S1 Supplementary methods

S1.1 Systematic coding protocol

The following coding protocol was used to systematically code relevant information from each of the papers included in the review. In each study there was one or more intervention; the ecosystem health outcomes of each intervention were reported with one or more ecosystem health metrics. All coding was done in Excel.

1. **Paper**
   Free text: Paper title

2. **Intervention ID**
   Free text: same ID as used in the coding for Chausson, Turner et al. (2020)

3. **Ecosystem health metric ID**
   Free text: 1a, 1b, 1.1a, 1.1b etc. where '1' refers to the first paper, '.1' refers to the second intervention of the first paper, and 'a' refers to the first ecosystem health metric of that intervention.

   - We made a new row for each new biodiversity metric (a, b, c etc.)
   - We made a new row for each new intervention within a paper (1, 1.1, 1.2 etc.)

4. **Intervention description**
   Free text: we described the intervention, e.g. 'creation of trenches to aid plant establishment in degraded rangeland'. We gave a high level of detail and copy and pasted text from paper if appropriate.

5. **Intervention type**
   Free text: same intervention types as used in the coding for Chausson, Turner et al. (2020)

6. **Habitat type**
   Free text: same habitat types as used in the coding for Chausson, Turner et al. (2020)

7. **Broad habitat type**
   Selected: Forest/other terrestrial/coastal/marine/freshwater/terrestrial combination/terrestrial-aquatic/coastal combination

8. **Country**
   Free text: country where the intervention took place

9. **Geographical region**
   Selected: Africa/Asia + Pacific/Europe/Latin America + Caribbean/North America/West Asia

10. **Assessment of potential bias**
    Conflict of interest?

    - Selected: no conflict declared/conflict declared

    Author-intervention-funding relationship?

    - Selected: yes/no/unclear
    - Selected from:

    - Field experiment implemented by author(s)
    - Author(s) associated with intervention
    - Author(s) associated with intervention funding
    - Same funding body for research & intervention
    - Field experiment implemented by author(s) & intervention associated with research funding
    - Author(s) associated with intervention & intervention funding

    Note: this was coded separately for different interventions from the same study when appropriate.

11. **Ecosystem health outcome data type**

    The type of evidence provided for the effect of the intervention on biodiversity. Marked X for all that apply:

    - **Social:** data collected through social science methods such as interviews and focus groups, as well as non-numerical data collected by other means, such as statements by authors without defined methods. This also includes anecdotal data, such as statements of change by authors – these studies will have 'absent' or 'poor detail' in their methodology (see question 7).
    - **Biophysical:** numerical data on an element of ecosystem health, including binary outcomes e.g. presence/absence data, economic assessments of the effects of changes in biophysical variables (e.g. effects of drought on agricultural productivity used to infer the prevalence of drought), and quantitative variables without supporting data, but that were not collected with social methods (e.g. increase in fragmentation, from remote sensing or unclear methods).
    - **Secondary:** authors cited evidence from another source and *did not conduct their own analysis*. In this case, pasted the link to or citation of the source(s) in the next column. This does not include analyses which include some secondary data e.g. satellite imagery sourced online but analysed by authors.

12. **Ecosystem health outcome data quality**

    Methodology presence and level of detail. Selected:

    - Detailed - sufficient detail to repeat, or close to sufficient
    - Poor detail - insufficient detail to repeat, with major uncertainties or ambiguities
    - Absent - no methods given
    - Secondary data so NA - when secondary data used (don't assess the methodology in the secondary source)

    Counterfactual or baseline. Marked X:

    - Control: use of an area which did not undergo the intervention to estimate what would have happened in the absence of the intervention - the control area must have been monitored in the same way as the test/intervention area. E.g. before-after-control-impact method. Do not count e.g. space for time substitution, or historical baselines.
    - Modelled scenario: use of modelling to estimate what would have happened in the absence of the intervention.
    - Historical baseline: use of historical data from the area of study as a proxy for what would have happened in the absence of the intervention.
    - Other: enter brief description, e.g. space for time substitution, meeting set of criteria.
    - None: if no form of counterfactual or baseline is used.

Intervention comparison. Marked X if two or more interventions were compared (which may be instead of, or in addition to, using a no-intervention scenario as a baseline).

Threshold. Marked X if a threshold was used to assess success with respect to an ecological parameter.

13. Ecosystem health metric type
Identify which ecosystem health metric types were used (i.e. were mentioned by name or met the definition given in Table 2, in the methods of the main text).

We made a new row for each ecosystem health metric, and answered each of the following questions for individual biodiversity metrics.

14. Broad ecosystem health metric category
Assigned according to definitions in Table 2.

15. Ecosystem health metric details

- Free text entry: noted the name of the specific metric used (e.g. Shannon-Weiner species diversity index), or described it concisely (e.g. percentage canopy cover). This distinguishes between metrics which fall within the same metric type. This column was not always filled in since the same information can be found in the outcome details column.

16. Taxa:

- **Kingdoms** of taxa studied:
    - Selected from: Plants, Animals, Microbes (general), Fungi, Protists, Archaea, Eubacteria, NA (for e.g. sediment accretion).
    - Chose 'microbes (general)' when microbial kingdom was not specified.
- **Broad taxa:**
    - Selected from: Animals/plants/plants, animals & other/microorganisms, fungi & other
- **Animal class:**
    - Free text: recorded class of any animals studied, or give the paraphyletic group if more appropriate e.g. pisces, if the study was on unspecified fish classes.
- **Plant info**:
    - Free text: for any plants studied, listed from: trees, grasses, other plants.
    - Note: other plants can include unidentified plants (that might be trees and grasses).
- **Taxa description**:
    - Free text: recorded a description of taxa (or equivalent) at the lowest level recorded by the authors and any additional description e.g. fodder species. For surveys of species richness, did not list all species recorded.
    - This also included e.g. 'streams' where e.g stream length was recorded for habitat extent.
- **Includes a non-living component:**
    - Mark X if the metric involved assessing a non-living component of the ecosystem such as water body extent, water quality, or accretion rate.
- **Native composition:**
    - Selected: native/non-native/mixed/native assisted migrant/native invasive/unclear

- Where a 'native assisted migrant' includes migration of the same species from a different country; and 'native invasive' is a native species which has extended its range or abundance due to human disturbance, leading to imbalance in the ecosystem.
- In the review text, 'non-native/invasive' is used to refer to both the non-native species and the native species that have become 'invasive' (or 'over-dominant'; in some cases 'native assisted migrant' is grouped under 'native'.

- **Intervention level native composition:**
  - Selected: Native/at least some non-native/invasive/unclear/partially unclear/NA
- **Location of biodiversity:**
  - List: soil, aboveground, aquatic
  - Note: aboveground includes marshland, mangroves, as well as leaf litter

17. Outcomes

- **Direction of effect on ecosystem health metric**

  Selected from:

  - **Positive** – when the authors were explicit in describing the nature-based intervention as having benefits for the ecosystem health metric in question, compared to a counterfactual or other comparator (e.g. threshold). If over a given landscape the intervention had both neutral and positive effects, code as positive.
  - **Negative** – when the authors were explicit in describing the nature-based intervention as having negative effects for the ecosystem health metric in question, compared to a counterfactual or other comparator (e.g. threshold). If over a given landscape the intervention had both neutral and negative effects, code as negative.
  - **Mixed** – when both negative and positive outcomes were reported (e.g. over the spatial or temporal scope of the intervention).
  - **No effect/neutral** – code when the authors were explicit in describing the nature-based intervention as having no effect on ecological outcomes compared to a counterfactual or other comparator (e.g. threshold).
  - **Unclear** – when the authors did not derive an explicit conclusion as to whether the nature-based intervention has either negative, positive, mixed, or neutral outcomes as per the above definitions.

  The decision over direction of outcome depended on the framing given by authors as to what constituted an improvement in the situation with respect to the specific metric in question. Increases in e.g. incidence of non-native species, was coded as a positive outcome (for a non-native species) unless the authors framed the increase as a negative outcome.

- **Details of reported ecosystem health outcomes:**
  - Free text - copy and pasted details of outcomes from the study and/or wrote a detailed description, e.g. 'increase in number of plant species per hectare from 10 to 30 after 10 years of forest protection'.

18. **Data collection methods:**
    Listed all that apply (usually just one per ecosystem health metric:

    - **Visual surveys** on the ground, e.g. line transects, quadrats, mist netting, pitfall traps.
    - **Ex situ** i.e. samples followed by *ex situ* analysis, e.g. camera traps, eDNA samples.
    - **Remote sensing** e.g. LiDAR, satellite imagery.
    - **Modelling** e.g. dynamic vegetation models, model of survival.
    - **Secondary data,** where it is used directly as a biodiversity measure.
    - **Social or qualitative primary data** e.g. interviews.
    - **Unclear,** where the methods of the paper are insufficiently detailed to deduce how the data were obtained.

Excluded coding

The following data were systematically collected and can be found in the dataset (Dataset 1 ecosystem health outcomes), but were not used in the final evidence synthesis: information on declared conflicts of interest (of which there were none); relationships between authors, intervention and funding; ecosystem health outcome data type; type of comparator; whether a comparison was made between interventions; data collection methods.

Some further information was initially collected but is not presented here due to our concerns over its quality. Firstly, the goal of the interventions was coded because we hypothesised that it may effect ecosystem health outcomes, but we found that it was often difficult to discern what the primary objective(s) of the intervention were since this was usually only very briefly mentioned in the introduction of papers, with no supporting citation of evidence, meaning it relied on the authors' words accurately reflecting the intentions of potentially multiple groups of stakeholders. We were not sufficiently confident that our recorded intervention goals accurately reflected the true goals of interventions and so excluded this from our analysis; a more in depth study of intervention goals would be worthwhile.

For the subset of studies that passed a set of quality criteria (no conflict of interest, detailed methodology, and quantitative data), we initially extracted a further set of information. This included aspects of the intervention that might influence their effect on ecosystem health: the funding body for the intervention, the areas of society responsible for planning and governing the intervention, the areas of society responsible for managing the intervention on the ground, and the level of community involvement in the governance and management of the project. This information was often ambiguous or not mentioned in the papers themselves, and would require more extensive investigation to accurately discern, hence it was not included in the review.

S1.2 The relevance of native identity, and approach to analysis of relationships between outcome directions and features of the interventions or assessment method

The following ecosystem health metric types and sub-groups were excluded for the assessment of native/non-native species (3.2.2) because this is either not relevant to the metric or could be incorporated within it: habitat diversity, habitat cover when with respect to waterbodies, habitat density when with respect to water bodies, litter cover, elevation rate, habitat quality, connectivity, unspecified.

Our chi-squared analyses met all the assumptions specified by McHugh (2013) except for that only 60-66% of the cells had expected values of five or greater, rather than the recommended 80%. We therefore tested the same relationships with Fisher's exact test, for which all the assumptions were met, which produced the same results as chi-squared. The dependent and independent variable categories are defined in the Table B. It was necessary to pool outcome directions in order to permit statistical analysis; we decided to group positive with no effect since these interventions are united by all lacking a negative effect, and because in some cases an intervention could qualify as an NbS whilst having no effect on ecosystem health (such as protection of an intact forest). When assessing the relationship between binary outcome direction and broad taxonomic group, four interventions were excluded from the analysis which only had outcomes for either abiotic features of the environment (sediment accretion rate) or in the unspecified category (e.g. perceptions of change in overall ecosystem health from interviews). The intervention that only assessed abiotic features was also excluded from the assessment of the relationship between binary outcome direction and native/non-native identity.

Table B: Definitions of variables used in chi-squared tests.

Variable	Category	Definition
Binary outcome direction	Positive/no effect	Interventions with overall ecosystem health outcomes that are either positive or no effect, following the definitions of Chausson, Turner et al. (2020).
	Negative/mixed	Interventions with overall ecosystem health outcomes that are either negative or mixed, following the definitions of Chausson, Turner et al. (2020).
Habitat type	Forest	Interventions taking place in any forest ecosystem
	Other terrestrial	Interventions taking place in any terrestrial ecosystem other than forests. Includes savannah and saltmarsh.
	Coastal/freshwater	Interventions taking place in any coastal or freshwater ecosystem, where coastal includes mangroves but not saltmarsh.
	Combination	Interventions taking place in any combination of the groups above.
Broad intervention type	Protection	The intervention involves site-specific protection, e.g. protected areas and their management, private land conservation measures, reserves, conservancies, areas protected by indigenous or local communities. Same definition as in Chausson, Turner et al. (2020).
	Restoration	The intervention is described as active or passive restoration aiming to return degraded, damaged, or destroyed ecosystems to a pre-disturbance natural state; it must be specified that the intervention intends to restore or recover a natural or semi-natural state. Same definition as in Chausson, Turner et al. (2020).
	Management	Natural resource management interventions which do not include restoration or protection measures, e.g. ecosystem-based fire management. Same definition as in Chausson, Turner et al. (2020).
	Combination	When more than one type of intervention involving natural/semi natural ecosystems

		(management, restoration, or protection) is used. E.g. natural revegetation/reforestation in set asides/exclosures to protect abandoned land for regeneration. Same definition as in Chausson, Turner et al. (2020).
	Some/all created	An intervention which involves at least some establishment, protection or management of a 'created ecosystem', including the creation of a new ecosystem type in place of the naturally occurring one (e.g. afforestation of former grasslands, or creation of wetlands such as in peri-urban spaces), or where an ecosystem is modified such that it does not resemble its natural ecological state (e.g. rehabilitating degraded land with exotic species or reforesting an area with a single species where there used to be a diverse forest). This category merges the 'created ecosystems' and 'mixed created/non-created ecosystems' category from Chausson, Turner et al. 2020.
Broad taxonomic group	Just plants	Intervention outcomes that only involve plants; leaf litter is included in this category.
	Just animals	Intervention outcomes that only involve animals.
	Plants & animals	Intervention outcomes that involve both plants and animals.
	Plants & other	Intervention outcomes that involve both plants and other ecosystem components that are not plants or animals – this includes fungi, microbes and abiotic features such as rivers, as well as outcomes in the unspecified category.
	Plants, animals & other	Intervention outcomes that involve plants, animals and 'other' as described above.
Native/non-native identity	Native	Intervention outcomes that the authors specify assess only native species.
	At least some invasive/non-native	Intervention outcomes that the authors specify assess one or more invasive or non-native species; this includes species that are native but have become over-dominant. This could be in combination with native species.
	Unclear/partially unclear	Intervention outcomes for which the native or non-native/identity of the species studied is unclear for at least one of the species.

S1.3 Robustness assessment

The robustness assessment involved sequentially excluding interventions with: poor or absent methods; no comparator; one, two or three broad ecosystem health metrics. Data for the robustness assessment were obtained through systematic coding at the intervention level (see S1.1). Methods were categorised as detailed, poor (i.e. methods were present but not detailed enough for the methods to be repeatable), no methods, or use of secondary data. The type of comparator used was categorised as: control, baseline, space for time substitution, statement of change, threshold (use of non-exceedance of a threshold as evidence of a positive outcome), or no comparator. Broad ecosystem health metrics were assigned as in Table 2, and summed to find the number used per intervention.

S1.4 Meta-analysis supplementary methods

S1.4.1 Effect size choice and use

Hedges' *d* is commonly used as the effect size for ecological meta-analyses because it corrects for small sample sizes and adjusts for differences in sampling effort (Rosenberg, Rothstein, & Gurevitch, 2013). However, when pooling across ecological studies with different designs, scales or taxonomic groups, because Hedges' *d* is standardised against within-treatment variance, calculated differences in effect size between studies may be more associated with study design or taxonomic group (which are associated with the variance) than with the treatment itself (Spake & Doncaster, 2017). However, because the studies in our dataset are varied in design and taxonomic coverage, we chose to use the log response ratio (Ln*R*) as the effect size, which does not include a measure of variance and so is not confounded by systematic differences in variance like Hedges' *d* (Spake & Doncaster, 2017). This has the additional benefit of allowing more studies to be included since sufficient data to calculate variance was not given for all studies. Transforming the response ratio (*R*) with a natural logarithm normalises the skewed distribution, and linearises the metric (Hedges, Gurevitch, & Curtis, 1999).

Species richness data were obtained from studies, using WebPlot Digitiser (Rohatgi, 2020) to extract data from graphs where necessary. Studies that did not quantify the number of species involved, despite reporting a change in richness, were not included in the meta-analysis. When authors subdivided richness outcomes of interventions by either location within an intervention site, or by groups of species, response ratios were calculated for these subgroups separately; the mean of effect sizes *within* interventions were found before finding the mean effect sizes across interventions.

S1.4.2 Sensitivity analysis

A sensitivity analysis was performed to assess the effect on the mean response ratio of the following: two different methods of weighting, exclusion of non-native species and interventions with zero values for control/baseline values, and addition of fractions of one to control/baseline values of zero and their corresponding treatment values (Table A). The fact that none of these adjustments changed the sign of the mean effect size estimate increases the robustness of our analysis. We sought to include as many studies as possible to minimize publication bias (Koricheva, Gurevitch, & Mengersen, 2013), whilst being conservative in our estimates of effect size. This led us to using the non-weighted mean effect size, and excluding data points with zero values or non-native species; the full justification for this is given below.

Zero values for $\bar{x}_t$ and $\bar{x}_c$ cannot be used in Ln*R* without adjustment, since this would result in values of zero or infinity (Kalies, Chambers, & Covington, 2010). Adding 0.1, 0.5 and 1 to both $\bar{x}_t$ and $\bar{x}_c$ for the two interventions with zero $\bar{x}_c$ resulted in considerable inflation of Ln*R* compared to excluding these interventions. Therefore, to make our estimate of Ln*R* conservative, we chose to exclude these interventions for our mean effect size estimate.

The most standard weighting method for meta-analyses is by the inverse of the variance, however this would have required exclusion of a number of studies and can perversely assign heavy weights to studies with pseudoreplicated or little-replicated designs due to their small variances (Spake & Doncaster, 2017). Therefore, we tested the effect of two alternative weighting methods, based on study design and sample size respectively. The study design weighting, designed by Christie et al. (2019) simply weights before-after study designs with 0.226, and control-impact study designs with 0.206, based on tests of how different study designs perform on the same dataset; other study designs such as before-after-control-impact and randomised control trials were not found in our review. To weight for number of samples we used the following formula: $(N_c N_t/(N_c+N_t)$ where N_c is the number of control/baseline sites and N_t is the number of treatment sites (Ma & Chen, 2016). Since one intervention had no control sites, it was weighted with zero and hence excluded from the analysis. Both methods of weighting led to a slight inflation in mean Ln*R.* We chose to present the unweighted analysis in this review because both weighting systems exclude factors that considerably influence study accuracy, such as the size of individual samples, and weighting by number the samples alone gives several-fold higher weights to some studies due to great variation in the number of samples among studies. Moreover, unweighted analyses have been shown to reduce type I error and underestimates of effect size (Lajeunesse & Forbes, 2003; Marczak, Thompson, & Richardson, 2007).

The species richness change reported exclusively for non-native species was excluded from the analysis, resulting in a slight decrease in the mean effect size. The majority of the species richness comparisons were either for a mixture of native and non-native species, or the native identity was unclear. Since interventions involving created habitats are more likely to involve non-native species, for created habitat interventions with unclear native/non-native species identity, we looked up species ranges using Kew Science's tool: Plants of the World Online (http://www.plantsoftheworldonline.org/), but 13 interventions with unclear native/non-native species identity still remained. Only one intervention in the meta-analysis had confirmed species identity as only native, so we could not restrict our analysis to these data.

Table A: Sensitivity analysis for log response ratios. Testing the effects of excluding datapoints and adding 0.1, 0.5, and 1 to zero values.

Treatment	**Mean Ln*R***			**Geometric mean**		
	No weighting	**Study design weighting**	**Sample size weighting**	**No weighting**	**Study design weighting**	**Sample size weighting**
Excluding zero values	0.54	0.54	0.58	1.72	1.71	1.79
Adding 0.1 to zero values	1.45	1.50	1.91	4.25	4.46	6.74
Adding 0.5 to zero values	1.13	1.16	1.47	3.10	3.20	4.34
Adding 1 to zero values	1.00	1.03	1.28	2.72	2.79	3.60
Excluding non-native species and zero values	0.52	0.52	0.56	**1.67**	1.67	1.75
Excluding non-native species, and adding 1 to zero values	0.98	1.01	1.27	2.67	2.74	3.55

S1.4.3 Meta-analysis coding protocol

Here species richness is defined as a measure of the number of species in a given area, whether this is an estimate of the true number of species or simply a report of the number of species found from surveys; the latter is commonly referred to as 'species density'. Each of the following was recorded for each species richness outcome of an intervention:

1. Intervention ID and ecosystem health metric ID:

- Usually these were the same as those used for the systematic coding described in S1.1, but in some cases an intervention or metric needed to be subdivided for this stage of data collection due to authors reporting effects on species richness in disaggregated groups such as groups of species or locations. If subdivision by location reflected differences in the intervention itself then the intervention was subdivided, rather than the metric.
- Renaming was done as follows:
  - For interventions: INT-068-1 became INT-068-1.1m and INT-068-1.2m
  - For ecosystem health metrics: 44a became 44a1m and 44a2m

2. Direction of effect:

- Copied from first round of coding. A positive change corresponds with an increase in species richness between treatment and control/baseline, with either a statistically significant change or when no statistical test was done.
- However, if the ecosystem health metric has been subdivided for the meta-analysis coding, the direction may be different for each of the new rows.

3. Inclusion:

- Selected: Include/exclude; wrote an explanation for exclusion in the notes column.
  - Studies were excluded from the meta-analysis if they had insufficient data available to estimate species richness in the treatment and control/baseline. We did not exclude studies due to small sample sizes, in order maximise inclusion and reduce type I error (Lajeunesse & Forbes, 2003; Marczak et al., 2007). We included one intervention in the meta-analysis which used an alternative intervention in place of a no-intervention state as the control/baseline (and so the outcome direction was coded as unclear in the previous analysis), because this alternative intervention was the historical state at the site (Biel, Hacker, Ruggiero, Cohn, & Seabloom, 2017).
  - The supplementary material of studies originally coded for species diversity but not species richness was checked for species richness data, but none had sufficient information to be included in the meta-analysis.
  - Note that only included studies are shown in the data table provided.

4. Native/non-native:

- This was initially copied from first round of coding, however due to the prevalence of 'unclear' native/non-native identity, for created and mixed created/non-created interventions, we revisited the studies which were 'unclear' to see if they could be assigned to native, non-native or mixed. This was sometimes possible by looking up species ranges online e.g. using Kew Gardens' tool (plantsoftheworldonline.org).

5. Metric description:

- A precise description of the metric e.g. mean species richness per 10m^2.
- If an average must have been taken, but the authors did not state if they used the mean or another average, then we just referred to it the average.

6. Control for sample size:

- Description: e.g. rarefaction by bootstrapping; identified minimum sampling area with species accumulation curve.

7. Species richness in treatment and control:

- When species richness data was not reported in the form needed for the meta-analysis, wherever possible estimations were made. The following methods were used for estimation:
  - Extracted data using WebPlotDigitiser; in these cases, we rounded the number of species to nearest 0.1.
  - When species richness was given as a function of age, species richness before and after the intervention was estimated from the equations (Lennox et al., 2011).
  - When a time series was used, starting with one time step before the intervention was initiated, the first and last time steps were taken.
  - When several years before and after the intervention were measured, we took the average of before and after, although for Cardoso et al. (2008) where species richness declined after recovering post-intervention due to a flood, we used the data points before the flood.
  - Some interventions did not have a control, but only a baseline – see 8.

8. Species richness in treatment and control sites in the pre-intervention state:

- As with 7, but for the pre-intervention state (baseline) if given. Most interventions had values either for 7-treatment and 8-control, or 7-treatment and 7-control; few had data for all four.

12. Calculations behind treatment and control species richness:

- Only filled in if the coder had to do a calculation to estimate the species richness, e.g. if averaged across three sites within the study.

13. Number of treatment and control sites:

- If a range was given, we took the mid-point (e.g. 3-8 plots in each of 20 sites: used 5.5x20).
- For use in weighting.

14. Study type:

- Selected: BACI/BA/CI, where BACI is Before-After-Control-Impact experiment, BA if Before-After, and CI is Control-Impact experiment
- We treated the only study which had a BACI design like a before-after design (the mean species richness at the baseline site was used for $\bar{x}_c$) because the control site was not a close match to the treatment site before the intervention took place i.e. it was not a true BACI design.

15. Measure of spread:

- Free text: give the measure of spread (i.e. variance or standard deviation if available, or range or interquartile range) for both treatment and control/baseline

16. Notes on measure of spread:

- If a measure of spread given in previous 2 columns, state what type
  - Selected: Variance/SD/SE/Interquartile range/insufficient data, where SD is Standard Deviation, and SE is Standard Error; or if it did not fit into a category we described it.

17. Spatial scale of the intervention:

- Number of hectares over which the intervention took place (not over which data were gathered). If there were multiple exclosures over a larger site, but the outcomes are meaningful at the exclosure level, then we recorded the exclosure size.
- We took the mean size if appropriate, and estimated the area from maps given in papers if necessary.

18. Temporal scale of the intervention:

- Years passed since the intervention was initiated at the point when species richness data was collected.
- We counted years non-inclusively, e.g. if the authors said the intervention was present from 2003-2012, we recorded this as 9 years, unless more precise dates are available.

- If a range was given (e.g. 1-3 years) then we took the mid-point, or if possible the mean.
- For some interventions, the temporal scale may reflect time since the intervention, rather than the length of time for which the intervention has been in place – e.g. time since prescribed burns.

19. Statistical significance:

- Selected: Significant/not significant/not tested

20. Notes:

- We gave details of how each of the above factors were filled out for individual studies when assumptions had to be made.
- One study was excluded from the review because it was outside the scope; the ecosystem health outcome that had originally been coded as an ecosystem health outcome lacked an indication of a change due to the intervention (Tran & Brown, 2019).
- See Dataset 2 meta-analysis coding for data collected using this protocol.

References

Biel, R. G., Hacker, S. D., Ruggiero, P., Cohn, N., & Seabloom, E. W. (2017). Coastal protection and conservation on sandy beaches and dunes: Context-dependent tradeoffs in ecosystem service supply. *Ecosphere*, *8*(4). https://doi.org/10.1002/ecs2.1791

Cardoso, P. G., Raffaelli, D., Lillebø, A. I., Verdelhos, T., & Pardal, M. A. (2008). The impact of extreme flooding events and anthropogenic stressors on the macrobenthic communities' dynamics. *Estuarine, Coastal and Shelf Science*, *76*(3), 553–565. https://doi.org/10.1016/j.ecss.2007.07.026

Chausson, A., Turner, B., Seddon, D., Chabaneix, N., Girardin, C. A. J., Kapos, V., … Seddon, N. (2020). Mapping the effectiveness of nature-based solutions for climate change adaptation. *Global Change Biology*, (June), 1–22. https://doi.org/10.1111/gcb.15310

Christie, A. P., Amano, T., Martin, P. A., Shackelford, G. E., Simmons, B. I., & Sutherland, W. J. (2019). Simple study designs in ecology produce inaccurate estimates of biodiversity responses. *Journal of Applied Ecology*, *56*(12), 2742–2754. https://doi.org/10.1111/1365-2664.13499

Hedges, L. V., Gurevitch, J., & Curtis, P. S. (1999). The meta-analysis of response ratios in experimental ecology. *Ecology*, *80*(4), 1150–1156. https://doi.org/10.1890/14-2402.1

Kalies, E. L., Chambers, C. L., & Covington, W. W. (2010). Wildlife responses to thinning and burning treatments in southwestern conifer forests: A meta-analysis. *Forest Ecology and Management*, *259*(3), 333–342. https://doi.org/10.1016/J.FORECO.2009.10.024

Koricheva, J., Gurevitch, J., & Mengersen, K. (Eds.). (2013). *Handbook of Meta-analysis in Ecology and Evolution*. *Handbook of Meta-analysis in Ecology and Evolution*. Princeton University Press. https://doi.org/10.1515/9781400846184-002/PDF

Lajeunesse, M. J., & Forbes, M. R. (2003). Variable reporting and quantitative reviews: A comparison of three meta-analytical techniques. *Ecology Letters*, *6*(5), 448–454. https://doi.org/10.1046/j.1461-0248.2003.00448.x

Lennox, M. S., Lewis, D. J., Jackson, R. D., Harper, J., Larson, S., & Tate, K. W. (2011). Development of Vegetation and Aquatic Habitat in Restored Riparian Sites of California's North Coast Rangelands. *Restoration Ecology*, *19*(2), 225–233. https://doi.org/10.1111/j.1526-100X.2009.00558.x

Ma, Z., & Chen, H. Y. H. (2016). Effects of species diversity on fine root productivity in diverse ecosystems: a global meta-analysis. *Global Ecology and Biogeography*, *25*(11), 1387–1396. https://doi.org/10.1111/geb.12488

Marczak, L. B., Thompson, R. M., & Richardson, J. S. (2007). *META-ANALYSIS: TROPHIC LEVEL, HABITAT, AND PRODUCTIVITY SHAPE THE FOOD WEB EFFECTS OF RESOURCE SUBSIDIES*. *Ecology* (Vol. 88). https://doi.org/10.1890/0012-9658

McHugh, M. L. (2013). The Chi-square test of independence. *Biochemia Medica*, *23*(2), 143. https://doi.org/10.11613/BM.2013.018

Rohatgi, A. (2020). WebPlotDigitizer: Version 4.4. . Retrieved October 1, 2021, from https://automeris.io/WebPlotDigitizer/

Rosenberg, M. S., Rothstein, H. R., & Gurevitch, J. (2013). 6. Effect Sizes: Conventional Choices and Calculations. *Handbook of Meta-Analysis in Ecology and Evolution*, 61–71. https://doi.org/10.1515/9781400846184-008

Spake, R., & Doncaster, C. P. (2017, September 15). Use of meta-analysis in forest biodiversity research: key challenges and considerations. *Forest Ecology and Management*. Elsevier B.V. https://doi.org/10.1016/j.foreco.2017.05.059

Tran, L., & Brown, K. (2019). The importance of ecosystem services to smallholder farmers in climate change adaptation: learning from an ecosystem-based adaptation pilot in Vietnam. *Agroforestry Systems*, *93*(5), 1949–1960. https://doi.org/10.1007/s10457-018-0302-y