

1 From theory to experiment and back again — Challenges in  
2 quantifying a trait-based theory of predator-prey dynamics

3 K. L. Wootton<sup>1,2\*</sup>, Alva Curtsdotter<sup>3</sup>,  
Tomas Jonsson<sup>1,4</sup>, H.T. Banks<sup>5</sup>, Riccardo Bommarco<sup>1</sup>, Tomas Roslin<sup>1</sup> and A. N. Laubmeier<sup>6</sup>

4 <sup>1</sup>Swedish University of Agricultural Sciences, Department of Ecology, Uppsala, Sweden

5 <sup>2</sup>BioFrontiers Institute, University of Colorado, Boulder, Boulder, CO, USA

6 <sup>3</sup>Insect Ecology Lab, Zoology, The University of New England, Armidale, NSW, Australia

7 <sup>4</sup>Ecological modelling group, University of Skövde, Skövde, Sweden

8 <sup>5</sup>Center for Research in Scientific Computation, North Carolina State University, Raleigh, NC, USA

9 <sup>6</sup>Department of Mathematics & Statistics, Texas Tech University, Lubbock, TX, USA

10 \* Corresponding author:

11 Email: [kate.l.wootton@gmail.com](mailto:kate.l.wootton@gmail.com)

12

## Abstract

13

Food webs map feeding interactions among species, providing a valuable tool for understanding and predicting community dynamics. Trait-based approaches to food webs are increasingly popular, using e.g. species' body sizes to parameterize dynamic models. Although partly successful, models based on body size often cannot fully recover observed dynamics, suggesting that size alone is not enough. For example, differences in species' use of microhabitat or non-consumptive effects of other predators may affect dynamics in ways not captured by body size.

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

Here, we report on the results of a pre-registered study (Laubmeier et al., 2018) where we developed a dynamic food-web model incorporating body size, microhabitat use, and non-consumptive predator effects and used simulations to optimize the experimental design. Now, after performing the mesocosm experiment to generate empirical time-series of insect herbivore and predator abundance dynamics, we use the inverse method to determine parameter values of the dynamic model. We compare four alternative models with and without microhabitat use and non-consumptive predator effects. The four models achieve similar fits to observed data on herbivore population dynamics, but build on different estimates for the same parameters. Thus, each model predicts substantially different effects of each predator on hypothetical new prey species. These findings highlight the imperative of understanding the mechanisms behind species interactions, and the relationships mediating the effects of traits on trophic interactions. In particular, we believe that increased understanding of the estimates of optimal predator-prey body-size ratios and maximum feeding rates will improve future predictions. In conclusion, our study demonstrates how iterative cycling between theory, data and experiment may be needed to hone current insights into how traits affect food-web dynamics.

34

35

36

*Key words: body size, microhabitat use, non-consumptive predator-predator effects, allometric scaling, food webs*

37

## 1 Introduction

38

39

40

Mapping feeding interactions among species in food webs is a crucial first step for understanding how ecological communities function, for gauging the impacts of anthropogenic stress on community structure and stability, and for evaluating how ecosystems might be managed to conserve biodiver-

41 sity and ecosystem functioning (Thompson et al., 2012). However, to achieve quantitative food-web  
42 understanding and predictions, we need a second step of formulating mechanistic models capable  
43 of replicating food-web abundance dynamics, and to develop feasible approaches to parameterize  
44 such models (e.g. Portalier et al., 2019; Schneider et al., 2012). Only by deriving robust parameter  
45 estimates are we then prepared to predict dynamics beyond the range of the existing data, such as  
46 what happens when a new species enters the system.

47 Historically, parameterization of a food-web model required the strength of every trophic link  
48 to be independently estimated experimentally, a laborious and often imprecise process (Roslin  
49 and Majaneva, 2016). Additionally, some elements, such as non-consumptive interactions among  
50 multiple predators, cannot be understood solely from a pairwise predator and prey perspective  
51 (Terry et al., 2020). This complexity has made it unwieldy to use dynamic food-web models to  
52 map the abundance dynamics of diverse predator-prey assemblages in nature. Recent developments  
53 in food-web ecology are now offering a potential cure for this ‘plague of parameters’ (Hudson  
54 and Reuman, 2013) through trait-based and, especially, allometric (body-size based) approaches  
55 (Yodzis, 1998; Schneider et al., 2012; Boit et al., 2012; Curtsdotter et al., 2019). Such models  
56 assume a general relationship between organismal body size and metabolism (Brown et al., 2004;  
57 Peters, 1983), and from this infer a relationship between body size and trophic interaction strength  
58 (Brose, 2010). Allometric Trophic Network (ATN) models (Otto et al., 2007; Schneider et al., 2012;  
59 Berlow et al., 2009) have been formulated based on this idea. They show promising predictions of  
60 observed trophic interaction strengths and abundance dynamics of interacting species (Boit et al.,  
61 2012; Schneider et al., 2012; Curtsdotter et al., 2019; Jonsson et al., 2018), as well as replication  
62 of observed community patterns such as the mass-abundance relationship (Hudson and Reuman,  
63 2013). However, while body size can explain a large portion of observed interaction strengths, in  
64 most cases there remains some substantial unexplained variation, which is potentially attributable  
65 to other, as yet unmeasured traits (Schneider et al., 2012; Jonsson et al., 2018). Thus, although  
66 promising, the general applicability of the ATN modelling approach and the extent to which direct  
67 and especially indirect trophic interactions are determined by traits other than body size, remains  
68 to be explored.

69 Among the more successful applications of the ATN model, Schneider et al. (2012, 2014) discuss  
70 the potential importance of species' 'habitat domain' (Schmitz, 2007), suggesting that differences  
71 in predators' and prey's microhabitat use may explain residual variation where the model did not  
72 accurately capture the experimental data. Motivated by this, Jonsson et al. (2018) combined the  
73 microhabitat use of species with their body size to parameterize an ATN model, thereby successfully  
74 predicting experimentally observed population-level interaction strengths when a predator species  
75 was alone with its prey (i.e. in the absence of indirect effects from other species). While these and  
76 other studies have pointed to the importance of predator and prey habitat use (e.g. Schmitz, 2007;  
77 Knop et al., 2014; Staudacher et al., 2018), it has, to our knowledge, never been explicitly incor-  
78 porated into a dynamic model and parameterized by experimental data. This lack of integration  
79 between theory and empirical validation limits our ability to quantify the effect or importance of  
80 habitat use.

81 The trophic interaction modifications (or indirect trait-based effects) observed in treatments  
82 with more than two species and pinpointed by Jonsson et al. (2018) are often behaviour-mediated  
83 effects of population-level interaction strengths, where changes in the behaviour of a predator and/or  
84 its prey is induced by the presence of another species, thereby modifying the per capita interaction  
85 strength between the predator and its prey (Terry et al., 2017). Mechanisms include avoidance of  
86 intraguild predation and interference among predator species as well as facilitation (Preisser et al.,  
87 2007; Kéfi et al., 2012; Sih et al., 1998; Losey and Denno, 1998; Knop et al., 2014). Such interactions  
88 are not described by the ATN model, and so the model poorly captures their population-level effects  
89 (Terry et al., 2020; Jonsson et al., 2018). Jonsson et al. (2018)'s results strongly suggested that  
90 it is a lack of behavior-based non-consumptive interspecific interference effects in the ATN model  
91 that is the main cause for its inability to accurately predict trophic interaction strength in more  
92 complex webs. Hence, two promising model developments might improve predictions: to consider  
93 the spatial niche of species and/or to account for non-consumptive intra-guild interactions.

94 Here we report on the post-experiment findings of a pre-registered experiment (Laubmeier et al.,  
95 2018) aimed at testing the importance of both microhabitat use and non-consumptive predator-  
96 predator effects in food-web dynamics. To this purpose, we altered the ATN model to include

97 these factors. We introduced a term for microhabitat use, where predators and prey will encounter  
98 each other more frequently the more time they spend in the same area, thereby showing a stronger  
99 interaction strength. We also included a term for non-consumptive predator-predator effects, where  
100 avoidance of other predators due to the fear of intraguild predation (e.g. Lima, 2002) or interference  
101 by other predators decreases predation rate. Finally, to establish whether the effects of microhabitat  
102 use and non-consumptive predator-predator effects were sufficiently strong to be observed across a  
103 diverse range of predators, we intentionally selected diverse predators covering a range of guilds and  
104 feeding modes. To accommodate effects of variation in e.g. feeding mode on optimal predator-prey  
105 body-mass ratio for different types of predators, we allowed the value of the optimal predator-prey  
106 body mass ratio to vary from predator to predator in the parameter estimation that followed -  
107 thereby departing from other studies utilizing the ATN model, such as Schneider et al. (2012) and  
108 Jonsson et al. (2018).

109 To arrive at an optimized design for generating empirical data to inform theoretical models,  
110 we (Laubmeier et al., 2018) *a priori* explored the optimal and minimal timing and frequency of  
111 experimental sampling to provide sufficient data on population dynamics to enable the use of the  
112 inverse method (Chowell, 2017; Banks et al., 2014) for parameter estimation and model testing  
113 parameter values of the model. In the current paper, we now perform these steps and test our  
114 model against the resulting empirical data, developing a habitat overlap metric in the process.

## 115 **2 The model**

116 We model predator-prey population dynamics in a food web, inferring interaction strengths from  
117 body size and microhabitat use. Stronger interactions occur when prey are close to a predator's  
118 optimal prey size, or when predator and prey overlap more in their microhabitat use. To develop  
119 our model, we started with the Allometric Trophic Network (ATN) model, which uses body sizes  
120 of predator and prey to dictate interaction strengths (Brose, 2010; Otto et al., 2007; Schneider  
121 et al., 2012). We then modified the ATN model to include habitat overlap and non-consumptive  
122 predator-predator interactions.

123 Our modified model was published in Laubmeier et al. (2018). Subsequent to publication, we

124 observed that our original formulation for similarity in microhabitat use did not always capture the  
125 amount of time predator and prey were in the same location and therefore their likely frequency of  
126 interaction. We have, therefore, modified the habitat overlap index to account for this. Below, the  
127 entire, updated model is presented and described.

128 To account for differences in habitat use across species, we divide the habitat into microhabitat  
129 zones, quantify the amount of predation that occurs in each microhabitat zone, and sum across all  
130 zones. The amount of predation increases with the amount of time spent in the microhabitat ( $p$ )  
131 and decreases with the size of the microhabitat ( $A$ ). When  $p_{i,h}$  is large, species  $i$  spends more time  
132 in habitat  $h$ , and if  $p_{i,h}$  and  $p_{j,h}$  are both large, we expect species  $i$  and  $j$  to encounter one another  
133 more often in habitat  $h$ . Here, we measure  $p_{i,h}$  empirically.

134 We also introduce a term that describes the decrease in predation by a predator due to non-  
135 consumptive effects of other predators. This may include fear of predation, leading to decreased  
136 foraging, or physical interference (Preisser et al., 2007; Sih et al., 1998). We propose that the  
137 magnitude of this effect depends on the likelihood of predator  $j$  being intraguild prey to predator  
138  $l$ , and therefore depends on the expected attack rate of  $l$  on  $j$  ( $a_{jl}$ ). Microhabitat overlap will also  
139 affect predator encounters and should therefore affect the magnitude of non-consumptive effects  
140 (e.g. Knop et al., 2014). We account for the effects of microhabitat overlap on non-consumptive  
141 predator-predator effects in the same way as described above for predator-prey interactions. We  
142 sum over the potential attack rates of all species  $l$  on a single individual of species  $j$  to account for  
143 time spent avoiding or evading species  $l$  while species  $j$  is attempting to capture its own prey. The  
144 importance of non-consumptive predator-predator effects is described by the scaling constant  $t_0$ ,  
145 where a large value indicates a high penalty to attack rates due to non-consumptive effects. Non-  
146 consumptive effects from a conspecific individual may not be distinguishable from non-consumptive  
147 effects from another predator species, and so we remove the intraspecific competition term as used  
148 in Schneider et al. (2012) from this version of the ATN model and replace it by the more general  
149 expression for non-consumptive effects from other predator individuals of any species.

150 In total, dynamics for the number of individuals  $N_i$  of species  $i$  are therefore given by:

$$\frac{dN_i}{dt} = r_i N_i - \sum_j \sum_h \frac{\frac{a_{ij}}{A_h} p_{i,h} N_i p_{j,h} N_j}{1 + \sum_k \frac{a_{kj}}{A_h} h_{kj} p_{k,h} N_k + t_0 \sum_l \frac{a_{jl}}{A_h} p_{l,h} N_l} \quad (1)$$

151 where species  $i$  increases in proportion to its intrinsic growth rate  $r_i$  and decreases due to  
 152 predation. We assume the intrinsic growth rate ( $r_i$ ) for predators to be zero due to their much  
 153 longer generation time (a year) compared with the duration of our experiment. The realized per  
 154 capita attack rate of predator  $j$  on species  $i$  in a microhabitat  $h$  ( $\alpha_{ijh}$ ) increases with the intrinsic  
 155 attack rate determined by the predator-prey body-mass ratio ( $a_{ij}$ , see below) and decreases with  
 156 the size of the microhabitat, ( $\alpha_{ijh} = a_{ij}/A_h$ ), because predator and prey encounter each other less  
 157 frequently in the larger area. Total predation in a microhabitat increases as the proportion of prey  
 158 species  $i$  ( $p_{i,h} N_i$ ) and predator species  $j$  ( $p_{j,h} N_j$ ), in habitat  $h$  increases, but decreases dependent  
 159 on the time predator  $j$  spends handling prey of the same or other species ( $h_{kj}$ ), or spends avoiding  
 160 or interfering with other predators  $l$ .

161 As in Schneider et al. (2012), we assume that for species body masses  $W_i$  and  $W_j$  (corresponding  
 162 to prey  $i$  and predator  $j$ ), the allometric parameters (i.e. those dependent on body mass) are given  
 163 by:

$$a_{ij} = a_0 W_i^{1/4} W_j^{1/4} \left( \frac{W_j/W_i}{R_{opt,j}} e^{1 - \frac{W_j/W_i}{R_{opt,j}}} \right)^\phi \quad (2)$$

$$h_{ij} = h_0 W_i^{1/4} W_j^{-1/4}$$

164 The derivation of allometric parameters is described in Schneider et al. (2012). We note the  
 165 importance of scaling parameters  $a_0$ ,  $h_0$ , and  $R_{opt,j}$ .  $a_0$  scales the frequency of attacks when species  
 166 encounter one another, with larger values of  $a_0$  indicating more frequent attacks.  $h_0$  scales the  
 167 time spent handling alternative prey items. Larger values of  $h_0$  indicate more time spent handling  
 168 prey which results in attacks on a lower portion of the prey population when prey are abundant.  
 169  $R_{opt,j}$  indicates the optimal predator-prey body-mass ratio for a successful attack by predator  $j$ ,

170 where  $R_{opt,j} = 1$  indicates that predator  $j$  is most successful when attacking prey as large as itself  
171 and  $R_{opt,j} \gg 1$  indicates that predator  $j$  is most successful when attacking prey much smaller than  
172 itself. Parameter  $\phi$  ( $\phi \geq 0$ ) tunes the width of this success curve, with  $\phi = 0$  indicating that  
173 attack success is independent of prey size, while the greater the value of  $\phi$  the more restricted the  
174 attack success around  $R_{opt}$ . In contrast to Schneider et al. (2012), we allow the value of  $R_{opt}$  to  
175 vary from predator to predator. This is to account for differences in traits not accounted for in the  
176 model that may affect predator foraging behavior.

177 To determine the importance of the terms we introduce — microhabitat overlap and non-trophic  
178 predator-predator effects — we compare four variations of the model:

- 179 1. the full model (Eq. 1, i.e. with  $t_0 > 0$  and including  $p_{i,h}$  and  $A_h$ )
- 180 2. an intermediate model with habitat use but without predator interference (setting  $t_0 = 0$ )
- 181 3. an intermediate model with predator interference but without habitat use (removing  $p_{i,h}$  and  
182  $A_h$ )
- 183 4. a minimal model without habitat use or predator interference (removing  $p_{i,h}$  and  $A_h$  and  
184 setting  $t_0 = 0$ )

185 The different parameters of Eqs 1-2 affect (i) which prey a predator is most likely to consume  
186 and (ii) to what extent. As such, the values that we estimate here for  $a_0$ ,  $h_0$ ,  $R_{opt}$  etc. for our  
187 four model variants will, to some extent, reflect how each model emphasizes the importance of  
188 different factors for each interaction. To begin with, a predator with an  $R_{opt}$  value of 1 will interact  
189 most strongly with prey of its own size, while a predator with an  $R_{opt}$  of 100 will more effectively  
190 consume prey 100 times smaller than itself. Next, if a predator spends all its time on bean plants  
191 ( $p_{i,beans} = 1$ ) it will have the strongest interaction with prey that also reside predominantly on  
192 beans, and have no interaction with prey that are never on beans. Crucially, when the model  
193 includes habitat use, it is the combination of both  $R_{opt}$  and  $p_{i,h}$  that dictates trophic interaction  
194 strength. For example, consider prey  $x$  (size = 10 and  $p_{x,beans} = 1$ ) and prey  $y$  (size = 1 and  
195  $p_{y,barley} = 1$ ) with predator  $i$ . If predator  $i$  (size = 100 and  $p_{i,beans} = 1$ ) only interacts with prey  $x$ ,



196 this is already captured entirely by the microhabitat use terms since the predator spends all its time  
197 in the same microhabitat zone as prey  $x$  and never overlaps with prey  $y$ . In a model accounting  
198 for habitat use (as in models 1 and 2),  $R_{opt,i}$ , therefore, will likely have an estimated value near  
199 10, since predator  $i$  is ten times larger than prey  $x$ . If, however, we do not account for habitat  
200 use (as in models 3 and 4), the model needs some other way to capture that predator  $i$  does not  
201 interact with prey  $y$  in order to optimize the fit to empirical data. In the parameter estimation this  
202 could be achieved by a larger value of  $R_{opt,i}$  (moving its optimal prey size further from prey  $y$ ),  
203 or a higher value of  $\phi$  (narrowing the effective feeding range), thus absorbing differences based on  
204 which terms are present in the model and producing good model fit without necessarily reflecting  
205 the 'true value' of a parameter. This is important to remember when interpreting the results of the  
206 model fitting below.

207 We do not explicitly include non-consumptive mortality in this model. For the aphid (or basal)  
208 prey, mortality not due to predation is included in the growth rate term  $r_i$ , while for predators the  
209 experiment is not long enough that we expect mortality other than that due to intraguild predation.  
210 Furthermore, without single individual controls, it would be difficult to separate "natural" mortality  
211 from that due to predation or cannibalism.

## 212 **3 Methods**

### 213 **3.1 The mesocosm experiment**

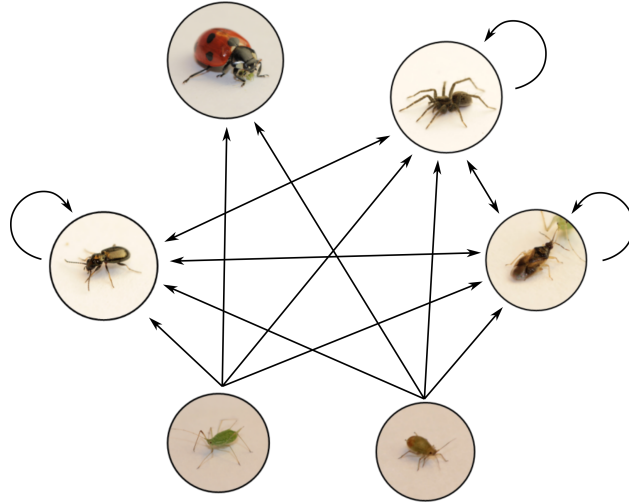
214 To empirically test and parameterize this model required a study system with rapid growth of  
215 the prey population, a range of body sizes of both predators and prey, and distinct habitat zones.  
216 With this in mind, and with the benefit of data from a previous experiment (Jonsson et al., 2018),  
217 we assembled a six-species terrestrial arthropod community (Laubmeier et al., 2018) (figure 1)  
218 dependent on two species of plants; barley (*Hordeum vulgare*) and fava beans (*Vicia faba*).

219 As primary consumers we chose one large (*Acyrtosiphon pisum*) and one small (*Rhopalosiphum padi*)  
220 species, both aphids. Next, to explore the importance of body mass and microhabitat  
221 use in trophic interactions, we chose four predators on these prey, differing in body size and/or  
222 habitat preference; one large and one small predominantly foliage-dwelling predator (*Coccinella*

223 *septempunctata* and *Orius majusculus* respectively), and one large and one small predominantly  
224 ground dwelling predator (*Pardosa* spp. and *Bembidion* spp. respectively, where spp. signals  
225 the potential inclusion of several congeneric but morphologically indistinguishable species). Each  
226 mesocosm contained both barley (*Hordeum vulgare*) (as a host for *R. padi*) and fava beans (*Vicia*  
227 *faba*) (as a host for *A. pisum*), one or both aphid species and zero, one or two predator species. All  
228 combinations of predator and prey were replicated six times in a fully factorial design (figure 2).  
229 This resulted in 30 predator-prey combinations, plus three control treatments with no predators.

230 Plants were sown in 60x40cm, 20cm deep, plastic containers, sown with two rows of 10 fava bean  
231 seedlings and three rows of 15 barley seedlings. A 60cm high mesh cage, with one side resealable  
232 to allow aphid counting, was placed on top of each container to prevent insects from entering or  
233 escaping the microcosm.

234 150 wingless adult aphids, placed on Petri dishes, were introduced per microcosm two days  
235 before the experiment began. One third of the mesocosms (66 mesocosms) were inoculated with  
236 150 *R. padi* (zero *A. pisum*), one third with 150 *A. pisum* (zero *R. padi*), and the final third with  
237 75 *R. padi* and 75 *A. pisum*. Predators were introduced at the beginning of the experiment. The  
238 number of predators was determined using a combination of short term (8hr) feeding trials and  
239 pilot studies to reach a density where predators would impact the prey, but not eliminate them too  
240 quickly. Predator numbers in single-species mesocosms were: *C. septempunctata*: 4 individuals; *O.*  
241 *majusculus*: 40 individuals; *Pardosa*: 20 individuals; *Bembidion*: 40 individuals. Mesocosms with  
242 two predators species contained half the number of individuals of each predator species as single  
243 predator-species mesocosms, i.e. mesocosms with both *Pardosa* and *Bembidion* together contained  
244 10 *Pardosa* individuals and 20 *Bembidion* individuals



245

Figure 1: The food web including all possible interactions that we allowed in the model. Species are, from top left: lady beetle (*Coccinella septempunctata*); wolf spiders (*Pardosa* spp.); minute pirate bug (*Orius majusculus*); bird cherry-oat aphid (*Rhopalosiphum padi*); pea aphid (*Acyrtosiphon pisum*); and ground beetle (*Bembidion* spp). Arrows indicate potential feeding interactions which we then parameterized using the inverse method. Arrows point from prey to predator. Double headed arrows indicate that species could potentially eat each other and arrows beginning and ending with the same species indicate cannibalism. We removed all interactions to and from *C. septempunctata* except for *C. septempunctata* preying on aphids, and assumed that the aphids did not consume any predators. This arthropod community was dependent on two species of plants; barley (*Hordeum vulgare*) and fava beans (*Vicia faba*).

246 Frequency and timing of aphid counts were determined based on our pre-experimental analyses  
247 (Laubmeier et al., 2018). Due to a minor difference in method compared to previous experiments  
248 (Jonsson et al., 2018) that made data collection quicker than expected (in-cage rather than destruc-  
249 tive sampling), we increased sampling slightly from the minimum determined in Laubmeier et al.  
250 (2018). Aphid populations were counted on days 2, 4, 6 and 8. Treatments with *C. septempunctata*  
251 were also counted on days 1 and 3, because we realized that *C. septempunctata* decimated aphid  
252 populations so rapidly that we would require more data points in order to obtain an estimate of  
253 their parameters. Aphids were counted by opening the cage door and carefully counting the number  
254 of aphids on each plant.

255 The proportion of time predators spent in each habitat,  $p_{j,h}$ , was measured in single predator  
256 mesocosms. The location of each predator was marked on a mesocosm map before beginning to  
257 count aphids in these mesocosms. We then categorized these into four areas: walls/roof, ground,  
258 beans, and barley. Aphid habitat use was measured by separating aphid counts into each of those  
259 categories, but only recorded on days 2 and 6. While the absolute area of beans and barley changed

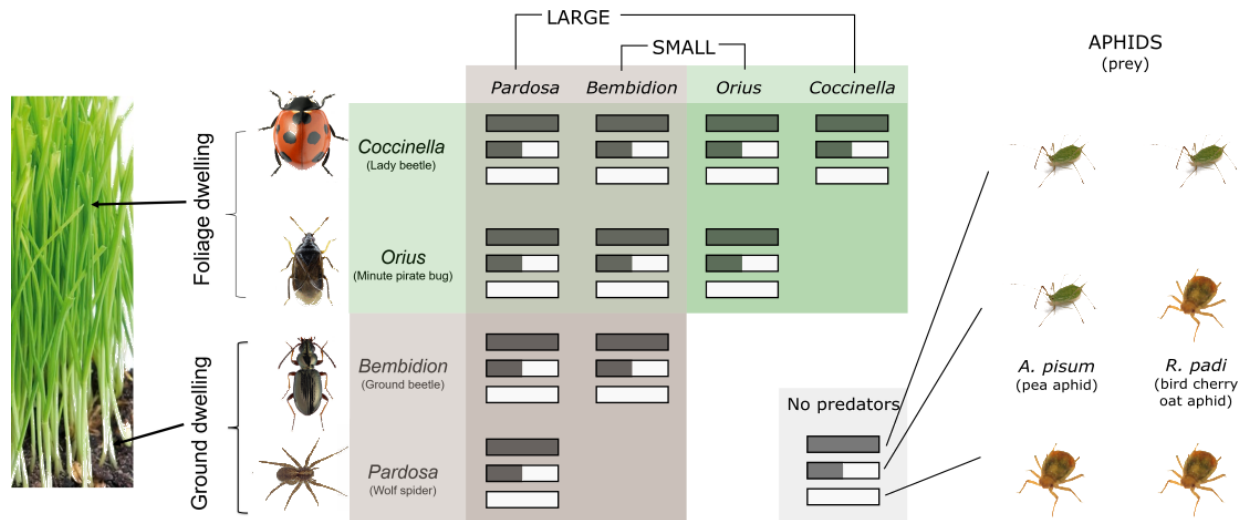


Figure 2: An overview of the predator-prey combinations used in the experiment. Each combination was replicated six times.

260 throughout the experiment, we estimated that, on average, the surface area of beans, of barley, and  
 261 of the ground were roughly equivalent, while the area of the walls and roof was six times larger  
 262 than each other habitat area (figure 4).

263 Predators could only be reliably counted through destructive sampling of the mesocosms, and  
 264 were therefore only counted on the final day (day 8). After the aphid count, predators were collected  
 265 by a thorough examination of cage and plant, and sifting through the soil. An additional predator  
 266 search was repeated the next day to catch any missed in the initial search.

267 Over the duration of the experiment, we found that *C. septempunctata* could occasionally  
 268 escape through gaps in the mesh cages. We assume that any *C. septempunctata* missing from cages  
 269 escaped in this manner, as other predators were never observed consuming *C. septempunctata*.  
 270 Because this change in the population is not described by our mathematical model, we added  
 271 replacement individuals to cages where *C. septempunctata* went missing and did not dynamically  
 272 model the population. Instead, we directly input *C. septempunctata* population densities into the  
 273 model for other species' population dynamics. We fixed these densities at constant levels for the  
 274 duration of the experiment by taking the average value of all observed abundances in each cage. We  
 275 used averages instead of time-series data due to the uncertainty associated with our observations;  
 276 it was impossible to know exactly when between observations the individuals went missing from

277 the mesocosm.

278 The ATN model describes species interaction strengths as a function of species traits (in our  
279 case body size and microhabitat use). Because the presence or absence of a food-web link is simply  
280 a binary interpretation of interaction strength, the ATN model also predicts the binary food-web  
281 structure. However, if a feeding interaction is prohibited due to traits not accounted for in the  
282 model, it cannot be expected to correctly predict the absence of such links. As neither body size  
283 nor microhabitat could explain why the other predators did not consume *C. septempunctata* (there  
284 was microhabitat overlap and predation of similar-sized intraguild prey), we removed feeding inter-  
285 actions between *C. septempunctata* and other predators from the network of potential interactions  
286 (figure 1). Similarly, *C. septempunctata* did not consume *Bembidion* or *Pardosa*, for reasons not  
287 necessarily explained by microhabitat use or body size (most likely *Bembidion*'s hard cuticle (e.g.  
288 Brousseau et al., 2018) and *Pardosa*'s speed), so we removed these interactions.

### 289 **3.2 Model fitting**

290 Using abundance data from our experiment, we parameterized four versions of the ATN model  
291 (based on different assumptions for habitat use and predator interference, see below) through the  
292 inverse method (Banks et al., 2014; Chowell, 2017). Under this method we first formulated a  
293 least squares cost criterion (JLS) (Banks and Tran, 2009; Banks et al., 2014), which describes the  
294 deviation of model predictions from empirical observations. Next, by minimizing this cost across  
295 all possible parameterizations, we found the best-fitting parameterization. In order to compare the  
296 importance of habitat use and predator interference, we repeated this fitting for each of the four  
297 models.

298 To fit the models, we first established a common baseline for aphid growth. Using the data  
299 from control treatments, we estimated the intrinsic growth rate ( $r_i$ ) for *A. pisum* and *R. padi*.  
300 The rate was distinct for each aphid species, but we assumed the same rate for each aphid species  
301 across all aphid treatments (single-species or combined) and replicates. After this, we estimated  
302 the remaining model parameters using data from predator-treated mesocosms. There were 10  
303 predator treatments, each with 18 aphid populations; 6 replicates of each of *A. pisum* and *R. padi*  
304 in isolation, as well as 6 replicates of *A. pisum* and *R. padi* when in combination with each other.

305 We simultaneously estimated constants for allometric relationships ( $a_0$ ,  $h_0$ ,  $R_{opt,j}$ ,  $\phi$ ) and predator  
306 interference ( $t_0$ ) alongside initial aphid abundances. Although model parameters must be the same  
307 across all treatments and replicates (but vary among the four models), initial aphid abundances were  
308 permitted to vary in every replicate mesocosm. This allowed for differences in population outcomes  
309 due to external, potentially stochastic factors, such as variation in plant growth. We constrained  
310 estimates of these initial abundances to a range determined by observations from control mesocosms  
311 on day 0 of the experiment (ranging from 125 to 775, with a median of 205). Other parameters  
312 were unconstrained, except  $R_{opt}$  which was bound between 1 and 1000. This was capped because  
313  $R_{opt}$  values could otherwise reach any upper bound permitted, but in reality increases over 1000  
314 had very little effect on  $a_{ij}$ . For the reduced models, we utilized the same values for  $r_i$  as in the  
315 full model and repeated the process for estimating parameters and initial abundances. To remove  
316 predator interference, we set  $t_0 = 0$  and did not estimate that parameter. To remove habitat use,  
317 we set  $p_{i,h} = 1$  and  $A_h = 1$ , removing the summation over all  $h$ .

### 318 **3.3 Model evaluation and prediction**

319 Notably, the above steps utilize the entire data set to derive parameter estimates. Each estimation  
320 problem yields a cost criterion (JLS) quantifying model fit, where a lower value of JLS indicates a  
321 better model fit. We can also evaluate the performance of each model according to the realism of  
322 estimated parameter values and associated processes (e.g. feeding rates), compared to literature or  
323 supplemental empirical testing. To summarize, we used a mix of statistical and expert-knowledge  
324 model performance criteria, specifically: (i) the JLS cost criterion, (ii) visual fit of predicted vs  
325 observed aphid abundance, (iii) realism of parameter values, (iv) realism of ecological processes,  
326 and (v) observed vs predicted final predator abundance.

327 To evaluate performance for these parameter values outside the observed data, we generated  
328 predictions for dynamics using an alternative food-web configuration. To demonstrate how the  
329 models might compare in their predictions for a new prey species, we used each model, with its  
330 resulting parameterization, to predict the population on days 2, 4, 6, and 8 of a hypothetical,  
331 entirely ground-dwelling prey species weighing 1mg (slightly larger than *A. pisum*) paired with

332 each of our four predator species.

## 333 4 Results

Table 1: Parameter values and model fit (JLS) for models with and without habitat use and non-consumptive predator-predator effects. Model 1 = full model with both habitat and non-consumptive effects. Model 2 = only habitat. Model 3 = Only non-consumptive effects. Model 4 = minimal model, neither habitat nor non-consumptive effects.  $a_0$ ,  $h_0$ ,  $t_0$  and  $\phi$  refer to parameters in Eqs. 1 and 2.  $R_{opt.P}$ ,  $R_{opt.O}$ ,  $R_{opt.C}$  and  $R_{opt.B}$  refer to the optimal predator-prey body-size ratio for *Pardosa* spp., *O. masculus*, *C. septempunctata*, and *Bembidion* spp. respectively.

Model	$a_0$	$h_0$	$t_0$	$\phi$	$R_{opt.P}$	$R_{opt.O}$	$R_{opt.C}$	$R_{opt.B}$	JLS
1	47.37	0.014	13.66	1.41	1	4	1000	108	1.20e+08
2	7.78	0.015	-	1.23	30	997	218	8	1.18e+08
3	0.32	0.063	9.14	1.64	1	3	360	389	1.22e+08
4	0.39	0.057	-	1.83	1	78	362	377	1.38e+08

334 All four models gave similar fits to the data (JLS values in table 1 and figure 3), but did so  
335 by having different parameter values, particularly  $a_0$  and  $R_{opt,j}$ . This is not surprising since both  
336 parameters can be expected to absorb differences based on which terms are present in the model,  
337 without necessarily reflecting the ‘true value’ of a parameter (see explanation at end of section  
338 2). For example,  $a_0$  scales the frequency of attacks when species encounter each other; a higher  
339 value of  $a_0$  signals a higher likelihood of an attack after an encounter and — all else being equal  
340 — will lead to a higher predation rate. At the same time, differences in habitat overlap between  
341 species decrease encounters, as do non-consumptive predator effects, which — all else being equal  
342 — would lead to a lower predation rate. Thus, in models from which the latter two terms are  
343 absent, a certain (observed) predation rate can only be achieved by a lower likelihood of attack (i.e.  
344 lower  $a_0$ ) than in models where these terms are present (because here a higher  $a_0$  can be countered  
345 by a decrease in attacks due to fewer encounters). In a similar vein,  $R_{opt,j}$  describes the optimal  
346 predator-prey body-size ratio for predator  $j$ . A larger value means predator  $j$  preferentially attacks  
347 prey much smaller than itself, while an  $R_{opt}$  of 1 means that predators preferentially attack prey  
348 the same size as themselves. Because we allowed  $R_{opt}$  to vary among predators, we found that a  
349 particular predator’s  $R_{opt}$  value could vary substantially among models, thereby compensating for



350 differences in attack rate which were or were not accounted for by habitat use or non-consumptive  
351 predator effects in that particular model.

352 As a result of all this, we saw different parameter values in the different models (Table 1). The  
353 different parameter values produced different null-model predictions of feeding rates (i.e. predictions  
354 for the scenario where predators and prey used all habitats in proportion to the habitat area; see  
355 colored lines in figure 5, here termed ‘*potential* feeding rates’). Despite this, all models produced  
356 similar ‘*realized* feeding rates’ of predators on aphids when actual habitat use was taken into account  
357 (larger dots at darker vertical lines in figure 5). Consequently, understanding how this is achieved  
358 will be useful for deciphering what each model assumes about the foraging behavior of the predators  
359 involved.

360 As expected, the scaling parameter for attack rate,  $a_0$  was highest in the models with habitat  
361 overlap (where encounters were limited), especially the model which also included non-consumptive  
362 predator-predator effects (Table 1).  $R_{opt,j}$  values also varied significantly among models, suggesting  
363 that this parameter is also absorbing differences based on which terms are present in the model.  
364 The optimal predator-prey body-mass ratio for *Pardosa* was equal to one in three out of four  
365 models, implying that *Pardosa* is capable of and willing to attack prey of its own size. Such  
366 behaviour by *Pardosa* is in line with our own observations made during the experiment, where  
367 *Pardosa* individuals readily consumed each other. In models with non-consumptive effects, the  
368 optimal prey body size for the small *O. majusculus* was large as well (i.e. small  $R_{opt}$ ) - a pattern  
369 also in line with our observations from the feeding trials, where *O. majusculus* individuals readily  
370 consumed *A. pisum* individuals of approximately the same size as themselves. In stark contrast,  
371 model 2, which includes habitat use but lacks non-consumptive predator effects, has a value of  
372  $R_{opt,O}$  very close to the upper limit of 1000. In this model with habitat overlap but not non-  
373 consumptive predator effects, a small value of  $R_{opt}$  (corresponding to large optimal prey), would  
374 result in a much stronger effect of *O. majusculus* on *A. pisum* than we observe in the data. This  
375 is for two reasons. First, because *O. majusculus* had relatively high microhabitat overlap with *A.*  
376 *pisum* (figure 4) the model would predict many attacks by *O. majusculus* on *A. pisum* due to many  
377 encounters and high  $a_0$ . Second, the model predicts that non-consumptive predator effects strongly



378 decrease foraging (and therefore encounters) when  $R_{opt,j}$  is close to one and conspecific individuals  
379 become a threat. A model without non-consumptive predator effects, therefore, would not show a  
380 decrease in attacks on *A. pisum* due to *O. majusculus* avoiding conspecific potential predators. The  
381 high value of  $R_{opt}$  which the model in fact fitted, accounted for the lower-than-otherwise-expected  
382 effect of *O. majusculus* on *A. pisum*, by driving down attack rates of *O. majusculus* on all prey. *C.*  
383 *septempunctata* seemed to prefer smaller prey, with  $R_{opt}$  values ranging from 218 to 1000. *R. padi*,  
384 which *C. septempunctata* had a very strong effect on, has a body-size ratio with *C. septempunctata*  
385 of 242, so this range of  $R_{opt}$  values reflects *C. septempunctata*'s impact on *R. padi*. *Bembidion* also  
386 tended to have larger values of  $R_{opt}$ , especially when the model did not include habitat use. This  
387 may reflect the fact that *Bembidion* has a negligible impact on *A. pisum*. When habitat use was  
388 included in the model, the negligible effect of *Bembidion* on *A. pisum* was accounted for by the  
389 fact that they had very little overlap. Without habitat, the larger value of  $R_{opt}$  predicted a lower  
390 interaction strength with *A. pisum*.

391 Other parameters varied less among models. Handling time was uniformly low across all models,  
392 ranging from 0.014 to 0.063. When included in the model,  $t_0$  was high (13.66 and 9.14 with and  
393 without habitat respectively), suggesting that non-consumptive predator effects are important in  
394 those models.  $\phi$  ranged from 1.23 to 1.41, which tells us that the importance of predator-prey body  
395 size was roughly consistent across models.

396 While the models fit very similarly to the dynamics of aphid populations, they gave different  
397 predictions of predator populations (figure 6). The full model usually gave the best prediction of  
398 final predator population sizes. Omitting only non-consumptive predator-predator effects from the  
399 model tended to estimate a stronger decline of predator populations than observed in the data,  
400 while omitting both non-consumptive effects and habitat tended to underestimate the decline. The  
401 exception here is *Pardosa*, where the full model over-estimated the decline. This is because *Pardosa*  
402 had a small  $R_{opt,j}$  value in the full model, making cannibalism a common occurrence.

403 Despite the different parameter values estimated for each model, they combined to yield rel-  
404 atively similar *realized* feeding rates (dots in figure 5), especially on the aphids *R. padi* and *A.*  
405 *pisum* (with the exception of *Bembidion* and *O. majusculus* in the habitat-only model). *Potential*

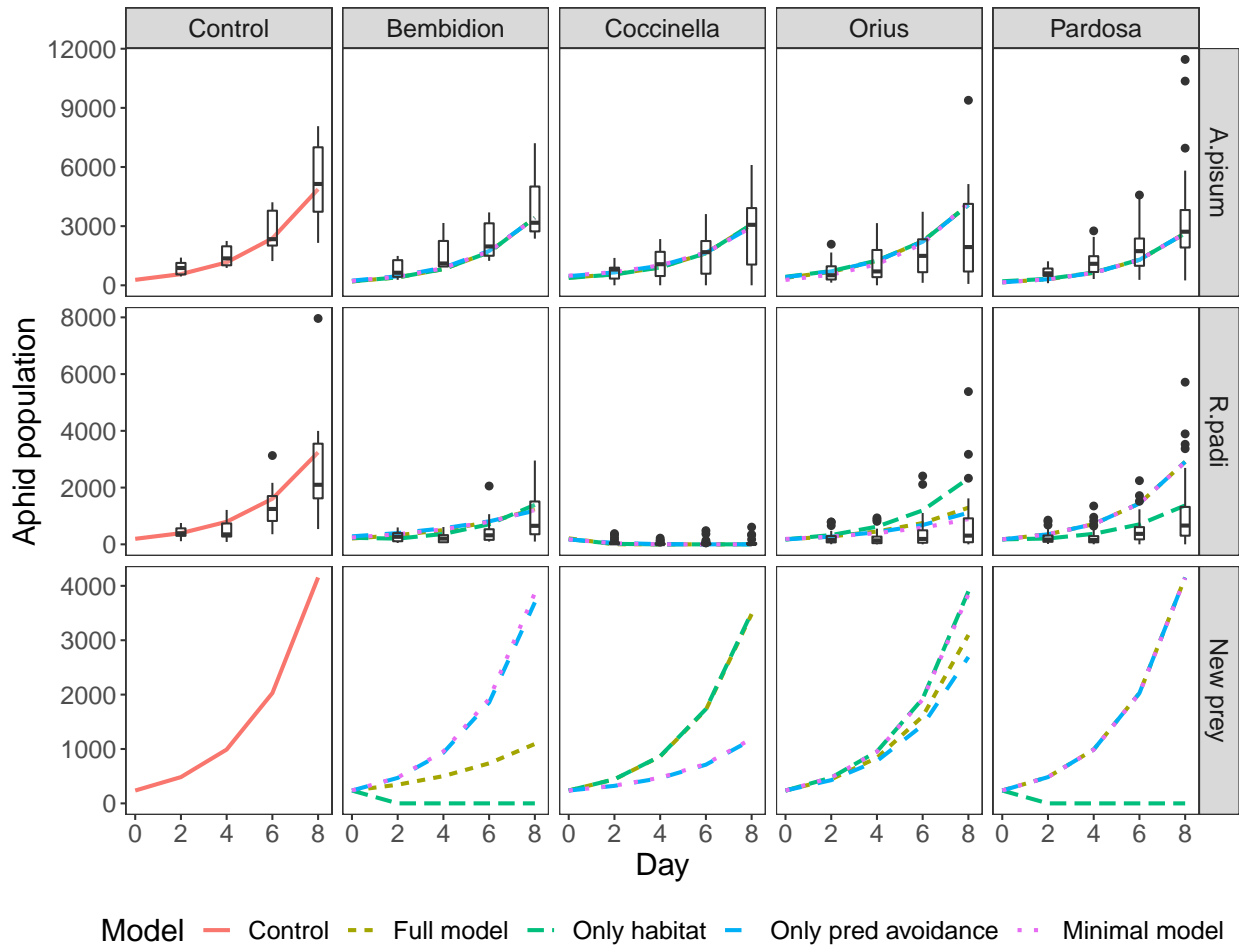


Figure 3: Model predictions of aphid population growth (lines) across time, compared to data of aphid counts per day (boxes) in single-aphid (rows) single-predator (columns) treatments. Lines show predictions of the different models. The final row shows model predictions when fitting each model to a hypothetical new prey species that resides entirely on the ground and has a body size of 1mg.

406 feeding rates (lines in figure 5) were higher in the models including habitat overlap, but were (in  
407 most cases) dramatically decreased when accounting for the actual habitat use of species. This  
408 decrease is shown by the difference between the dots and the lines in figure 5. The few cases where  
409 the *realized* feeding rate was larger than the *potential* feeding rate (e.g. the interaction strength  
410 of *Bembidion* with itself) occurred when species had similar (or in this case, identical) habitat use,  
411 especially in a relatively small habitat area. These *potential* feeding rates are primarily relevant  
412 when extrapolating to other species combinations or habitat configurations; when fitting the data  
413 from our experiment, it is the *realized* feeding rates, including the effects of habitat overlap and  
414 non-consumptive predator-predator effects, that matter. This is why such different models can  
415 produce such similar fits to observations.

416 However, the fact that the different models all produced similar fits to observations, especially  
417 of aphid dynamics, does not mean the models are equivalent. When we used the models to make  
418 predictions for a new species, each model gave very different predictions (bottom row in figure  
419 3). Using a hypothetical, entirely ground-dwelling prey species slightly larger than *A. pisum*, each  
420 model predicted vastly different impacts of the predators on the prey population. Without habitat  
421 use, the models predicted that *Bembidion* and *Pardosa* would have no impact on the prey, *C.*  
422 *septempunctata* would be a strong predator, and *O. majusculus* would fall somewhere in between.  
423 With habitat, *C. septempunctata* was predicted to have very little impact on the prey (since *C.*  
424 *septempunctata* was almost never on the ground). *Pardosa* and *Bembidion* were predicted to have  
425 the strongest effect when habitat, but not non-consumptive effects, was included in the model.

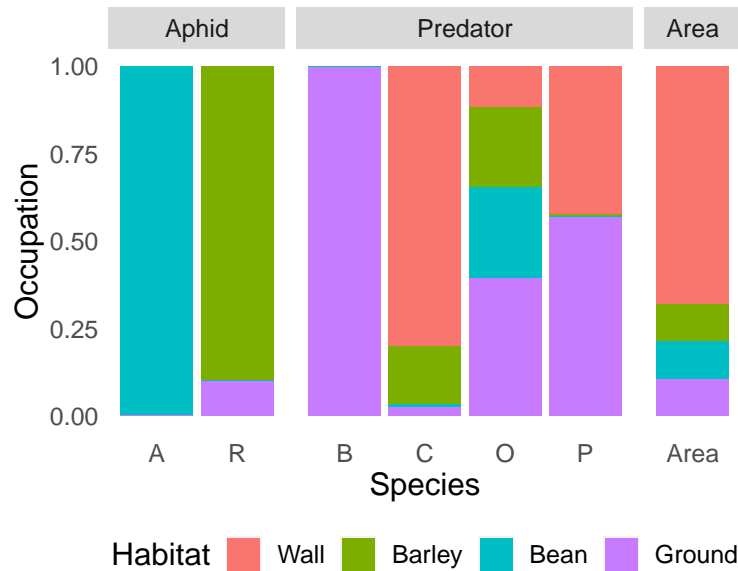


Figure 4: Proportion of time each species spent in each of the four microhabitat zones. The final column shows the relative size of each area. The prey species were A = *A. pisum*, R = *R. padi*, while the predators were B = *Bembidion*, C = *C. septempunctata*, O = *O. majusculus*, P = *Pardosa*.

## 5 Discussion

Here, we have reported on the post-experiment results of a pre-registered study aimed at developing a dynamic food-web model, taking into account body size, habitat use, and non-consumptive predator-predator effects. We used the inverse method to determine parameter values of the dynamic model from time-series data from mesocosm experiments. When comparing the fits of four alternative models with or without microhabitat use and non-consumptive predator-predator effects to the dynamics of two aphid species and their predators, we found that objectively (based on the JLS criterion alone) the four models fit equally well and predicted similar aphid population dynamics. Yet, the four models had different values for key parameters, making it difficult to determine which model, if any, could be interpreted as being the best. By having different parameter values, the models ascribed different mechanisms to similar dynamic outcomes. These different mechanisms matter when models are applied beyond the data range, as we showed with vastly different predicted effects of each predator on a hypothetical new prey species across the different models. Below, we examine each finding in turn.

With respect to our model performance criteria, we found that all four models performed rel-

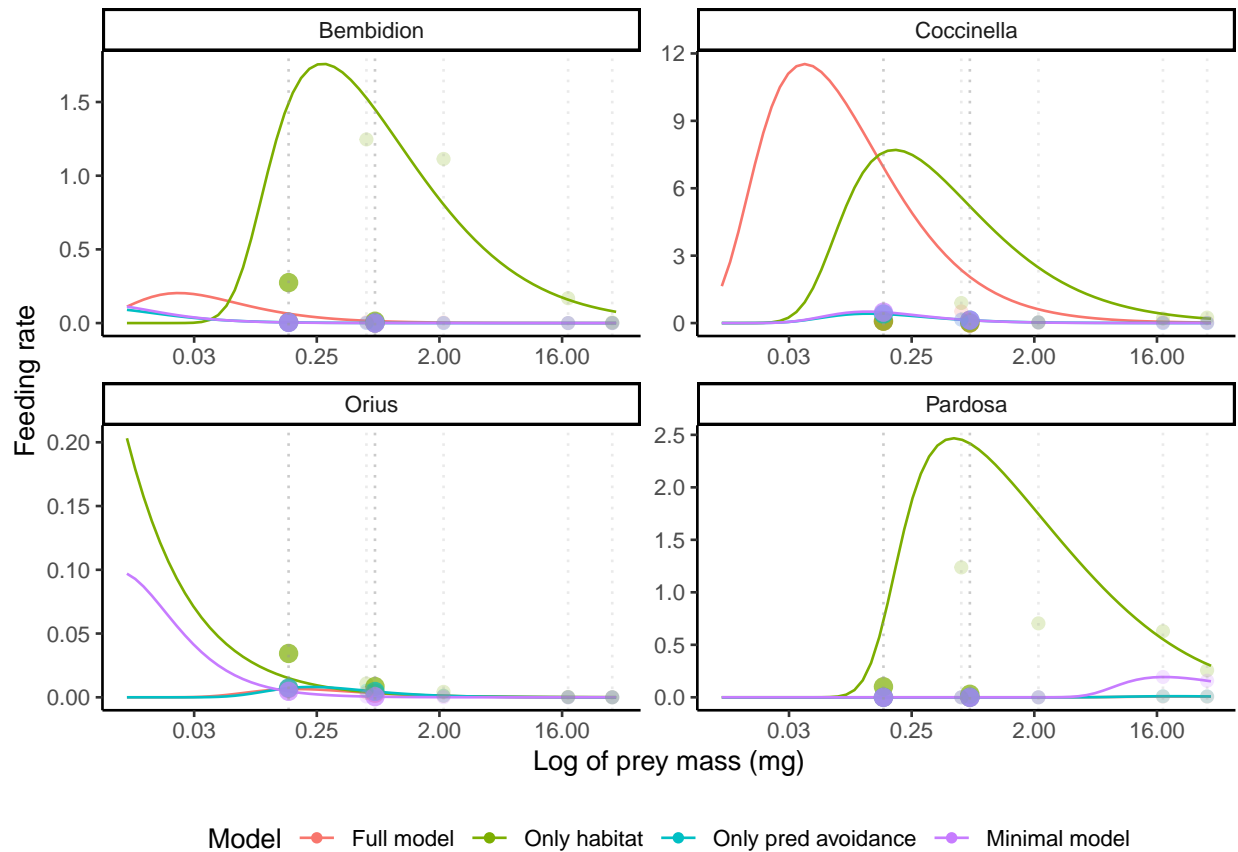


Figure 5: Predictions of each model for each predator population’s feeding rate (y axis) on prey of different body sizes (x axis). Lines show the null expectation or *potential* feeding rate, i.e. predicted feeding rate as a function of prey size if predator and prey used all habitats in proportion to the area of the habitat (no difference in habitat use between predator and prey). This can also be understood as the likelihood that a predator will successfully attack a prey individual *after* they encounter each other, and is clearly much higher for the models that include habitat use. Line color corresponds to different models. Points show *realized* feeding rates of predators on each species in the experiment, based on the prey’s body size (position along the x-axis, denoted by vertical dashed lines) and the actual amount of time predators and prey spend in different habitats. This can be understood as the likelihood that a predator first encounters and then successfully attacks a prey individual. Points are colored according to model, the same as lines. The two aphid species are identified by larger points and darker vertical lines, while predator species (which become intraguild prey to other predators) are shown by smaller points and lighter colored lines. Body sizes were *R. padi* = 0.155mg, *A. pisum* = 0.67mg, *O. majusculus* = 0.58mg, *Bembidion* = 2.15mg, *Pardosa* = 18mg, and *C. septempunctata* = 37mg. Relationships here are shown for predator populations of 20 *Bembidion*, 2 *C. septempunctata*, 20 *O. majusculus* or 10 *Pardosa* individuals with an aphid population of 100 individuals. Despite having a higher  $a_0$  value, full model feeding-rate predictions are lower than the model with only habitat for all species except *C. septempunctata* due to predator-predator non-consumptive effects. Note the varying scales of the y-axis.

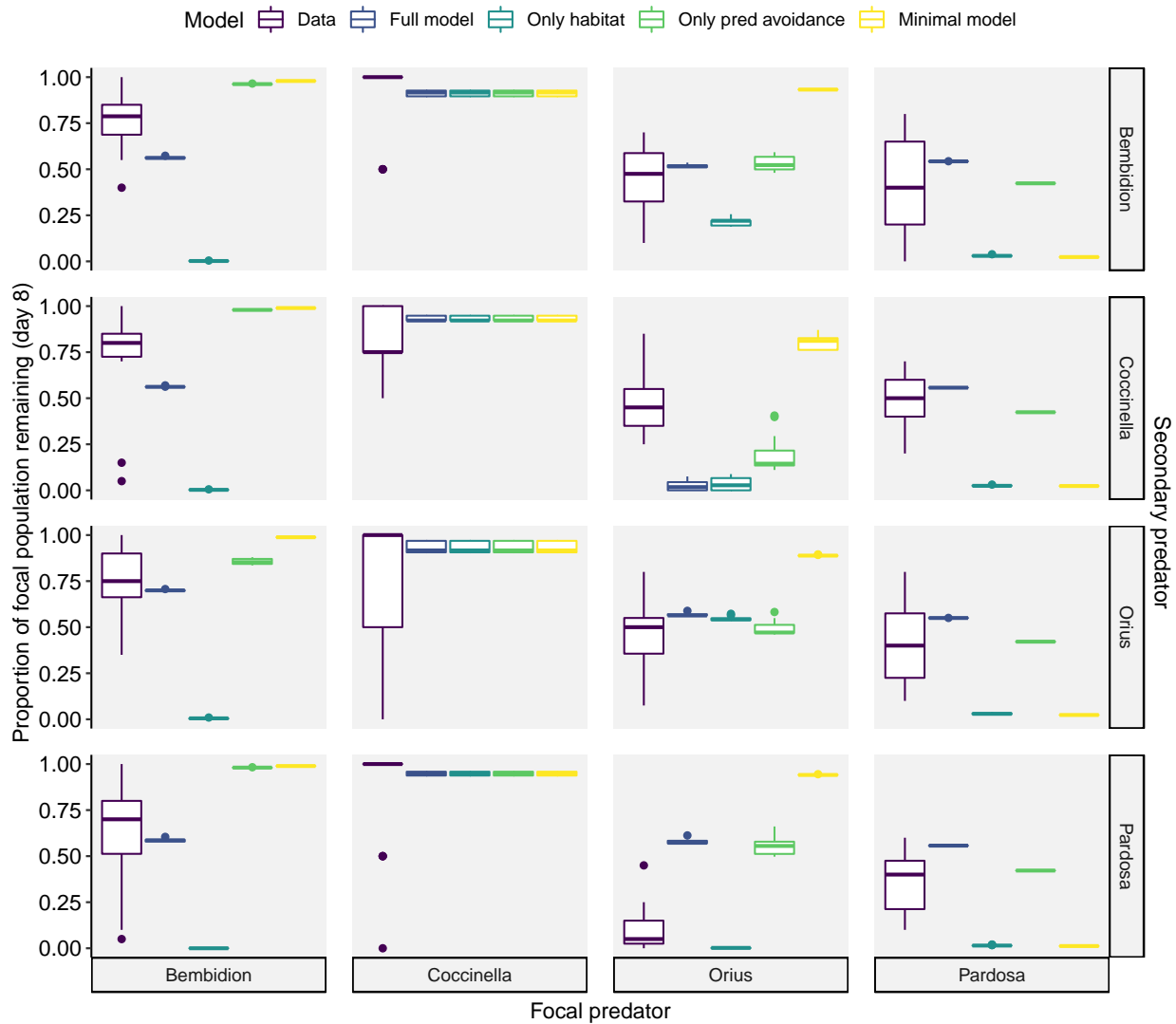


Figure 6: Experimental results (purple boxes) and model predictions for the proportion of each predator population remaining on the final day of the experiment. Data is grouped (facets along the right-hand y-axis) according to the second predator, i.e. the top right panel shows the proportion of the *Pardosa* population remaining at the end of the experiment when combined with *Bembidion*. Note that *C. septempunctata* was not modelled dynamically, so all model predictions are the same. Model predictions show different predictions based on different prey treatments and aphid population initial conditions.

442 atively similarly. First, the models matched aphid abundances to a comparable extent (JLS range  
443 from  $1.18 \times 10^8$  to  $1.38 \times 10^8$ , table 1), and second, the visual fit was good overall (figure 3). Third,  
444 all parameter values were — given our current knowledge — within reasonable ranges (Table 1).  
445 Fourth and similarly, processes (e.g. realized feeding rate), also appeared to be within reasonable  
446 ranges (figure 5). We found the largest difference between models in our final performance criteria,  
447 observed versus predicted final predator abundance. Generally the full model and the model with  
448 only predator avoidance gave the closest predictions to the data, but this depended on the predator  
449 species (figure 6). Overall, with regards to all five model performance criteria, the models per-  
450 formed well at describing aphid population dynamics for the studied predator-prey combinations.  
451 Nonetheless, estimated parameter values varied among models. This variation related especially to  
452 attack rate,  $a_0$ , and to optimal predator-prey body-size ratio,  $R_{opt,j}$ , with more variation in some  
453 species than others. This implies that we may ascribe different mechanisms to similar outcomes,  
454 whereas the models themselves are far from exchangeable and cannot all be correct at the same  
455 time.

456 That variation in performance between models is so small makes it difficult to confidently  
457 assess which model is the most accurate and appropriate. Two criteria point to the minimal  
458 model (without habitat use and predator-predator interference) as being the worst: it provided  
459 the worst fit (table 1), and it frequently over-estimated both aphid populations (figure 3) and  
460 predator populations. (Here, *Pardosa* formed an exception, as the the population of this predator  
461 was typically underestimated, figure 6). The full model and the model with non-consumptive  
462 predator-predator effects but no habitat use most often yielded the best predictions of the predator  
463 populations, while the model with habitat use but no non-consumptive predator-predator effects  
464 usually under-estimated the predator population (figure 6). Based on our observations during the  
465 experiment, the model with only habitat use also seems to predict the most biased  $R_{opt,j}$  values  
466 for both *Pardosa* and *O. majusculus*. We observed that both predator species will readily feed on  
467 individuals of a similar size as themselves, yet this model predicts that the optimal prey are 30 or  
468 997 times smaller than *Pardosa* and *O. majusculus* respectively. This would mean that the optimal  
469 prey mass of *O. majusculus* is 0.58 micrograms, approximately the mass of a grain of maize pollen

470 (Sheridan, 1982; Porter, 1981). The minimal model also predicted a relatively high  $R_{opt,j}$  value for  
471 *O. majusculus*, in contrast to experimental observations. These inaccuracies of the simpler models  
472 would hint that the full model, or the model with non-consumptive predator-predator effects but not  
473 habitat, is the best model, and that non-trophic predator-predator effects and possibly habitat use  
474 are important mechanisms driving trophic dynamics. Each line of evidence is, however, inconclusive,  
475 and the variation in performance between models is small. These considerations make it difficult to  
476 effectively determine which model is the most accurate and appropriate, and therefore to evaluate  
477 the importance (or lack thereof) of habitat use and non-consumptive predator-predator effects.

478 What our findings suggest is that we are yet to reach the point where we can fit ATN models  
479 to data for more diverse species, where traits other than body size also impact interactions, then  
480 say with any degree of confidence that this or that model is the correct one and the right one  
481 to base predictions upon. Each model points to different mechanisms behind the same outcomes.  
482 Models with high  $a_0$  values (those including habitat overlap) ascribe the realized feeding rate to a  
483 low frequency of encounters, but predict that when encounters do occur, then an attack is likely.  
484 Models with low  $a_0$  values, in contrast, suggest that encounters are common, but that when an  
485 encounter occurs, an attack is less likely. Similarly, differences in  $R_{opt,j}$  values attribute different  
486 mechanisms to similar predator-prey interaction dynamic outcomes. In both models that include  
487 non-consumptive predator-predator interference, the  $R_{opt}$  value of *O. majusculus* is close to the  
488 actual body-size ratio between *O. majusculus* and *A. pisum*. This suggests that *A. pisum* is a  
489 favoured prey of *O. majusculus*, and that the reason *O. majusculus* does not decimate *A. pisum*  
490 is cannibalism. Based on the  $R_{opt,j}$  value of *O. majusculus*, the feeding rate of this species is  
491 likely limited by non-consumptive predator-predator effects by conspecifics. In models without  
492 this mechanism (non-consumptive predator-predator effects), however, a similarly low value of  $R_{opt}$   
493 would indicate that *O. majusculus* should annihilate *A. pisum* — which it clearly does not. In  
494 these models, *O. majusculus* had a larger  $R_{opt}$  value, suggesting that the reason it did not feed on  
495 *A. pisum* so strongly was not due to the effects of other predators, but because *A. pisum* was not  
496 its optimal prey size. Given that we observed during initial feeding trials that *O. majusculus* would  
497 readily feed on *A. pisum*, it seems most likely that non-consumptive predator-predator effects are



498 in fact the limiting mechanism, not incompatible body sizes.

499 To help differentiate between models and improve prediction accuracy, we will need a closer  
500 understanding of which mechanisms are actually at play. Such insight is particularly important  
501 for making predictions of novel prey species, as we showed here (figure 3), as well as predictions  
502 of changes in existing species. If predator and prey change their habitat use to overlap more or  
503 less (as may be caused by loss of habitat for example, e.g. Carroll et al. (2019)), then a model  
504 where habitat use is important will have vastly different predictions from one where habitat use  
505 is not important. Improving our understanding of the mechanisms underlying food-web dynamics  
506 is therefore crucially important before we begin to make predictions. One way to resolve this  
507 conundrum, and to enable selection among models with a similar fit to the data, is to explore  
508 certain parameters more explicitly, and to estimate their value through supplementary experiments  
509 or investigations. Through such added data, we should be able to put limits on the ranges of  
510 parameter values, and thereby enable determination of which models are, in fact, most accurate.

511 In our case,  $R_{opt}$  would be a good candidate. This parameter absorbed a large portion of  
512 residual variance in our models, but should be relatively straight-forward to measure in itself (e.g.  
513 Brousseau et al., 2018). In our model fitting, we allowed  $R_{opt}$  to vary among predators to account  
514 for traits, such as feeding mode, that we had not included in the model. This, however, probably  
515 allowed too much flexibility. As a result, the predicted  $R_{opt}$  values varied significantly (in the case  
516 of *O. majusculus*,  $R_{opt}$  covered almost the entire allowed range, from 4 to 997). Feeding trials of  
517 each predator with otherwise similar prey covering a range of body sizes should enable a restriction  
518 of each predator's  $R_{opt}$  range. Such additional data would enable these parameters to be fixed or  
519 limited, aiding in discerning which models actually are most accurate.

520 Our study has shown that different model formulations and combinations of parameter values  
521 can produce similar outcomes in terms of predicted community dynamics. This implies that success  
522 in predicting observed trophic interaction strength and community dynamics of the ATN model so  
523 far (e.g. Schneider et al., 2012; Jonsson et al., 2018) must perhaps be interpreted cautiously. More  
524 specifically, although it is well established that body size does affect many interactions between  
525 species, the generality and exact nature and quantitative form of this relationship needs to be

526 further established. Only then can we rule out that a particular model formulation does not  
527 produce the ‘correct results’ (i.e. good fit to the data) for the wrong reasons. For this, the  
528 underlying biology of what drives foraging ecology (such as prey preference and feeding behaviour  
529 and efficiency) needs to be better understood, especially how it may vary among species and be  
530 affected by other traits than body size (e.g. Brousseau et al., 2018). Furthermore, we have shown  
531 that it is important to understand model limitations before applying a model to an experimental  
532 system, so that some aspect of the system is not more complex than that assumed by the model.  
533 In hindsight, we realise that some aspects of our model design might seem overambitious, if the  
534 goal were to discriminate between alternative model formulations and estimate parameter values  
535 for some universal relationship. More specifically, we explicitly aimed for predator diversity in  
536 our experimental design, leading us to include predators like *O. majusculus*, *Pardosa*, and *C.*  
537 *septempunctata* into the parameterization. As the foraging behavior of some of these predators  
538 differs from more traditional ‘grab-and-chew’ predators like *Bembidion*, it is probably too simplistic  
539 to describe trophic interaction strengths of all these predators using the same universal relationship  
540 and based on one trait (body size) only. On the other hand, our attempt at doing so has revealed  
541 the need to develop models that accommodate a diversity of foraging behaviour in predators.

542 In the current case, we will stop short of these proposed added steps, as the current study was  
543 explicitly designed to span the steps reported here (Laubmeier et al., 2018). Our intent is not to  
544 arrive at the final solution, but to point to the next step in the iterative process between theory  
545 and empirical insight.

## 546 **5.1 Conclusion**

547 In this study, we tested an approach explicitly developed in Laubmeier et al. (2018). Our aim  
548 was to arrive at an optimized design for generating empirical data to inform theoretical models.  
549 While carefully designed, the real-life implementation of the approach reveals the limitations of the  
550 information gained when trying to discriminate between different model versions. This emphasizes  
551 the challenges in developing the ATN model approach that need to be tackled. While we were  
552 able to arrive at a set of alternative parameter combinations plausible in the light of the data,

553 we were still unable to pinpoint which model is correct without locking down the exact value of  
554 one or several further parameters. The importance of doing so was revealed by our exercise of  
555 predicting the dynamics of a hypothetical novel species arriving in the system. Depending on  
556 which model is true, *Bembidion* and *Pardosa* may have the strongest effect on the novel prey,  
557 driving their population size nearly to zero, or they may have no impact and *C. septempunctata* and  
558 *O. majusculus* may be the more effective predators. Before we can convert the proposed models  
559 to predictive tools, we thus need to do more ground-work and conduct smaller experiments to  
560 estimate parameters such as optimal prey body size and attack rate scaling parameters — thereby  
561 gaining the resolution to select among multiple models. In conclusion, our study demonstrates how  
562 iterative cycling between theory, data and experiment may be needed to hone current insights into  
563 how traits affect food-web dynamics.

## 564 **6 Acknowledgements**

565 We owe a huge thank you to the many people who assisted in setting up this experiment, building  
566 cages, collecting predators, and counting so many aphids! Especial thanks to Gerard Malsher and  
567 Carol Högfeldt for their patient assistance and expertise. Thanks also to the research assistants Joe  
568 Bliss, Carly Miranda, Elin Ljunggren, Josefin Farrer-Sundberg, Maylis Moro, Josephine Biro, Lina  
569 Wu, Kevin Cestrieres, and Michail Artamonov. Thanks also to Anna Eklöf for valuable feedback  
570 on the manuscript.

571 We acknowledge funding from the Swedish University of Agricultural Sciences, Faculty of Nat-  
572 ural Resources and Agricultural Sciences (KLW, TR and RB), Swedish University of Agricultural  
573 Sciences, August T. Larsson guest researchers programme (RB), VR 2016-04580 (TR, RB, and  
574 TJ) and FORMAS 2016-01168 (TR and RB). AL and HTB were supported by grants AFOSR  
575 FA9550-18-1-0457 and NSF DMS-1246991.

## 576 **References**

577 Banks, H. and Tran, H. (2009). *Mathematical and Experimental Modeling of Physical and Biological*  
578 *Processes*. CRC Press, New York.

- 579 Banks, H. T., Hu, S., and Thompson, W. C. (2014). *Modeling and inverse problems in the presence*  
580 *of uncertainty*. CRC Press.
- 581 Berlow, E. L., Dunne, J. A., Martinez, N. D., Stark, P. B., Williams, R. J., and Brose, U. (2009).  
582 Simple prediction of interaction strengths in complex food webs. *Proceedings of the National*  
583 *Academy of Sciences*, 106(1):187–91.
- 584 Boit, A., Martinez, N. D., Williams, R. J., and Gaedke, U. (2012). Mechanistic theory and modelling  
585 of complex food-web dynamics in Lake Constance. *Ecology Letters*, 15(6):594–602.
- 586 Brose, U. (2010). Body-mass constraints on foraging behaviour determine population and food-web  
587 dynamics. *Functional Ecology*, 24(1):28–34.
- 588 Brousseau, P. M., Gravel, D., and Handa, I. T. (2018). Trait matching and phylogeny as predictors  
589 of predator–prey interactions involving ground beetles. *Functional Ecology*, 32(1):192–202.
- 590 Brown, J. H., Gillooly, J. F., Allen, A. P., Savage, V. M., and West, B. (2004). Toward a metabolic  
591 theory of ecology. *Ecology*, 85(7):1771–1789.
- 592 Carroll, G., Holsman, K. K., Brodie, S., Thorson, J. T., Hazen, E. L., Bograd, S. J., Haltuch, M. A.,  
593 Kotwicki, S., Samhouri, J., Spencer, P., Willis-Norton, E., and Selden, R. L. (2019). A review  
594 of methods for quantifying spatial predator–prey overlap. *Global Ecology and Biogeography*,  
595 28(11):1561–1577.
- 596 Chowell, G. (2017). Fitting dynamic models to epidemic outbreaks with quantified uncertainty:  
597 A primer for parameter uncertainty, identifiability, and forecasts. *Infectious Disease Modelling*,  
598 2(3):379—398.
- 599 Curtsdotter, A., Banks, H. T., Banks, J. E., Jonsson, M., Bommarco, R., Jonsson, T., Laubmeier,  
600 A. N., and Traugott, M. (2019). Ecosystem function in predator-prey food webs — confronting  
601 dynamic models with empirical data. *Journal of Animal Ecology*, 88:196–210.
- 602 Hudson, L. N. and Reuman, D. C. (2013). A cure for the plague of parameters: constraining models  
603 of complex population dynamics with allometries. *Proceedings of the Royal Society B: Biological*  
604 *Sciences*, 280(1770):20131901.
- 605 Jonsson, T., Kaartinen, R., Jonsson, M., and Bommarco, R. (2018). Predictive power of food web  
606 models based on body size decreases with trophic complexity. *Ecology Letters*, 21(5):702–712.

- 607 Kéfi, S., Berlow, E. L., Wieters, E. a., Navarrete, S. a., Petchey, O. L., Wood, S. a., Boit, A., Joppa,  
608 L. N., Lafferty, K. D., Williams, R. J., Martinez, N. D., Menge, B. a., Blanchette, C. a., Iles,  
609 A. C., and Brose, U. (2012). More than a meal... integrating non-feeding interactions into food  
610 webs. *Ecology Letters*, 15(4):291–300.
- 611 Knop, E., Zünd, J., and Sanders, D. (2014). Interactive prey and predator diversity effects drive  
612 consumption rates. *Oikos*, 123(10):1244–1249.
- 613 Laubmeier, A., Wootton, K., Banks, J., Bommarco, R., Curtsdotter, A., Jonsson, T., Roslin, T.,  
614 and Banks, H. (2018). From theory to experimental design—Quantifying a trait-based theory of  
615 predator-prey dynamics. *PLoS ONE*, 13(4).
- 616 Lima, S. L. (2002). Putting predators back into behavioral predator–prey interactions. *Trends in*  
617 *Ecology and Evolution*, 17(2):70–75.
- 618 Losey, J. E. and Denno, R. F. (1998). Positive predator-predator interactions: Enhanced predation  
619 rates and synergistic suppression of aphid populations. *Ecology*, 79(6):2143–2152.
- 620 Otto, S. B., Rall, B. C., and Brose, U. (2007). Allometric degree distributions facilitate food-web  
621 stability. *Nature*, 450(7173):1226–1229.
- 622 Peters, R. (1983). *The Ecological Implications of Body Size*. Cambridge University Press, New  
623 York.
- 624 Portalier, S. M. J., Fussmann, G. F., Loreau, M., and Cherif, M. (2019). The mechanics of predator-  
625 prey interactions : First principles of physics predict predator-prey size ratios. *Functional Ecology*,  
626 33(2):323–334.
- 627 Porter, E. K. (1981). Origins and Genetic Nonvariability of the Proteins Which Diffuse from Maize  
628 Pollen. *Environmental Health Perspectives*, 37:53–59.
- 629 Preisser, E. L., Orrock, J. L., and Schmitz, O. J. (2007). Predator hunting mode and habitat  
630 domain alter nonconsumptive effects in predator-prey interactions. *Ecology*, 88(11):2744–51.
- 631 Roslin, T. and Majaneva, S. (2016). The use of DNA barcodes in food web construction – terrestrial  
632 and aquatic ecologists unite! *Genome*, 628(April):603–628.
- 633 Schmitz, O. J. (2007). Predator diversity and trophic interactions. *Ecology*, 88(10):2415–2426.

- 634 Schneider, F. D., Scheu, S., and Brose, U. (2012). Body mass constraints on feeding rates determine  
635 the consequences of predator loss. *Ecology Letters*, 15(5):436–443.
- 636 Schneider, F. D., Scheu, S., and Brose, U. (2014). Corrigendum to Schneider, Scheu & Brose (2012)  
637 DOI: 10.1111/j.1461-0248.2012.01750.x. *Ecology Letters*, 17(10):1339–1340.
- 638 Sheridan, W. F. (1982). *Maize for biological research*. Plant Molecular Biology Association, Char-  
639 lottesville, VA, USA.
- 640 Sih, A., Englund, G., and Wooster, D. (1998). Emergent impacts of multiple predators on prey.  
641 *Trends in Ecology and Evolution*, 13(9):350–355.
- 642 Staudacher, K., Rennstam Rubbmark, O., Birkhofer, K., Malsher, G., Sint, D., Jonsson, M., and  
643 Traugott, M. (2018). Habitat heterogeneity induces rapid changes in the feeding behaviour of  
644 generalist arthropod predators. *Functional Ecology*, 32(3):1–11.
- 645 Terry, J. C. D., Bonsall, M. B., and Morris, R. J. (2020). Identifying important interaction modi-  
646 fications in ecological systems. *Oikos*, 129(2):147–157.
- 647 Terry, J. C. D., Morris, R. J., and Bonsall, M. B. (2017). Trophic interaction modifications: an  
648 empirical and theoretical framework. *Ecology Letters*, 20(10):1219–1230.
- 649 Thompson, R. M., Brose, U., Dunne, J. A., Hall, R. O., Hladysz, S., Kitching, R. L., Martinez,  
650 N. D., Rantala, H., Romanuk, T. N., Stouffer, D. B., and Tylianakis, J. M. (2012). Food  
651 webs: Reconciling the structure and function of biodiversity. *Trends in Ecology and Evolution*,  
652 27(12):689–697.
- 653 Yodzis, P. (1998). Local trophodynamics and the interaction of marine mammals and fisheries in  
654 the Benguela ecosystem. *Journal of Animal Ecology*, 67:635–658.