, J.D., Vaughan, I.P., Sunderland, K.D. & Symondson, W.O.C. (2003) Collembola as alternative prey sustaining spiders in arable ecosystems: prey detection within predators using molecular markers. *Molecular Ecology*, 12(12), 3467–3475. Available from: https://doi.org/10.1046/j.1365-294X.2003.02014.x

Allesina, S., Alonso, D. & Pascual, M. (2008) A general model for food web structure. *Science*, 320(5876), 658–661. Available from: https://doi.org/10.1126/science.1156269

Bell, J.R., Wheater, C.P., Henderson, R. & Cullen, W.R. (2000) Testing the efficiency of suction samplers (G-vacs) on spiders: the effect of increasing nozzle size and suction time. In: Toft, S. & Scharff, N. (Eds.) *European arachnology*. Aarhus, Denmark: Aarhus University Press, pp. 285–290.

Blüthgen, N., Fründ, J., Vázquez, D.P. & Menzel, F. (2008) What do interaction network metrics tell us about specialization and biological traits. *Ecology*, 89(12), 3387–3399. Available from: https://doi.org/10.1890/07-2121.1

Böckmann, E. & Meyhöfer, R. (2017) Sticky trap monitoring of a pest-predator system in glasshouse tomato crops: are available trap colours sufficient? *Journal of Applied Entomology*, 141(5), 339–351. Available from: https://doi.org/10.1111/jen.12338

Bradley, R.A. (1996) Foraging activity and burrow distribution in the Sydney brown trapdoor spider (*Misgolas rapax* Karsch: Idiopidae). *The Journal of Arachnology*, 24(1), 58–67.

Briatte, F. (2021) *ggnetwork: Geometries to plot networks with "ggplot2"* (R package version 0.5.10).

Brook, A.J., Woodcock, B.A., Sinka, M. & Vanbergen, A.J. (2008) Experimental verification of suction sampler capture efficiency in grasslands of differing vegetation height and structure. *Journal of Applied Ecology*, 45(5), 1357–1363. Available from: https://doi.org/10.1111/j.1365-2664.2008.01530.x

Chao, A., Gotelli, N., Hsieh, T., Sander, E., Ma, K., Colwell, R. et al. (2014) Rarefaction and extrapolation with hill numbers: a framework for sampling and estimation in species diversity studies. *Ecological Monographs*, 84(1), 45–67.

Cherrill, A. (2015) Suction sampling of grassland invertebrates using the G-vac: quantifying and avoiding peripheral suction effects. *European Journal of Entomology*, 112(3), 520–524. Available from: https://doi.org/10.14411/eje.2015.058

Chittka, L. & Menzel, R. (1992) The evolutionary adaptation of flower colours and the insect pollinators' colour vision. *Journal of Comparative Physiology A*, 171(2), 171–181. Available from: https://doi.org/10.1007/BF00188925

Cooper, R.J. & Whitmore, R.C. (1990) Arthropod sampling methods in ornithology. *Studies in Avian Biology*, 13, 29–37.

Coyle, F.A. & Icenogle, W.R. (1994) Natural history of the Californian trapdoor spider genus *Aliatypus* (Araneae, Antrodiaetidae). *The Journal of Arachnology*, 22(3), 225–255.

Csardi, G. & Nepusz, T. (2006) The igraph software package for complex network research. *InterJournal Complex Systems*, 1695, 1–9.

Cuff, J.P., Drake, L.E., Tercel, M.P.T.G., Stockdale, J.E., Orozco-terWengel, P., Bell, J.R. et al. (2021) Money spider dietary choice in pre- and post-harvest cereal crops using metabarcoding. *Ecological Entomology*, 46(2), 249–261. Available from: https://doi.org/10.1111/een.12957

Cuff, J.P., Evans, S.A., Porteous, I.A., Quiñonez, J. & Evans, D.M. (2022) Candy-striped spider leaf and habitat preferences for egg deposition. *Agricultural and Forest Entomology*, 24(3), 422–431. Available from: https://doi.org/10.1111/afe.12504

Cuff, J.P., Kitson, J.J.N., Hemprich-Bennett, D., Tercel, M.P.T.G., Browett, S.S. & Evans, D.M. (2023) The predator problem and PCR primers in molecular dietary analysis: swamped or silenced; depth or breadth? *Molecular Ecology Resources*, 23(1), 41–51. Available from: https://doi.org/10.1111/1755-0998.13705

Cuff, J.P., Tercel, M.P.T.G., Drake, L.E., Vaughan, I.P., Bell, J.R., Orozco-terWengel, P. et al. (2022) Density-independent prey choice, taxonomy, life history and web characteristics determine the diet and biocontrol potential of spiders (Linyphiidae and Lycosidae) in cereal crops. *Environmental DNA*, 4(3), 549–564.

Cuff, J.P., Tercel, M.P.T.G., Windsor, F.M., Hawthorne, B.S.J., Hambäck, P.A., Bell, J.R. et al. (2023) *Sources of prey availability data alter interpretation of outputs from prey choice null networks* [dataset]. Zenodo. https://doi.org/10.5281/ZENODO.7908186

Cuff, J.P., Windsor, F.M., Tercel, M.P.T.G., Bell, J.R., Symondson, W.O.C. & Vaughan, I.P. (2023) Temporal variation in spider trophic interactions is explained by the influence of weather on prey communities, web building and prey choice. *Ecography*, 2023, e06737. Available from: https://doi.org/10.1111/ecog.06737

Cuff, J.P., Windsor, F.M., Tercel, M.P.T.G., Kitson, J.J.N. & Evans, D.M. (2022) Overcoming the pitfalls of merging dietary metabarcoding into ecological networks. *Methods in Ecology and Evolution*, 13(3), 545–559. Available from: https://doi.org/10.1111/2041-210X.13796

Cutajar, T.P. & Rowley, J.J.L. (2020) Surveying frogs from the bellies of their parasites: invertebrate-derived DNA as a novel survey method for frogs. *Global Ecology and Conservation*, 22, e00978. Available from: https://doi.org/10.1016/j.gecco.2020.e00978

Davies, S.R., Vaughan, I.P., Thomas, R.J., Drake, L.E., Marchbank, A. & Symondson, W.O.C. (2022) Seasonal and ontological variation in diet and age-related differences in prey choice, by an insectivorous songbird. *Ecology and Evolution*, 12(8), e9180. Available from: https://doi.org/10.1002/ece3.9180

Deagle, B.E., Thomas, A.C., McInnes, J.C., Clarke, L.J., Vesterinen, E.J., Clare, E.L. et al. (2019) Counting with DNA in metabarcoding studies: how should we convert sequence reads to dietary data? *Molecular Ecology*, 28(2), 391–406. Available from: https://doi.org/10.1111/mec.14734

Döring, T.F., Skellern, M., Watts, N. & Cook, S.M. (2012) Colour choice behaviour in the pollen beetle *Meligethes aeneus* (Coleoptera: Nitidulidae). *Physiological Entomology*, 37(4), 360–378. Available from: https://doi.org/10.1111/j.1365-3032.2012.00850.x

Dormann, C.F., Gruber, B. & Fruend, J. (2008) Introducing the bipartite package: Analysing ecological networks. *R News*, 8(2), 8–11.

Doxon, E.D., Davis, C.A. & Fuhlendorf, S.D. (2011) Comparison of two methods for sampling invertebrates: vacuum and sweep-net sampling: methods of sampling invertebrates. *Journal of Field Ornithology*, 82(1), 60–67. Available from: https://doi.org/10.1111/j.1557-9263.2010.00308.x

Drake, L.E., Cuff, J.P., Young, R.E., Marchbank, A., Chadwick, E.A. & Symondson, W.O.C. (2022) An assessment of minimum sequence copy thresholds for identifying and reducing the prevalence of artefacts in dietary metabarcoding data. *Methods in Ecology and Evolution*, 13(3), 694–710. Available from: https://doi.org/10.1111/2041-210X.13780

Drinkwater, R., Williamson, J., Clare, E.L., Chung, A.Y.C., Rossiter, S.J. & Slade, E. (2021) Dung beetles as samplers of mammals in Malaysian Borneo—a test of high throughput metabarcoding of iDNA. *PeerJ*, 9, e11897. Available from: https://doi.org/10.7717/peerj.11897

Evens, R., Conway, G., Franklin, K., Henderson, I., Stockdale, J., Beenaerts, N. et al. (2020) DNA diet profiles with high-resolution animal tracking data reveal levels of prey selection relative to habitat choice in a crepuscular insectivorous bird. *Ecology and Evolution*, 10(23), 13044–13056. Available from: https://doi.org/10.1002/ece3.6893

Gajski, D., Mifková, T., Košulič, O., Michálek, O., Serbina, L.Š., Michalko, R. et al. (2023) Brace yourselves, winter is coming: the winter activity, natural diet, and prey preference of winter-active spiders on pear trees. *Journal of Pest Science*, 97, 113–126. Available from: https://doi.org/10.1007/s10340-023-01609-5

PREY CHOICE NULL NETWORK DATA SOURCES

Ecological
Entomology

Royal
Entomological
Society

13

Gómez-Martínez, C., González-Estévez, M.A., Cursach, J. & Lázaro, A. (2022) Pollinator richness, pollination networks, and diet adjustment along local and landscape gradients of resource diversity. *Ecological Applications*, 32(6), e2634. Available from: https://doi.org/10.1002/eap.2634

Gotelli, N.J. (2001) Research frontiers in null model analysis: null model analysis. *Global Ecology and Biogeography*, 10(4), 337–343. Available from: https://doi.org/10.1046/j.1466-822X.2001.00249.x

Gotelli, N.J. & Graves, G.R. (1996) *Null models in ecology*. Washington, D.C, Smithsonian Institution Press.

Grimm, V. & Berger, U. (2016) Structural realism, emergence, and predictions in next-generation ecological modelling: synthesis from a special issue. *Ecological Modelling*, 326, 177–187. Available from: https://doi.org/10.1016/j.ecolmodel.2016.01.001

Gupta, N., Das, S.K. & Siliwal, M. (2015) Natural history of the trapdoor spider Idiops joida Gupta et al 2013 (Araneae: Idiopidae) from the Western Ghats in India. *Journal of Asia-Pacific Biodiversity*, 8(1), 38–42. Available from: https://doi.org/10.1016/j.japb.2015.01.004

Harper, C.A. & Guynn, D.C. (1998) A terrestrial vacuum sampler for macro-invertebrates. *Wildlife Society Bulletin*, 26, 302–306.

Harwood, J.D., Sunderland, K.D. & Symondson, W.O.C. (2001) Living where the food is: web location by linyphiid spiders in relation to prey availability in winter wheat: web location by linyphiid spiders. *Journal of Applied Ecology*, 38(1), 88–99. Available from: https://doi.org/10.1046/j.1365-2664.2001.00572.x

Harwood, J.D., Sunderland, K.D. & Symondson, W.O.C. (2003) Web-location by linyphiid spiders: prey-specific aggregation and foraging strategies. *Journal of Animal Ecology*, 72(5), 745–756. Available from: https://doi.org/10.1046/j.1365-2656.2003.00746.x

Hoback, W.W., Svatos, T.M., Spomer, S.M. & Higley, L.G. (1999) Trap color and placement affects estimates of insect family-level abundance and diversity in a Nebraska salt marsh. *Entomologia Experimentalis et Applicata*, 91(3), 393–402. Available from: https://doi.org/10.1046/j.1570-7458.1999.00507.x

Hsieh, T.C., Ma, K.H. & Chao, A. (2016) INEXT: an R package for interpolation and extrapolation of species diversity (hill numbers). *Methods in Ecology and Evolution*, 7(12), 1451–1456.

Ibanez, S. (2012) Optimizing size thresholds in a plant–pollinator interaction web: towards a mechanistic understanding of ecological networks. *Oecologia*, 170(1), 233–242. Available from: https://doi.org/10.1007/s00442-012-2290-3

Lamb, P.D., Hunter, E., Pinnegar, J.K., Creer, S., Davies, R.G. & Taylor, M.I. (2019) How quantitative is metabarcoding: a meta-analytical approach. *Molecular Ecology*, 28(2), 420–430. Available from: https://doi.org/10.1111/mec.14920

Littleford-Colquhoun, B.L., Freeman, P.T., Sackett, V.I., Tulloss, C.V., McGarvey, L.M., Geremia, C. et al. (2022) The precautionary principle and dietary DNA metabarcoding: commonly used abundance thresholds change ecological interpretation. *Molecular Ecology*, 31(6), 1615–1626. Available from: https://doi.org/10.1111/mec.16352

MacArthur, R.H. & Pianka, E.R. (1966) On optimal use of a patchy environment. *The American Naturalist*, 100(916), 603–609.

Méndez-Castro, F.E., Mendieta-Leiva, G., Rao, D. & Bader, M.Y. (2020) Island-biogeographic patterns of spider communities on epiphytes depend on differential space use among functional groups. *Journal of Biogeography*, 47(6), 1322–1332. Available from: https://doi.org/10.1111/jbi.13812

Moorhouse-Gann, R.J., Vaughan, I.P., Cole, N.C., Goder, M., Tatayah, V., Jones, C.G. et al. (2022) Impacts of herbivory by ecological replacements on an Island ecosystem. *Journal of Applied Ecology*, 59(9), 2245–2261. Available from: https://doi.org/10.1111/1365-2664.14096

Nakamura, S., Mukai, H. & Tokuda, M. (2022) A trapdoor spider, *Latouchia typica* (Araneae: Halonoproctidae), uses vibrational cues as a trigger for predatory behavior. *Invertebrate Biology*, 141(4), e12388. Available from: https://doi.org/10.1111/ivb.12388

Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B. et al. (2016) *vegan: Community Ecology Package* (R package version 2.4-1).

R Core Team. (2021) *R: a language and environment for statistical computing (4.1.0)*. R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org/

Ramirez, I.E., Causton, C.E., Gutierrez, G.A., Mosquera, D.A., Piedrahita, P. & Heimpel, G.E. (2022) Specificity within bird–parasite–parasitoid food webs: a novel approach for evaluating potential biological control agents of the avian vampire fly. *Journal of Applied Ecology*, 59(8), 2189–2198. Available from: https://doi.org/10.1111/1365-2664.14235

Roswell, M., Dushoff, J. & Winfree, R. (2021) A conceptual guide to measuring species diversity. *Oikos*, 130, 321–338.

Sétamou, M., Sanchez, A., Saldaña, R.R., Patt, J.M. & Summy, R. (2014) Visual responses of adult Asian citrus psyllid (Hemiptera: Liviidae) to colored sticky traps on citrus trees. *Journal of Insect Behavior*, 27(4), 540–553. Available from: https://doi.org/10.1007/s10905-014-9448-2

Sunderland, K.D., Fraser, A.M. & Dixon, A.F.G. (1986) Distribution of linyphiid spiders in relation to capture of prey in cereal fields. *Pedobiologia*, 29, 367–375.

Tercel, M.P.T.G., Symondson, W.O.C. & Cuff, J.P. (2021) The problem of omnivory: a synthesis on omnivory and DNA metabarcoding. *Molecular Ecology*, 30(10), 2199–2206. Available from: https://doi.org/10.1111/mec.15903

Vaughan, I.P., Gotelli, N.J., Memmott, J., Pearson, C.E., Woodward, G. & Symondson, W.O.C. (2018) Econullnetr: an R package using null models to analyse the structure of ecological networks and identify resource selection. *Methods in Ecology and Evolution*, 9(3), 728–733. Available from: https://doi.org/10.1111/2041-210X.12907

Vázquez, D.P. & Aizen, M.A. (2003) Null model analyses of specialization in plant–pollinator interactions. *Ecology*, 84(9), 2493–2501. Available from: https://doi.org/10.1890/02-0587

Verschut, V., Strandmark, A., Esparza-Salas, R. & Hambäck, P.A. (2019) Seasonally varying marine influences on the coastal ecosystem detected through molecular gut analysis. *Molecular Ecology*, 28(2), 307–317. Available from: https://doi.org/10.1111/mec.14830

Villsen, K., Corse, E., Archambaud-Suard, G., Héran, K., Meglécz, E., Ereskovsky, A.V. et al. (2022) Diet Metabarcoding reveals extensive dietary overlap between two benthic stream fishes (*Zingel asper* and *Cottus gobio*) and provides insights into their coexistence. *Diversity*, 14(5), 412. Available from: https://doi.org/10.3390/d14050412

Wickham, H. (2016) *ggplot2: elegant graphics for data analysis (3.3.0)*. New York, Springer-Verlag.

Zentane, E., Quenu, H., Graham, R.I. & Cherrill, A. (2016) Suction samplers for grassland invertebrates: comparison of numbers caught using Vortis™ and G-vac devices. *Insect Conservation and Diversity*, 9(5), 470–474. Available from: https://doi.org/10.1111/icad.12185

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Data S1: Supplementary Information 1: Molecular analysis and bioinformatics.

Figure S1. Diversity detected during suction sampling, and the associated sample coverage. Red lines with terminal circles, green with terminal triangles and blue with terminal squares denote Hill-species richness, Hill–Shannon and Hill–Simpson, respectively. Solid lines

represent observed diversity, and dashed lines extrapolated diversity. Light zones surrounding lines denote 95% confidence intervals.

Figure S2. Diversity detected during sticky trapping and the associated sample coverage. Red lines with terminal circles, green with terminal triangles and blue with terminal squares denote Hill-species richness, Hill-Shannon and Hill-Simpson, respectively. Solid lines represent observed diversity, and dashed lines extrapolated diversity. Light zones surrounding lines denote 95% confidence intervals.

Figure S3. Diversity detected by the combination of suction sampling and sticky trapping, and the associated sample coverage. Red lines with terminal circles, green with terminal triangles and blue with terminal squares denote Hill-species richness, Hill–Simpson, respectively. Solid lines represent observed diversity, and dashed lines extrapolated diversity. Light zones surrounding lines denote 95% confidence intervals.

Figure S4. Standardised effect sizes from the comparison of null models against observed interactions. Mean SES values for each prey taxon across the individual consumers were used. Each colour denotes SESs from a different prey choice model. Links are drawn between points representing the same prey taxon.

Figure S5. Degree of each node within the observed (presence–absence) and the null networks generated from the different prey availability datasets.

Figure S6. Relative read abundance-based prey choice standardised effect sizes (SESs) for each prey taxon and individual consumers for the five different null models. Larger points reflect larger deviations of SESs from zero (i.e., stronger or weaker selection). Red, white and blue points denote significantly more frequent (stronger selection), non-significant and significantly less frequent (weaker selection) interactions compared with the null model. Absent points are those for which data were not available. The overall patterns were consistent between presence–absence and relative read abundance data, but the latter identifies more significant deviations from the null model and stronger SESs generally (Figure 3).

Figure S7. Relative read abundance-based standardised effect sizes (SESs) exported from the comparison of null models against observed interactions. Mean SES values for each prey taxon across the individual consumers were used. Each colour denotes SESs from a different prey choice model. Links are drawn between points representing the same prey taxon. The data very closely resemble those of the presence–absence dietary models, but SESs were generally more positive in the relative read abundance data (Figure S4).

Figure S8. Spider plot derived from non-metric multi-dimensional scaling of observed (relative read abundance-based) and null model-expected spider diets. Colours denote the data source (grey = abundance; other colours are null-model-predicted diets based on the different prey availability datasets). Axes represent a two-dimensional variation in spider diet. In the left plot, the terminal end of each line represents the diet of a spider (whether observed or simulated), joined by the centroids of diets from each data source (larger nodes; mean coordinates in that group), with distance between them indicating their dissimilarity (i.e., proximate points are similar, distant points are dissimilar). In the right plot, each point represents the diet

of a spider, with lines joining points representing the same sample and those lines meeting at the mean coordinates for that sample. Mean Euclidean coordinates of the diets, differed in their distance from the mean observed dietary composition based on relative read abundance, with equal prey abundance (Euclidean distance = 0.052), FOO combined (0.055), abundance (0.071), proportional combined (0.115) and activity density (0.244) being progressively further from the observed data, respectively. This relationships between the models and the observed data were similar to those from the presence–absence dietary data, but the equal prey abundance model simulated diets more similar to the observed diets whereas its performance was intermediate from the presence–absence data (Figure 4).

Figure S9. Observed (relative read abundance-based) and null networks produced from the different prey availability datasets. Link weights represent the number of observed interactions for the observed network, and otherwise the frequency of interactions expected based on prey availability according to null models. Red links represent those for which observed interaction frequencies significantly exceeded those expected from null models. No interactions are plotted which were significantly less frequent than expected.

Figure S10. Weighted nestedness and linkage density of the observed (relative read abundance-based) network and the null networks generated from the different prey availability datasets. The observed relative read abundance network has a lower linkage density than the presence–absence network (Figure 6). The null models based on relative read abundance data differ in their relationship to the observed network compared with those of the presence–absence data. The activity density network is marginally more similar to the observed network than the abundance network in terms of nestedness, but the reverse is true of linkage density; both of these relationships are inverted for the presence absence data.

Figure S11. Degree of each node within the observed (relative read abundance-based) and the null networks generated from the different prey availability datasets. The degrees of resource nodes in the relative read abundance-based networks were far more variable and generally much higher, especially compared with the observed network, than those of the presence–absence networks, particularly the degrees of the equal prey abundance null network nodes (Figure S5).

Figure S12. Non-metric multi-dimensional scaling of observed (presence–absence) and null model expected spider diets for the two subfamilies of spiders. These data were generated in models exactly like those used for Figure 4, but with consumers represented as subfamilies (Linyphiinae and Erigoninae; one Lycosidae individual was removed). Colours denote the data source (grey = abundance; other colours are null-model-predicted diets based on the different prey availability datasets). Axes represent a two-dimensional variation in spider diet. The points are centroids (mean coordinates) of diets from each data source for each subfamily/family with distance between them indicating their dissimilarity (i.e., proximate points are similar, distant points are dissimilar). The diet of both taxa more closely resembled the abundance-based null diets than activity density-based null diets, with activity density null diets being most dissimilar to observed diets for both (Euclidean distance = 0.397 for Erigoninae

PREY CHOICE NULL NETWORK DATA SOURCES

Ecological
Entomology

Royal
Entomological
Society

15

and 0.780 for Linyphiinae). Erigoninae diets much more closely resembled abundance (Euclidean distance = 0.328) than Linyphiinae (0.422), but this was the most similar null diet for Linyphiinae while Erigoninae diets more closely resembled equal prey abundance (0.284), FOO combined (0.317) and proportional combined (0.320). Replication of each spider group is not, however, even and larger datasets would be required to fully explore this (Linyphiinae = 57, Erigoninae = 13).