Price/VanBlaricom Geoduck Aquaculture Harvesting Study Lacks Statistical Rigor

Gary A. Ritchie, Ph.D.

Washington State Shorelines Hearings Board Testimony

August 12, 2013

Impact of geoduck harvesting on benthic communities

- Proponents claim that there are no significant effects of geoduck harvesting on benthic communities in Puget Sound.
- Claim is based largely on a Master of Science thesis written by Jennifer Price at the University of Washington.

Jennifer Price MS Thesis

- "Quantifying the Ecological Impacts of Geoduck (Panopea generosa) Aquaculture Harvest Practices on Benthic Infauna," Jennifer Price; August 18, 2011.
- Study Objectives… (page 8) "to quantify the [geoduck] harvest impact (if any) on benthic organism abundance, diversity, and community structure…"
- Conclusion (page 33): "…the effects of geoduck harvest on the benthic community in Puget Sound are at most minimal."

Result is widely reported in meetings, hearings and the media

- Examples:
  - Longbranch SHB hearing. March 1, 2012. Dr. Glenn VanBlaricom presented results and conclusions of Price thesis.
  - Sea Grant Symposium, Alderbrook WA, March 6, 2012. Dr. VanBlaricom again presented results and conclusions of Price thesis.
  - Kitsap Sun, March 10, 2012. "Glenn VanBlaricom, a researcher with the UW School of Aquatic and Fishery Sciences, said benthic organisms, which live on or in the sediments, don't appear to be affected much by geoduck harvesting."
  - Henderson Inlet Hearing November 26, 2012. ENVIRON (a consulting firm) presented a review of environmental effects of geoduck aquaculture. Cited the Jennifer Price thesis as indicating that geoduck harvesting has no significant effect on benthic organisms.

I will argue that…

- the conclusion "…*the effects of geoduck harvest on the benthic community in Puget Sound are at most minimal.*" is not supportable…
- …because the statistics (experimental design and analysis) were not done correctly.

What are statistics?

- Descriptive statistics
  - Baseball scores
  - Populations of cities
  - Average income of a specific group
- Inferential statistics
  - Highly developed field of applied math that helps people "get their heads around" large volumes of data.
  - Essential in design of experiments and analysis/interpretation of experimental results.
  - Enables you to draw inferences regarding the results of an experiment.

Example

- Test of dog food supplement:

Treatment	Weight gain (lbs.)
Control (no supplement)	10.2
Level 1 (1x supplement)	10.8
Level 2 (2x supplement)	10.7

- Are differences "real" or result of random chance?
- Compute a statistic called an "F-ratio."
- F-ratio estimates probability of a "real" result.

Price study overview

Three sites

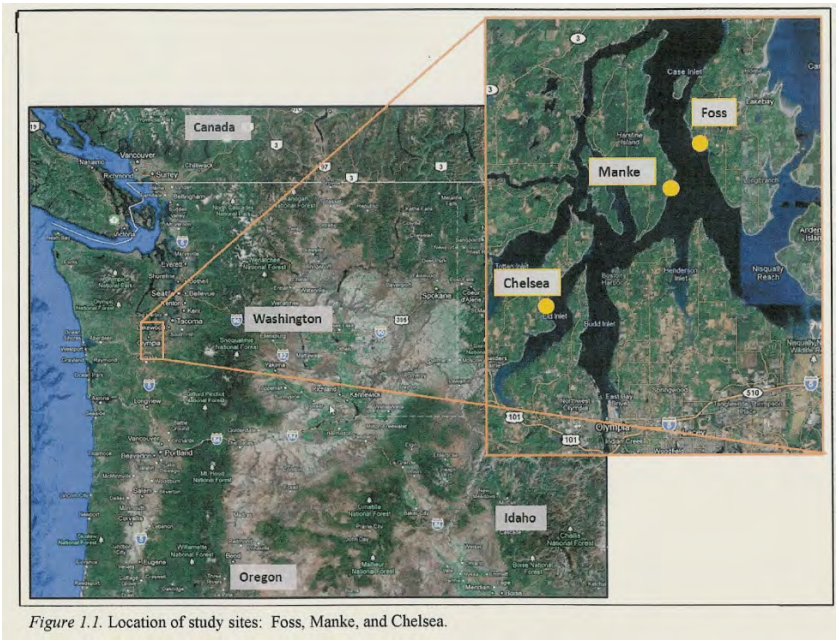


Figure 1.1. Location of study sites: Foss, Manke, and Chelsea.

One Cultured/Reference plot on each site

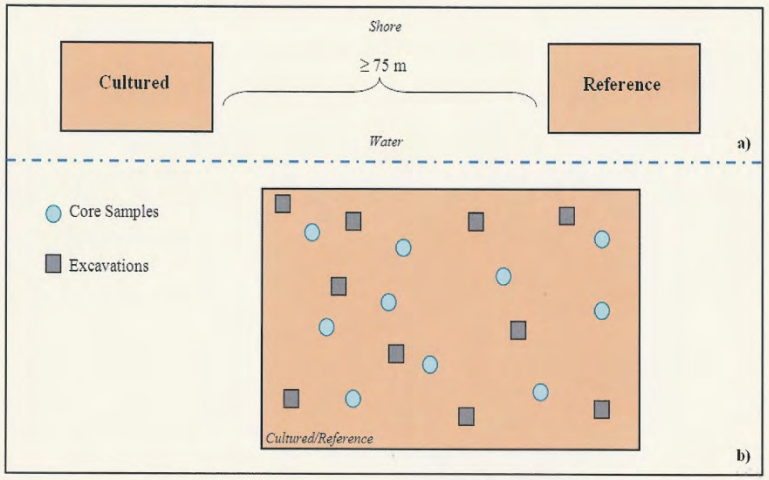


Figure 1.2. a) Relative layout of cultured and reference plots in the intertidal zone, b) sampling strategy within cultured and reference plots.

40

So, what's the problem?

- The researchers unwittingly fell into a statistical trap called

"Pseudoreplication"

- **Replication** is the independent application of a treatment.

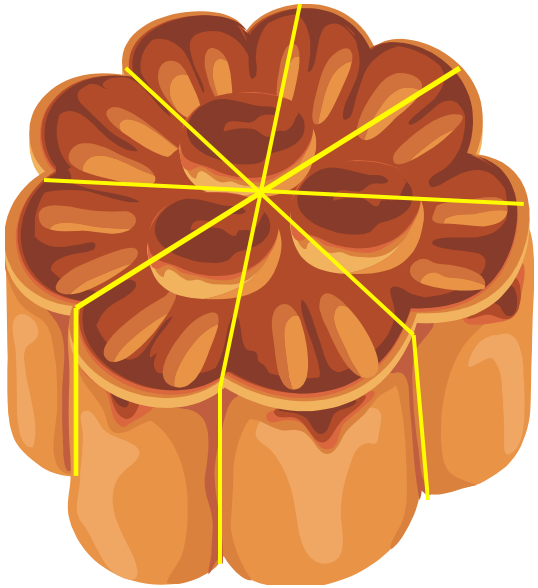
- **Pseudoreplication** is treating something as a replicate that really isn't a replicate (sub-samples, repeated measures, etc.).

Four examples of

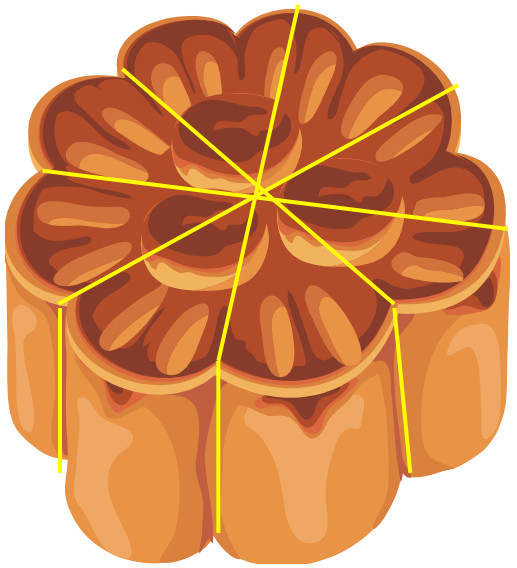
pseudoreplication

1. Cake experiment

Betty Crocker



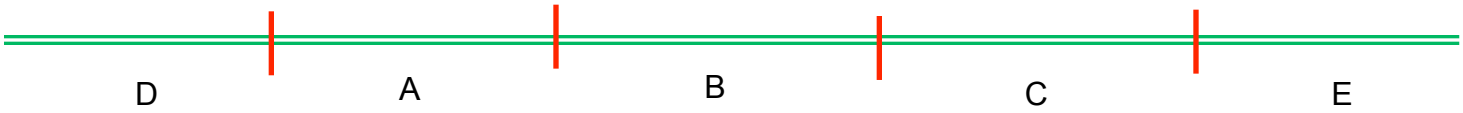
Brand X



2. Row plot experiment

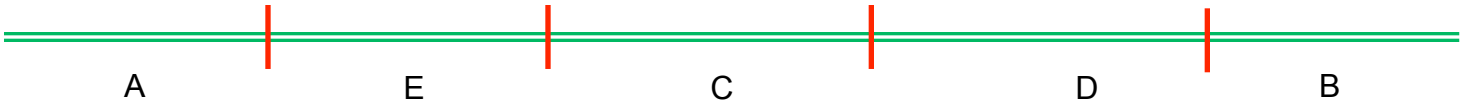
- Control

- Var. 1

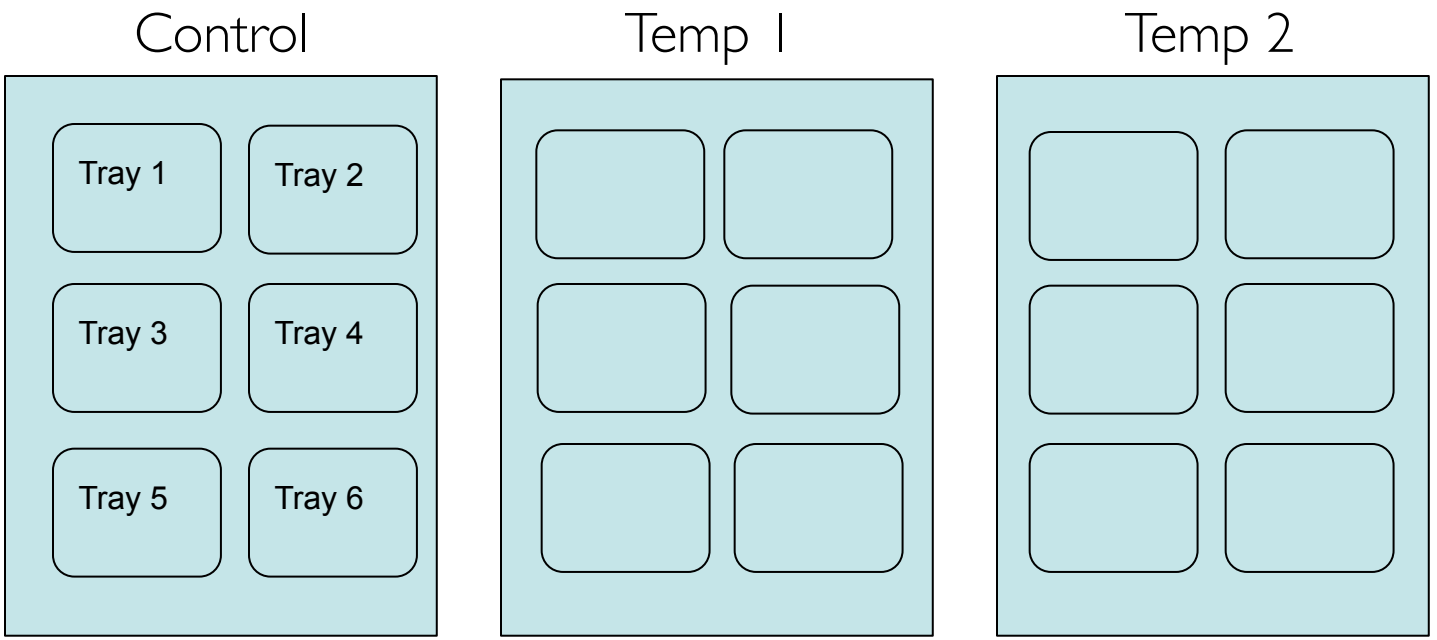


- Var. 2

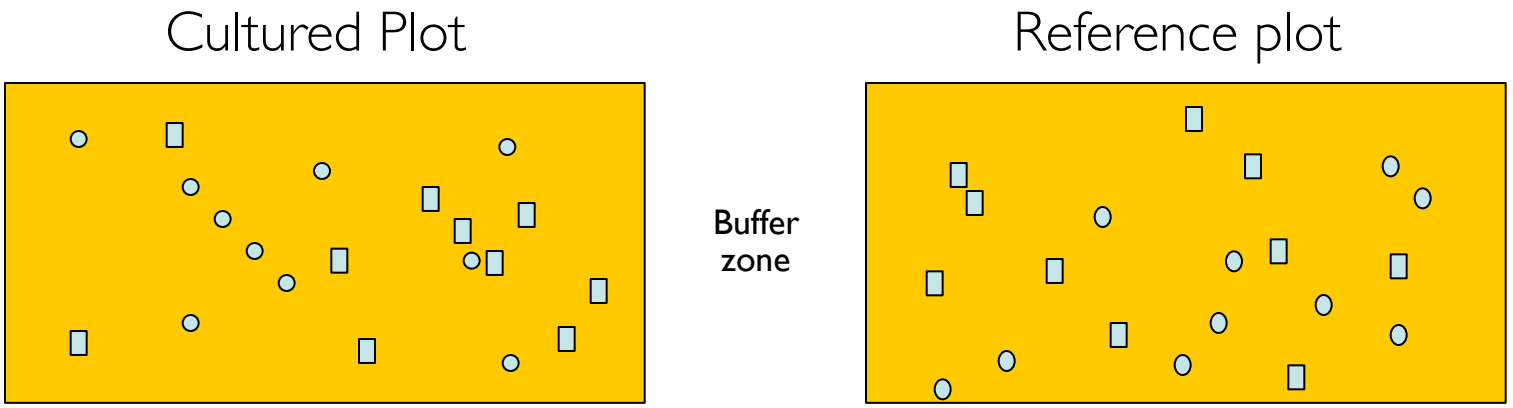
- Var. 3



3. Controlled environment chamber experiment



Geoduck harvesting study



What's wrong with pseudoreplication?

- Without true replication you *cannot test for significance between treatments* because you cannot compute "error variance" necessary to conduct the F-test of statistical significance (Milliken and Johnson 1997*).

- R.A. Fisher: "No one would…dream of testing the response to a treatment by comparing two plots, one treated and the other untreated."

- George Milliken, Pers. Comm.: "Pseudoreplication is like going back to the same doctor for a second opinion."

*Milliken, G.A. and D.E. Johnson. 1997. Analysis of Messy Data: Vol. 1, Designed Experiments, Chapman and Hall, London, New York, 473 p.

Pseudoreplication is a common statistical trap

- I fell into the pseudoreplication trap early in my career with the Growth Chamber experiment I described earlier.
  - I served for 25 years as a referee for nine scientific journals.
    - *Canadian Journal of Forest Research*
    - *Forest Science*
    - *New Forests*
    - *Journal of Tropical Forest Science*
    - *Scandinavian Journal of Forest Research*
    - *Annales des Sciences Forestieres*
    - *Western Journal of Applied Forestry*
    - *Tree Physiology*
    - *Annals of Botany*
- Pseudoreplication was the most common statistical problem I encountered.
- The ecological literature contains many examples of pseudoreplication.

Back to the Price study

- They had no replication on any of the three sites, hence they cannot make valid statistical comparisons between treatments at any site.
- But they had 3 sites:
  - Page 11: "*Because of site differences, each site was analyzed independently and the sites were not considered replicates.*"
  - So there is no replication either at the site or plot level.
- With no replication anywhere in the study, statistical significance cannot be computed and no conclusions can be drawn.
- I conferred with two statisticians on this conclusion: ***both agreed***.

What should they have done?

- Achieve broader scope of inference by using more, carefully selected sites.
- Determine the size of the impact they were trying to detect using power analysis.
- Incorporate more replications on site based on results of power analysis.
- Perhaps user smaller plots, fewer subsamples, to achieve more replications.

Other concerns

Narrow scope of inference (moot)

- Three sites in lower Case Inlet would not seem to support the conclusion: "…the effects of geoduck harvest on the benthic community in Puget Sound are at most minimal."

Lack of statistical power (moot)

- How big an effect were they looking for?
- How was size of the experiment determined?
- Not discussed in Price thesis.

Data quality control

- What measures were put into place to ensure that no data errors occurred?
- Not discussed in Price thesis.

Summary and conclusions

- The Price thesis did not contain proper replication, so statistical tests conducted on the results are invalid and no conclusions can be drawn from them.
- Even if the study had been properly replicated, the narrow scope of inference would not support Puget Sound-wide conclusions.
- In his verbal presentations and on the Sea Grant web site, Dr. VanBlaricom has emphasized that the results of the Price thesis have neither been peer reviewed nor published, hence they **must not be cited.**
- Unfortunately his admonitions have fallen on deaf ears.