Supplementary Information for

Small mammal personalities generate context dependence in the seed dispersal mutualism.

Allison M. Brehm, Alessio Mortelliti

Allison M. Brehm
Email: allibrehm@gmail.com

This PDF file includes:

Supplemental Movie S1 Legend

Movie S1. How were interactions attributed to individual mice?

Appendix S1

Experiments in this study were approved by the University of Maine's Institutional Animal Care and Use Committee (IACUC numbers A2018-11-01 and A2015-11-02).

Supplementary Methods

Study site

The Penobscot Experimental Forest is a 1578 ha Forest Service experimental forest. Here, different forest units have been logged separately, managed with contrasting silvicultural treatments, and replicated twice in a randomized experimental design. Due to the contrasting silvicultural systems, these forest units differ greatly in the understory density, diameter of trees present, light levels, and quantity of downed woody material and snags. Approximately 25 ha of forest has been retained in two separate units and left unmanaged since the late 1800s to serve as reference. Common coniferous tree species include balsam fir (*Abies balsamea*), red spruce (*Picea rubens*), eastern white pine (*Pinus strobus*), and eastern hemlock (*Tsuga canadensis*), and deciduous species include red maple (*Acer rubrum*), American beech (*Fagus grandifolia*), Northern red oak (*Quercus rubra*), and birch and aspen (*Betula and Populus spp*.)

In this study, we performed small mammal trapping in six different areas of the experimental forest. Four areas were located in two contrasting silvicultural treatments (a uniform shelterwood cut using a two-stage overstory removal [treatment 1], and an irregular shelterwood cut using a two-stage overstory removal and retaining reserves, or trees from the older cohort [treatment 2]. We also used the two areas of unmanaged forest as reference sites. The area of the treatments used in this study was 12.8 ha on average (range: 9.49 – 19.39 ha).

Behavioral assays

Following capture an animal was transferred directly from the trap of capture into a clean, empty Longworth trap. This trap was then placed into a box sized 46 x 46 x 50 cm. To create a more natural environment, the inside of the box was painted brown with a small amount of debris (dead leaves and pine needles) placed on the floor. The box was placed underneath a tarp to control for light levels and canopy cover. A digital camera (Nikon CoolPix S3700) was mounted facing the opening of the Longworth trap, and the observer locked the trap door open before leaving the test area. A clear plexiglass lid was placed over the box to prevent escapes. After three minutes, the observer returned and ended the emergence test. Individuals were caught in a 4-liter plastic bag and then immediately released into the center of the open field arena.

A five-minute open-field test was performed in an arena of dimensions 46 x 46 x 50 cm, placed on a level platform with light levels controlled. After five minutes, the animal was caught in a 4-liter plastic bag and the observer suspended the bag into the open field test to control the visual surroundings. The observer then performed the handling bag test and measured the proportion of time that the individual spent immobile for 1 minute (referred to as handling time hereafter). Traps used for emergence tests and the open-field test box were cleaned thoroughly with 70% isopropyl alcohol and wiped with a dry cloth in-between all tests. After all three behavioral assays were complete (on average, this took approximately 10 minutes including time to transfer the animal between tests), the animal was processed (tagged and morphometrics measured) and promptly released at the point of capture.

Quantifying behavior

To quantify behavior from videotaped emergence and open-field tests, recordings were played back in the laboratory. For emergence tests, an observer recorded whether the animal emerged (defined as all four feet having left the Longworth trap), the latency to emerge, and the total time spent at the end of the Longworth tunnel before emerging. When an individual did not emerge from the test after the three-minute cutoff, the latency to emerge was set to 1.25x the maximum test length (this occurred in 156 tests out of 1164 total). Open-field tests were analyzed using the behavioral tracking software ANY-maze © (version 5.1; Stoelting CO, USA). The following behavioral variables were obtained from the behavioral assays: handling time (the number of seconds immobile in a handling test; note that the term handling time should not be confused with a term sometimes used in ecological literature to indicate the time spent handling, processing, and consuming food items), latency to emerge and time at tunnel end (from the emergence test), mean speed, rear rate, proportion of time grooming, and proportion of time in the center (from the open-field test). See Table S3 (modified from [1]) for a complete list of the behaviors used, their description and interpretation, and supporting sources.

Seed experiments – further details

Seeds used in this study included Northern red oak (*Quercus rubra*), Eastern white pine (*Pinus strobus*), and American beech (*Fagus grandifolia*). These three seed species were chosen because they are present at our study sites, they represent a variety of sizes (while being large enough to track using the fluorescent powder method), and they represent seeds whose primary dispersal method includes both wind and animal dispersal. Seeds were purchased from F.W. Schumacher Co., Inc. and cold stratified over the winter.

Presentation stations were placed on the forest floor and we mounted a trail camera (Reconyx XR6 Ultrafire) 1.25 to 1.5 meters above the station. Cameras were formatted to take 30 second 1080P HD videos (at 30 fps) and an 8-megapixel photograph prior to the start of the video and once per hour. The cameras were set to the shortest delay between videos (1 s). To identify individual small mammals, we used a permanent radio frequency identification (RFID) reader to scan and identify individuals marked with PIT tags (RFIDLOG dual animal tag rfid data logger). An antenna (Priority1 rfidcoil – 160a) was mounted to each presentation station and positioned to surround the seed presentation trays (Figure 4D). These antennas were built to operate at a frequency of 134.2 kHz for optimal reading of PIT tags used in this study. Records were stored automatically on an SD memory card along with a time stamp of the detection.

To allow for the relocation of seeds removed from the stations, we dusted the vinyl floor tiles with UV fluorescent tracking powder (TechnoGlow; yellow, firehouse orange, or magenta). These pigments are made of earth friendly materials and shine brightly under UV light. Small mammals would leave distinct fluorescent trails when dispersing seeds from the seed stations (2). To permit located caches to be matched to the individual disperser, each oak and beech seed was painted with non-toxic, UV fluorescent paint (Neon Glow, ASTMD-4236 certified). Each seed at a site was painted a different color, and the location of each color on the presentation tray (randomly chosen and rotated between stations), as well as the mass of beech and oak seeds were recorded when stations were set. When a seed was found, the color allowed us to match the cache to the individual small mammal observed dispersing it in trail camera videos. For white pines, five seeds were placed in each of the six wells and all five seeds in a well were painted the same color. Since several white pine seeds can be dispersed at a time, caches were identified by matching the count of each different color seed found within the cache to the seeds seen taken in videos (Figure 4F, G). Since all experiments were performed in complete darkness, seed color should not have influenced the initial decision to consume or cache the seed.

On average we ran four or five experiment sites at one time and performed the experiment at a trapping grid for 3-4 days. In total, we placed 52 seed sites in September (eight sites at two grids and nine sites at four grids) and 51 in October (eight sites at three grids and nine sites at three grids).

After completion of the seed experiments, we simulated five caches outside of four of our trapping grids. Half of these caches contained one painted beech seed, and the other half contained one unpainted beech seed. We monitored these caches using trail cameras over 11 days and confirmed using a two-sample t-test that the chance of a cache being pilfered or recovered is not influenced by our methods of tagging seeds with paint (p-value comparing the chance of painted vs. unpainted seeds being pilfered within 24 hours = 0.178; p-value comparing cache longevity of painted vs. unpainted seeds = 0.219). We acknowledge that we did not specifically test whether the different colors of paint would have an effect on an individual's choice to consume or cache a seed, but given the fact that all experiments were performed in the complete darkness of the forest, we have no reason to believe this should have occurred.

Playback of seed videos

From the trail camera videos obtained at the seed presentation stations, we recorded each visit to the seed stations, noting the identity of each individual and the nature of their interaction (i.e., which seed species they chose, whether they chose to consume it at the site or remove it, and the color of the seed they chose). We deemed behavior as an "interaction" with a seed if the individual intentionally made contact with the seed. Individual identification was confirmed using both the time-stamped RFID reads and the unique haircut seen in the trail cam videos.

Repeatability analyses

We calculated the adjusted repeatability (3) and associated 95% confidence intervals for seven key behaviors performed in the standardized assays using the 'rptR' package in R (4). For this analysis we used only individuals who had two or more repeated observations. Repeatability can be defined as the proportion of the total phenotypic variation that can be attributed to individual differences (5–7). Practically, this means that repeatability equals the between-individual variance divided by the total phenotypic variance (between-individual variance + within-individual variance). The random intercept of individual ID was included in the models and fixed effects included sex, body condition index (calculated using the scaled-mass index [6]), silvicultural treatment, and trapping session. In all models, we used 1000 parametric bootstraps and 100 permutations. We assessed normality by visually inspecting Q-Q plots and by plotting the fitted values against the residual values and used Box-Cox transformations on the response variable to approach normality when necessary (9, 10). We considered any behavioral trait with a 95% CI for repeatability that excluded zero to be a personality trait (6), but we emphasize that this classification as significantly repeatable does not say anything about the strength of repeatability.

After we confirmed that our behavioral variables were significantly repeatable, we took steps to account for the within-individual variability that occurs between repeated behavioral observations (consistent with methods used by [11, 12]). We calculated an individual's best linear unbiased predictor (BLUP) for each of the behavioral variables after controlling for the variation due to sex, body condition, forestry treatment, and trapping session. We then calculated an individual's mean BLUP for each of the repeatable behavioral variables (estimated over 1000 simulations using the *sim* function in package arm [13]). We then used the mean BLUPs as the personality covariates in our linear models and mixed-effects models.

Further, before moving forward with model selection, we screened all seven behavioral variables for correlation before analysis (using R<0.7 as a threshold, consistent with [14]). See Table S6 for pairwise correlations between all behavioral variables used in model selection.

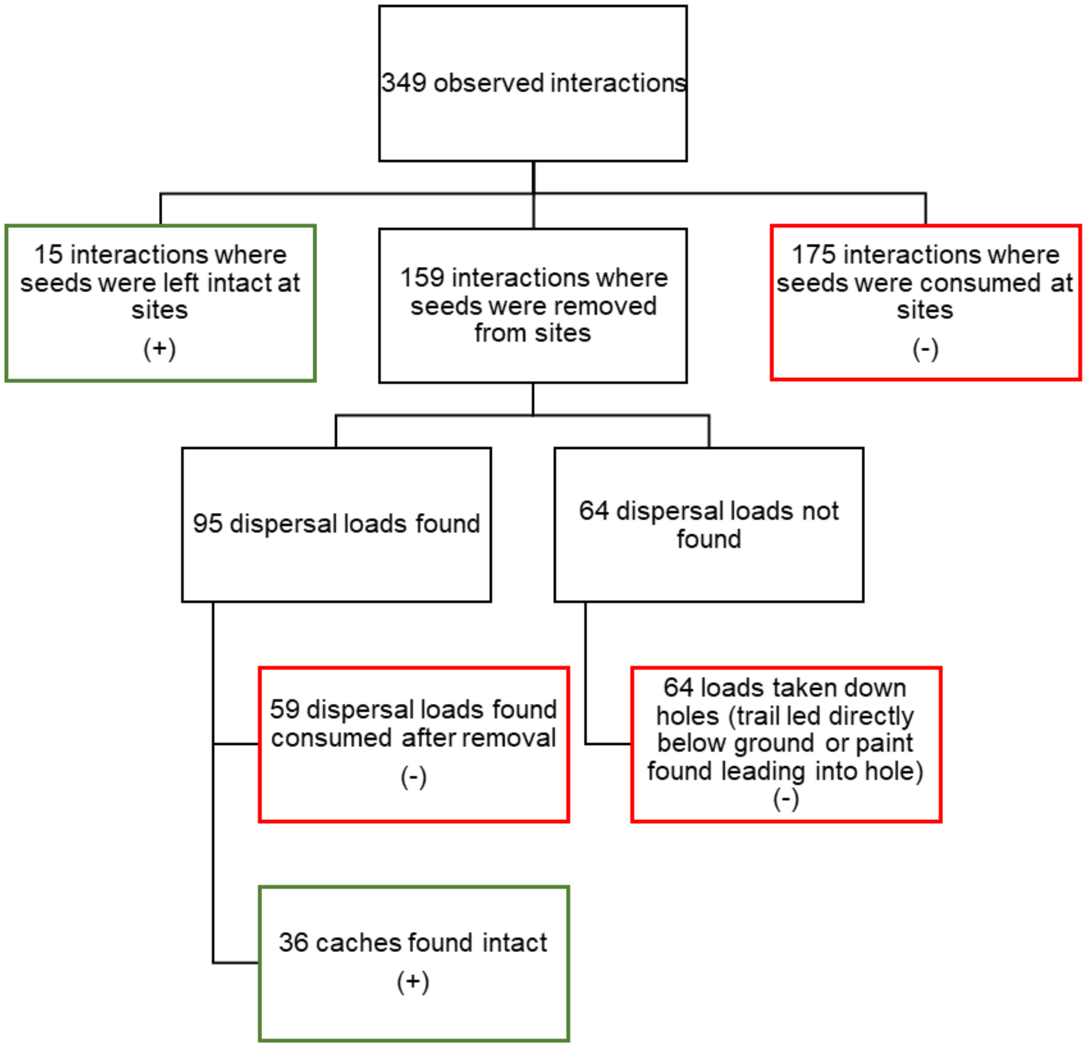

Figure S1. White pine (*Pinus strobus*) interaction diagram showing the number of interactions made by deer mice (*Peromyscus maniculatus*) classified as positive (outlined in green) or negative (outlined in red) and the nature of these interactions.

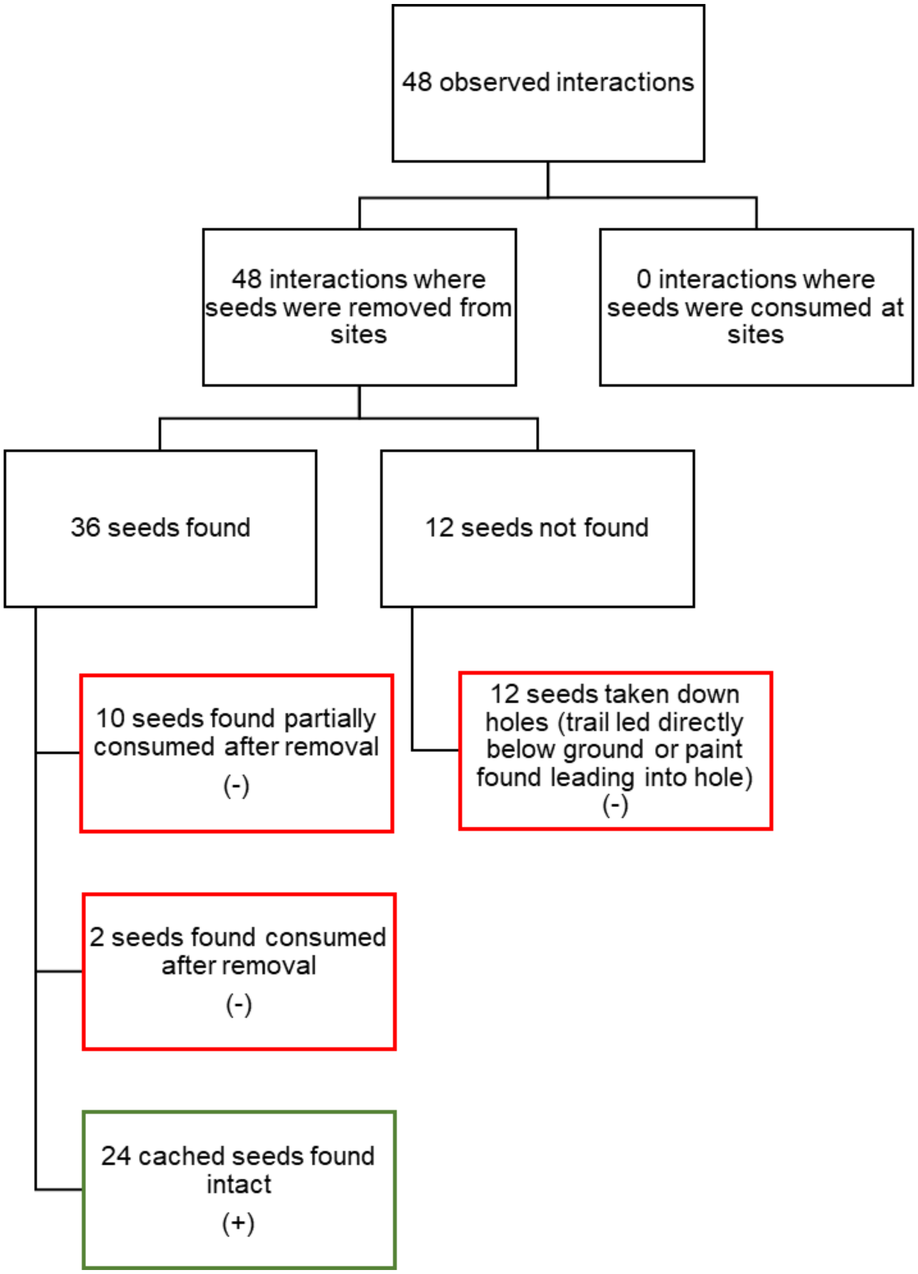

Figure S2. Red oak (*Quercus rubra*) interaction diagram showing the number of interactions made by deer mice (*Peromyscus maniculatus*) classified as positive (outlined in green) or negative (outlined in red) and the nature of these interactions.

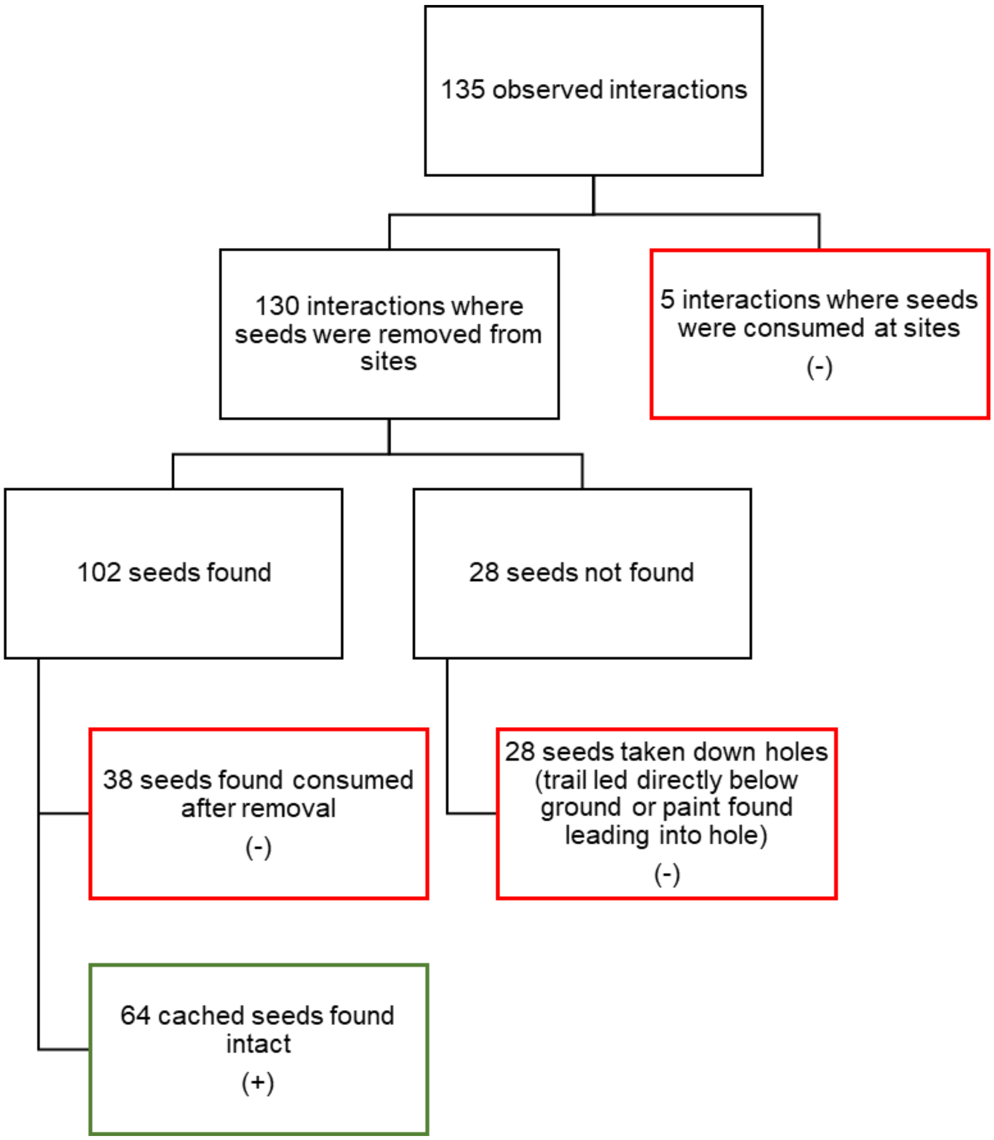

Figure S3. American beech (*Fagus grandifolia*) interaction diagram showing the number of interactions made by deer mice (*Peromyscus maniculatus*) classified as positive (outlined in green) or negative (outlined in red) and the nature of these interactions.

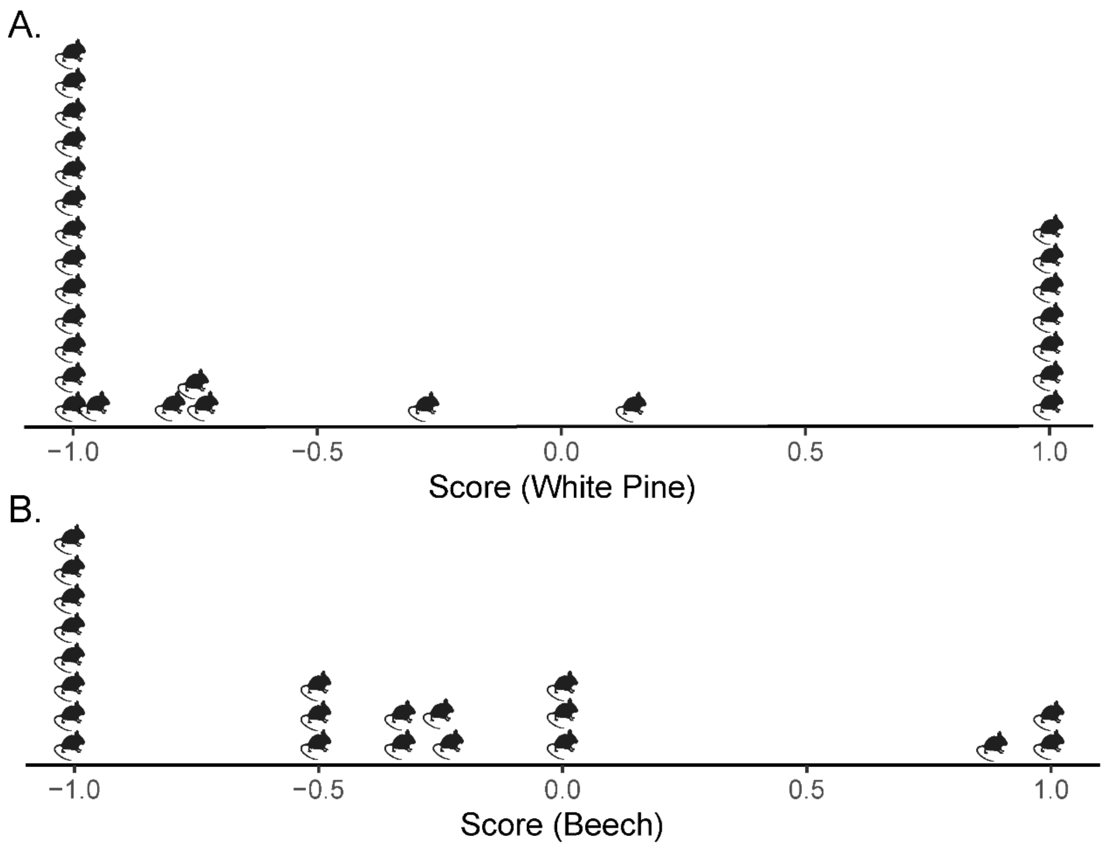

Figure S4. Individual scores for *Peromyscus maniculatus* along the antagonism-mutualism continuum for (A) eastern white pine (*Pinus strobus;* mean score = -0.36) and (B) American beech (*Fagus grandifolia;* mean score = -0.37).

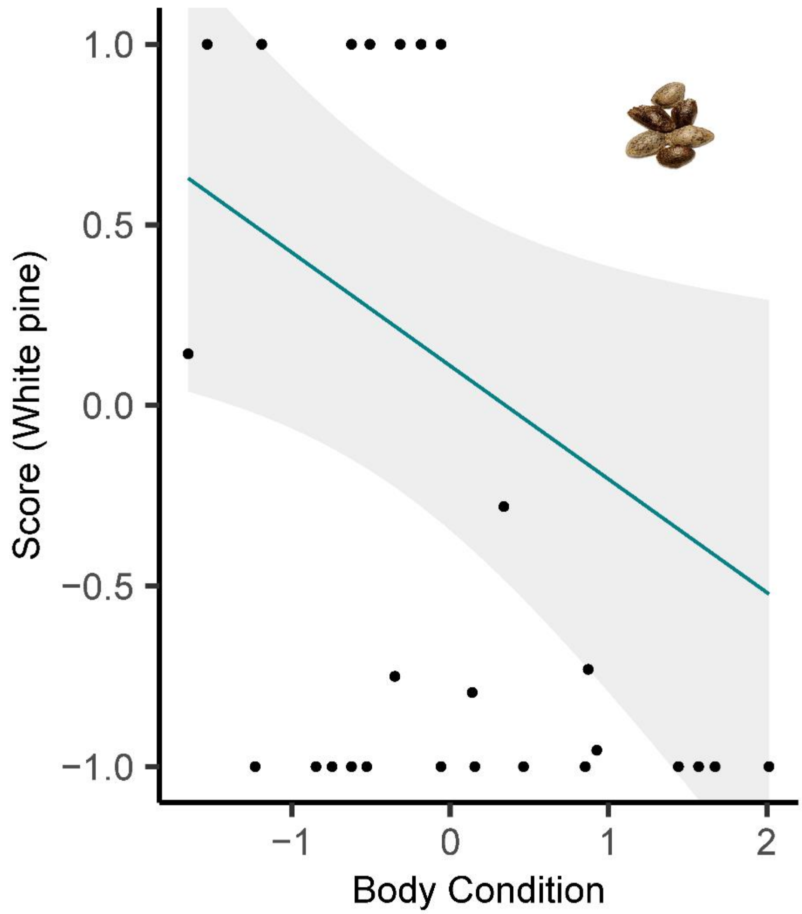

Figure S5. Predicted relationships (and 95% CIs) between an index of body condition and an individual's score along the antagonism-mutualism continuum. Individual *Peromyscus maniculatus* with a higher body condition index have lower (more antagonistic) scores for eastern white pine (*Pinus strobus;* $\beta = -0.31 \pm 0.15$ SE). Data points represent observed values (one point per individual). Body condition index represents a scaled-mass index and values were z-standardized. Relationship between body condition index and white pine score is shown for the treatment 2 forest type.

Table S1. Interactions used to calculate an individual's score along the antagonist/mutualist continuum.

Action	**Positive or Negative**
Seed consumed at the site	Negative
Seed removed from site and cached intact	Positive
Seed removed from site and then consumed	Negative
Seed removed from site and taken down a hole	Negative
*Seed left intact at the site after interaction	*Positive

The score can be quantified as the proportion of interactions that are positive minus the proportion of interactions that are negative.

* This behavior was observed regularly for white pine seeds only. Further, white pine seeds are primarily wind dispersed, so for this species a non-consumption/rejection can be considered positive.

Table S2. Repeatability estimates for target behaviors measured in three behavioral assays (open field, handling bag, and emergence tests) in deer mice (*Peromyscus maniculatus*).

Behavior	Mean	Range	RPT	(95% CI)	Observations	Individuals
Mean speed (m/sec)	0.09	(0, 0.29)	**0.447**	**(0.371, 0.515)**	815	300
Rear rate (rears/sec)	0.19	(0, 0.65)	**0.361**	**(0.285, 0.445)**	819	301
Prop time groom[1]	0.20	(0, 0.97)	**0.389**	**(0.312, 0.467)**	818	301
Prop time in center[2]	0.02	(0, 0.53)	**0.280**	**(0.211, 0.368)**	816	301
Handling time (sec)	12.82	(0, 60)	**0.342**	**(0.268, 0.426)**	749	268
Latency to emerge (sec)	47.44	(0, 225)	**0.300**	**(0.221, 0.386)**	689	250
Time at tunnel end (sec)	4.67	(0, 180)	**0.191**	**(0.104, 0.295)**	617	229

[1]Proportion of time spent grooming in the open-field test.
[2]Proportion of time spent in center portion of open-field arena.

Table S3. Repeatable behavioral traits from three standard behavioral assays performed on deer mice (*Peromyscus maniculatus*). This table provides the description of the behavior, the personality trait that this behavior represents, the assay that each trait was obtained from, and how values were interpreted. Included is a non-exhaustive list of citations supporting trait interpretation. This table was adapted from (1, 15).

Behavior	Personality trait	Behavioral assay	Description	Interpretation	Sources
Mean speed	Activity	Open field	Mean speed in the open field test in (m/s). Calculated by dividing the total distance traveled in the test by the test duration	Higher values indicate greater activity	(16, 17)
Rear rate	Activity and exploration	Open field	Rate of rearing (rears/s). Rearing is defined as forelegs leaving the arena floor	Higher values indicate greater activity and exploration	(18–21)
Proportion time grooming	Anxiety/ stress	Open field	Proportion of test duration spent grooming	Higher values indicate lower anxiety and better coping	(18, 22, 23)
Proportion time center	Bold/timid	Open field	Proportion of test duration spent in the center portion of the arena	Higher values indicate more boldness/less timidness	(17, 18, 24–27)
Handling time	Docility	Handling bag	Total number of seconds of inactivity during a 1-minute handling bag test	Higher values indicate more docility/less defensive aggression	(20, 28–30)
Latency to emerge	Bold/timid	Emergence	Latency (in seconds) to emerge from trap in the emergence test. An animal was considered to have emerged when all four feet left the trap tunnel	Higher values indicate more timidness/less boldness	(16, 17, 31)
Time at end of tunnel	Bold/timid	Emergence	Total number of seconds spent at the end of the tunnel before emerging	Higher values indicate more timidness/less boldness	(1, 32)

Table S4. Results from linear regression predicting scores along the predator-mutualist continuum in the deer mouse (*Peromyscus maniculatus*) for three target seed species: eastern white pine (*Pinus strobus*), red oak (*Quercus rubra*), and American beech (*Fagus grandifolia*). Models within 2.0 ΔAICc[1] of the top model are shown.

Species	**Model**	**ΔAICc[1]**	**Adj. R squared[2]**
Pinus strobus	Body condition + Forest type + Time at tunnel end	0	0.4
	Body condition[3] (-0.31 ± 0.15 SE) + Forest type	0.5	0.34
Quercus rubra	**Time at tunnel end (-0.44 ± 0.15 SE)**	0	0.29
Fagus grandifolia	Number of interactions	0	0.11
	Proportion of time in center	0.37	0.09
	Null	0.67	0
	Forest type	1.76	0.11
	Body condition	1.97	0.02

Models for *P. strobus* include scores from 26 individuals, models for *Q. rubra* include scores from 19 individuals, and models for *F. grandifolia* include scores from 21 individuals. Effect sizes and standard errors shown in parentheses for top models.
Behavioral variables in models above represent an individual's average BLUP (i.e. best linear unbiased predictor) estimated over 1000 simulations.
[1] Delta Akaike information criterion corrected for small sample sizes
[2] Adjusted coefficient of determination
[3] Body condition index calculated using the scaled-mass index (8)

Table S5. Results for logistic mixed-models predicting the probability of a positive seed interaction in the deer mouse (*Peromyscus maniculatus*). Models within 2.0 ΔAICc[1] of the top model are shown.

Species	Model*	ΔAICc[1]	Cond. R squared[2]
Pinus strobus	**Forest type + Body condition[3] (-0.43 ± 0.20 SE) + Time in center (-1.24 ± 0.41 SE)**	0	0.58
Quercus rubra	Sex	0	0.06
	Null	0.02	0
	Time at tunnel end	0.04	0.06
	Body mass	0.67	0.04
	Rear rate	1.81	0.01
	Seed mass	1.98	0.01
Fagus grandifolia	**Time in center (-0.89 ± 0.33 SE)**	0	0.30

Models for *P. strobus* include 349 observations from 26 individuals, models for *Q. rubra* include 48 observations from 19 individuals, and models for *F. grandifolia* include 135 observations from 21 individuals. Effect sizes and standard errors shown in parentheses for top models.

[1] Delta Akaike information criterion corrected for small sample sizes

[2] Conditional (theoretical) coefficient of determination calculated using the *r.squaredGLMM* command in the MuMIn package in R

[3] Body condition index calculated using the scaled-mass index (8)

Table S6. Pairwise correlations between all behavioral variables used in model selection and an individual's mean body condition index (averaged scaled-mass index), and mean body mass. Behavioral variables shown are the mean BLUP (best linear unbiased predictor) averaged over 1000 simulations.

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
1 - Mean speed	1	0.64	-0.62	0.18	-0.24	-0.33	-0.13	-0.06	-0.10
2 - Rear rate	0.64	1	-0.54	0.09	-0.18	-0.27	-0.12	-0.08	-0.13
3 - Prop. time grooming	-0.62	-0.54	1	-0.16	-0.05	0.11	-0.08	0.07	0.10
4 - Prop. time in center	0.18	0.09	-0.16	1	-0.04	-0.01	0	0.02	0.09
5 - Handling	-0.24	-0.18	-0.05	-0.04	1	0.31	0.22	0	0.05
6 - Latency to emerge	-0.33	-0.27	0.11	-0.01	0.31	1	0.41	0.02	0.09
7 - Time at tunnel end	-0.13	-0.12	-0.08	0	0.22	0.41	1	0.01	0.01
8 - Mean body condition	-0.06	-0.08	0.07	0.02	0	0.02	0.01	1	0.41
9 - Mean body mass	-0.10	-0.13	0.10	0.09	0.05	0.09	0.01	0.41	1

SI References

1. A. M. Brehm, A. Mortelliti, G. A. Maynard, J. Zydlewski, Land-use change and the ecological consequences of personality in small mammals. *Ecol. Lett.* **22**, 1387–1395 (2019).
2. W. S. Longland, C. Clements, Use of fluorescent pigments in studies of seed caching by rodents. *J. Mammal.* **76**, 1260–1266 (1995).
3. A. J. Wilson, How should we interpret estimates of individual repeatability? *Evol. Lett.* **2**, 4–8 (2018).
4. M. A. Stoffel, S. Nakagawa, H. Schielzeth, rptR: Repeatability estimation and variance decomposition by generalized linear mixed-effects models. *Methods Ecol. Evol.* **8**, 1639–1644 (2017).
5. N. J. Dingemanse, A. J. N. Kazem, D. Re, J. Wright, Behavioural reaction norms: animal personality meets individual plasticity. *Trends Ecol. Evol.* **25**, 81–89 (2009).
6. S. Nakagawa, H. Schielzeth, Repeatability for Gaussian and non-Gaussian data: A practical guide for biologists. *Biol. Rev.* **85**, 935–956 (2010).
7. N. J. Dingemanse, N. A. Dochtermann, Quantifying individual variation in behaviour: mixed-effect modelling approaches. *J. Anim. Ecol.* **82**, 39–54 (2013).
8. J. Peig, A. J. Green, New perspectives for estimating body condition from mass/length data: The scaled mass index as an alternative method. *Oikos* **118**, 1883–1891 (2009).
9. G. E. P. Box, D. R. Cox, An analysis of transformations. *J. R. Stat. Soc. Ser. B* **26**, 211–243 (1964).
10. Y. Yang, O. F. Christensen, D. Sorensen, Analysis of a genetically structured variance heterogeneity model using the Box-Cox transformation. *Genet. Res. (Camb).* **93**, 33–46 (2011).
11. N. J. Dingemanse, M. Moiron, Y. G. Araya-Ajoy, A. Mouchet, R. N. Abbey-Lee, Individual variation in age-dependent reproduction: fast explorers live fast but senesce young? *J. Anim. Ecol.* **89**, 601–613 (2020).
12. E. Gharnit, P. Bergeron, D. Garant, D. Réale, Exploration profiles drive activity patterns and temporal niche specialization in a wild rodent. *Behav. Ecol.* **31**, 772–783 (2020).
13. A. Gelman, Y.-S. Su, arm: Data analysis using regression and multilevel/ hierarchical models. *R Packag. version 1.10-1* (2018).
14. C. F. Dormann, *et al.*, Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography (Cop.).* **36**, 027–046 (2013).
15. A. M. Brehm, A. Mortelliti, Land-use change alters associations between personality and microhabitat selection. *Ecol. Appl.*, in press (2021).
16. A. J. Carter, W. E. Feeney, H. H. Marshall, G. Cowlishaw, R. Heinsohn, Animal personality: what are behavioural ecologists measuring? *Biol. Rev.* **88**, 465–475 (2013).
17. G. Gracceva, *et al.*, Turning shy on a winter's day: effects of season on personality and stress response in *Microtus arvalis*. *Ethology* **120**, 753–767 (2014).
18. E. Choleris, A. W. Thomas, M. Kavaliers, F. S. Prato, A detailed ethological analysis of the mouse open field test: effects of diazepam, chlordiazepoxide and an extremely low frequency pulsed magnetic field. *Neurosci. Biobehav. Rev.* **25**, 235–260 (2001).
19. L. Prut, C. Belzung, The open field as a paradigm to measure the effects of drugs on anxiety-like behaviors: a review. *Eur. J. Pharmacol.* **463**, 3–33 (2003).
20. J. G. A. Martin, D. Réale, Temperament, risk assessment and habituation to novelty in eastern chipmunks, *Tamias striatus*. *Anim. Behav.* **75**, 309–318 (2008).
21. S. Tanaka, J. W. Young, A. L. Halberstadt, V. L. Masten, M. A. Geyer, Four factors underlying mouse behavior in an open field. *Behav. Brain Res.* **233**, 55–61 (2012).
22. A. V. Kalueff, *et al.*, Neurobiology of rodent self-grooming and its value for translational neuroscience. *Nat. Rev. Neurosci.* **17**, 45–59 (2016).
23. A. Fernández-Teruel, C. Estanislau, Meanings of self-grooming depend on an inverted U-shaped function with aversiveness. *Nat. Rev. Neurosci.* **17**, 591 (2016).
24. S. Barnett, *The rat - a study in behavior* (ANU Press, 1976).
25. D. Treit, M. Fundytus, A. M. Fundytus, Thigmotaxis as a test for anxiolytic activity in rats. *Pharmacol. Biochem. Behav.* **31**, 959–962 (1989).

26. A. Ramos, O. Berton, P. Mormède, F. Chaouloff, A multiple-test study of anxiety-related behaviours in six inbred rat strains. *Behav. Brain Res.* **85**, 57–69 (1997).
27. J. A. Eccard, A. Herde, Consistency in boldness, activity and exploration at different stages of life. *BMC Ecol.* **103**, 61–68 (2013).
28. A. K. Boon, D. Réale, S. Boutin, The interaction between personality, offspring fitness and food abundance in North American red squirrels. *Ecol. Lett.* **10**, 1094–1104 (2007).
29. P. O. Montiglio, D. Garant, F. Pelletier, D. Réale, Personality differences are related to long-term stress reactivity in a population of wild eastern chipmunks, *Tamias striatus*. *Anim. Behav.* **84**, 1071–1079 (2012).
30. R. W. Taylor, S. Boutin, M. M. Humphries, A. G. Mcadam, Selection on female behaviour fluctuates with offspring environment. *J. Evol. Biol.* **27**, 2308–2321 (2014).
31. C. Brown, V. A. Braithwaite, Effects of predation pressure on the cognitive ability of the poeciliid *Brachyraphis episcopi*. *Behav. Ecol.* **16**, 482–487 (2004).
32. A. M. Brehm, S. Tironi, A. Mortelliti, Effects of trap confinement on personality measurements in two terrestrial rodents. *bioRxiv* (2019) https:/doi.org/10.1101/723403.