Supplemental Information on EcoEvidEx Study Design Fields

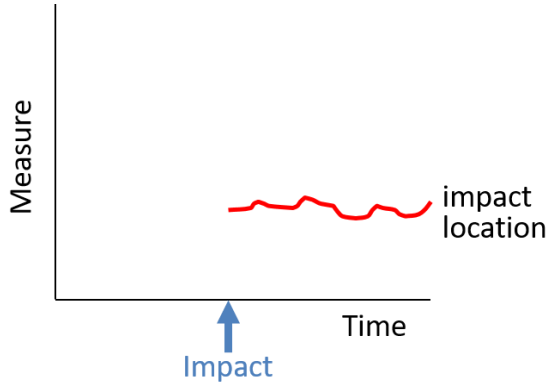
Specifying the study design fields in EcoEvidEx—in terms of overall design and how the data are analyzed, can be challenging. In this document, we provide additional clarification on the EcoEvidEx fields related to study design (Section 1), to build on the information provided in the general EcoEvidEx help file. We also provide example scenarios that show how different studies would be entered into the EcoEvidEx fields (Section 2).

Section 1. Additional information on selected study design fields in EcoEvidEx

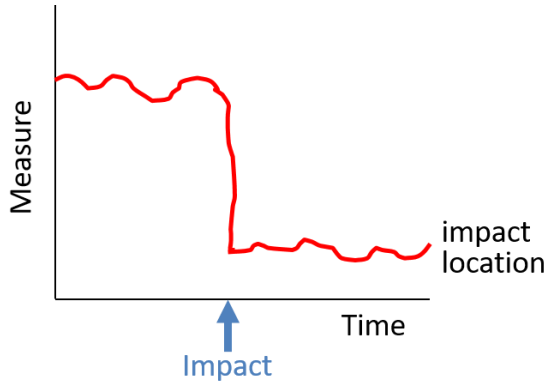
Study Design *

Brief description: Design of study used to generate relevant results
Field type: Drop-down list (9 options)
Options:

- ***After impact only***
  - Compares measures sampled at impact locations to a standard of some kind but not with control/reference location(s), and with no before impact or control/reference data.

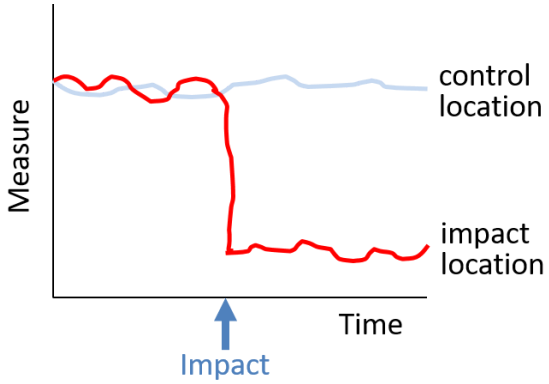


  - Example: dissolved oxygen measured at an impact location and compared to a water quality standard.
- ***Before vs. after (no control/reference)***
  - Compares measures sampled before and after some impact, with no control/reference data.



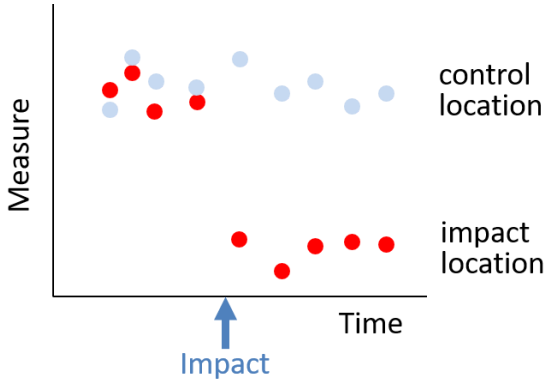
- Example: flow measured at the same location, before and after a wastewater treatment plant begins discharging immediately upstream.

- ***BACI/BARI (Before After Control Impact/Before After Reference Impact)***
  - Compares measures sampled at two locations (control/reference site vs. impact site), before and after impact.



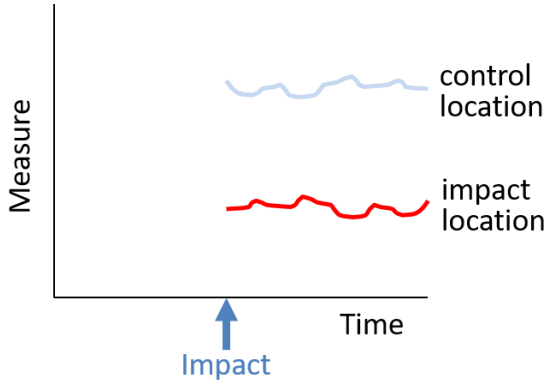
  - Looks at changes across two locations, each sampled once before the impact location is affected and once after the impact location is affected.
  - Example: total nitrogen measured in a reference stream and an agricultural stream, before and after a storm and subsequent runoff to both streams.

- ***BACIP (Before After Control Impact Paired)***
  - Compares measures taken at the same times (i.e., paired measures) at 1 control and 1 impact location.



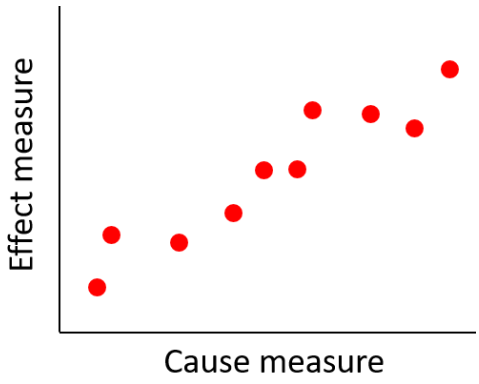
- Sampling over relatively long time intervals during both the before and after periods is used to estimate temporal variability in differences between control and impact locations.
- Time intervals should be long enough to prevent autocorrelation, and viewed as random samples of possible values in each time period.
- Example: river fish species richness measured at one reference site and one site with a lowhead dam installed; both sites are sampled at the same time once per season before and after installation of the dam.

- ***Control/reference vs. treatment/impact (no before)***
  - Compares control/reference locations to treatment/impact locations, with no before data at impact locations.



  - Example: benthic chlorophyll *a* is measured in 1 control treatment and one nutrient addition treatment in a lab experiment; benthic chlorophyll a is measured upstream (control) or downstream (treatment) of a paper mill effluent discharge.

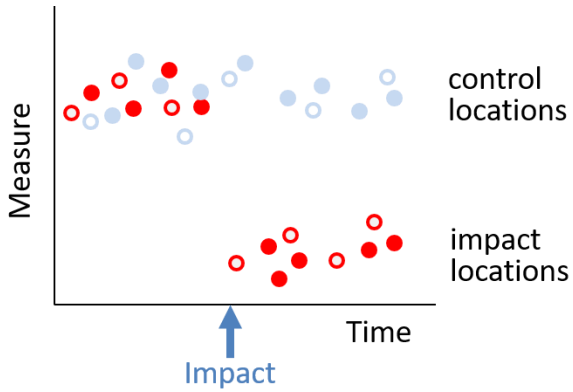
- ***Gradient***
  - Compares samples at multiple locations (spatial gradient), multiple times (temporal gradient), or multiple levels of a treatment (treatment gradient).



  - Gradient studies represent investigation of an association between a causal agent and a response (e.g., using correlation or regression analysis); results from a gradient study provide a dose-response gradient.
  - There is no requirement to distinguish between control and impact locations for a gradient design (although control or unimpacted sites may represent some of the sites along the gradient).

- If "gradient" is selected as the study design, specify the kind of gradient (spatial, temporal, treatment) under Design Description.
- Example (spatial gradient): total phosphorus and chlorophyll a are sampled at the same time, at 20 3rd-order streams in a drainage basin.
- Example (temporal gradient): total phosphorus and chlorophyll a are sampled monthly at the same site, over a 5-year period.
- Example (treatment gradient): mayfly survival is measured in the lab in three different temperature treatments.

- ***MBACI/MBARI (Multiple Before After Control Impact/Multiple Before After Reference Impact)***
  - Compares measures sampled at multiple control/reference locations and multiple impact locations.



  - MBACI/MBARI designs may also include replication in time.
  - Example: fish size is measured at 6 stream sites, over a 2-year period; in the first year all sites are similar, but at the start of the second year, wastewater treatment plants begin discharging at the upstream end of 3 of the rivers (i.e., there are 3 control locations and 3 impact locations followed through time, before and after impact).
- ***Other***
  - Any study design not represented by the other drop-down options.
  - If "other" is selected as the study design, provide a brief statement of the specific study design under Design Description.

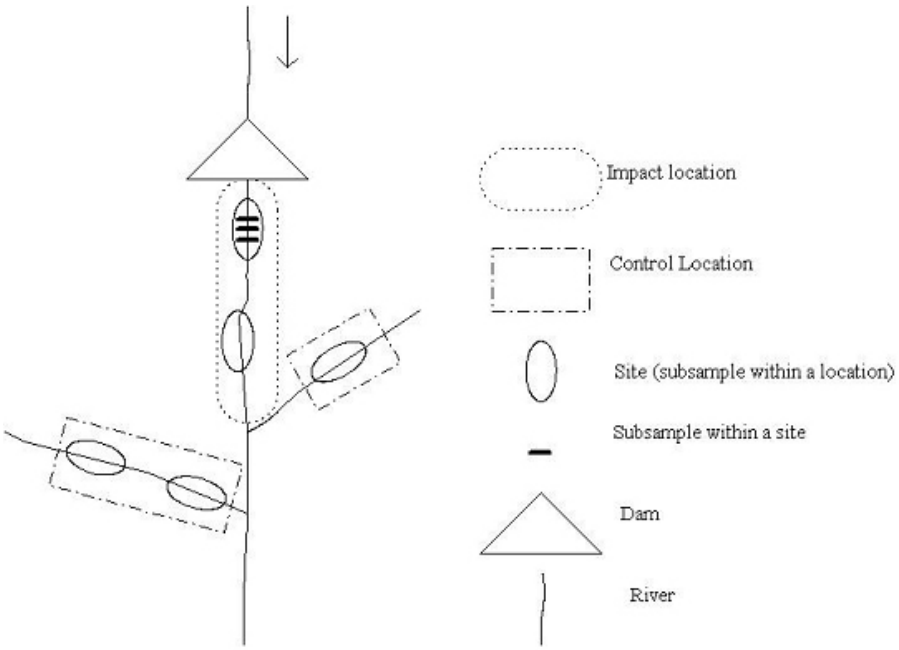
Number of Independent Sampling Units *

Brief description: Number of sampling units free of pseudoreplication
Field type: Free text (numeric)
Notes:

- This field represents the number of independent locations, sites, or experimental units evaluated in the study.
- In a gradient study, the number of samples might be equal to the number of independent sampling units (e.g., 1 sample taken at 10 independent streams), but often samples may be taken within locations or over time at the same location; thus, the number of independent sampling units may often be lower than the number of samples. For example, in the gradient examples provided above, the number of independent sampling units are 20 (total

phosphorus and chlorophyll a sampled at the same time, at 20 3rd-order streams in a drainage basin), 1 (total phosphorus and chlorophyll a sampled monthly at one site over a 5-year period), and 1 (mayfly survival measured in the lab at three different temperature treatments).

- When counting independent sampling units, do not count non-independent sites or sub-samples (they do not increase inferential power of the analysis). For example, in the figure below, there are 3 independent sampling units: 1 impact location and 2 control locations (this distinction between control and impact sampling units would be noted in the Control vs. Impact Sampling Units field, below). The sites within those locations, and sub-samples within those sites, do not represent independent samples.
- If you feel that authors used pseudoreplication in the analysis, provide your assessment of the number of independent units and explain your decision in the Replication field (see below).



- In determining sample units, consider how your study may be used in a synthesis study. These studies will often try to reach conclusions across sets of studies at a large experimental scale—so sites, rivers, watersheds, or other units might be the unit of replication.
- If in doubt, be conservative.
- Greater explanation of sampling units, replication, and sample size used in analysis can be provided in the three free text fields that follow (see below).

Control vs. Impact Sampling Units

Brief description: Number and/or explanation of control/reference vs. impact/treatment sampling units
Field type: Free text
Notes:

- This field should be completed only if a factorial design (i.e., BACI/BARI, BACIP, control/reference vs. treatment impact, MBACI/MBARI is selected under Study Design.

- It is important to consider the number of control vs. impact sampling units (e.g., locations or sites) actually used in the analysis for each effect; in some cases, not all control locations reported in a paper are used for every effect analysis.
- Replication of control sampling units reduces the likelihood of reaching spurious conclusions, whereas replication of impact sampling units reduces the likelihood of confounding.

Replication *

Brief description: Description of replication across or within sampling units
Field type: Free text
Notes:

- Use caution when assessing if there is replication across or within sampling units. Many observational studies may have 1 replicate per site, or may be pseudoreplicated (e.g., multiple samples taken at different locations along the same river); other study types may have true replication (e.g., a treatment may have 2 or more replicates).
- If you feel that authors used pseudoreplication to analyze a dataset, describe that here as an explanation for any discrepancy between what you have entered and the numbers reported in the paper.

Sample Size Used in Analysis *

Brief description: Sample size for the analysis that generates the response measure value (e.g., number of data points in a regression)
Field type: Free text (numeric)
Notes:

- This may be the same as the number of independent sampling units. However, these values may differ based the number of times sites were sampled, how samples are treated in the analysis, or the type of analysis used. For example, if all 5 replicates taken at each of 3 different rivers included in a simple linear regression analysis, the sample size used in analysis is 15, whereas the number of independent sampling units is 3.
- The sample size used in analysis can often be obtained by checking the degrees of freedom for the relevant statistical analysis.
- A number or range of numbers should be entered in this field.
- If further explanation is needed, describe under Design Description.

Section 2. Example scenarios for selected study design fields in EcoEvidEx

Scenario 1

A study samples 20 sites in each season (spring, summer, autumn, winter) in 2017. Each sampling event involves collecting 5 benthic macroinvertebrate samples (which are pooled into 1 composite sample) and 1 water sample at each site. Sites are the unit of analysis in a linear regression of total nitrogen and macroinvertebrate richness; each point in the regression represents the average, across all 4 seasons, for total nitrogen and macroinvertebrate richness at a site.

- Study design: gradient (spatial)
- Number of independent sampling units: 20
- Control vs. impact sampling units: not applicable

- Replication: 4 replicates per site (one replicate per season)
- Sample size used in analysis: 20

Scenario 2

A study examines species richness along a gradient in a floodplain, moving from the river channel to the far margin of the floodplain. Three transects are employed, with 20 sampling points along each transect. Samples are collected once. The data are analyzed as a linear regression, with species richness tested against vertical elevation of the sampling points.

- Study design: gradient (spatial)
- Number of independent sampling units: 1 (there is only one floodplain)
- Control vs. impact sampling units: not applicable
- Replication: 3 replicates per site (3 transects)
- Sample size used in analysis: 60

Scenario 3

Mayflies from two separate populations are collected and taken to the laboratory. 40 individuals from each population are divided into control and impact groups (20 each), with each group subdivided into 4 sub-groups of five. Each sub-group is either exposed to a single dose of an experimental herbicide or kept as a control, with body mass measured after 14 days. The data are analyzed using a nested ANOVA, with sub-groups nested within the combination of treatment (control or impact) and population (1 or 2).

- Study design: control/reference vs. impact (population source is not relevant to the hypothesis being asked)
- Number of independent sampling units: 2 (2 independent population of mayflies)
- Control vs. impact sampling units: 1 control, 1 impact
- Replication: 4 replicates per treatment (4 sub-groups)
- Sample size used in analysis: 80 (total n for the ANOVA)