

**Types of Experiments and Causal Process Tracing:
What Happened on the Kaibab Plateau in the 1920s?**

Presented at PSA 2014, November 2014, Chicago, IL

IMPORTANT NOTE:

Although this paper is an accurate reflection of what was presented at the PSA, it is still a draft. Thus, please do not cite or quote without permission

Roberta L. Millstein
Department of Philosophy
University of California, Davis
<http://www.RLM.net/>

How I came to this topic

An interest in ways of demonstrating causes in the field led to an interest in natural experiments, which led me to read Jared Diamond and James A. Robinson's (2010), *Natural Experiments in History*, which led me to re-read Diamond's (1986) classic "Overview: Laboratory Experiments, Field Experiments, and Natural Experiments" and to read Thad Dunning (2008), "Improving Causal Inference: Strengths and Limitations of Natural Experiments."

Dunning refers to "causal process tracing," which I found intriguing: what is that and how does it relate Diamond's typology of different types of experiment? Is it relevant outside of social science? What can it reveal that cannot be revealed by the different experiment types? To explore answers to these questions, I'll discuss a study of the classic wolf-deer Kaibab case from the 1920s. But first, some background.

Diamond's typology of experiments

Jared Diamond (1986) distinguishes between three main types of experiment that are performed in *community ecology*:

1. **Laboratory Experiment (LE)** - perturbations are produced by the experimenter in the laboratory.
2. **Field Experiment (FE)** - perturbations are produced by the experimenter in the field.
3. **Natural Experiment (NE)** - natural perturbations occur in the field; they are not produced by the experimenter. This includes perturbations due to "humans other than ecologists."

According to Diamond, in practice, LEs, FEs, and NEs form a continuum (as we shall see).

With LEs, one regulates the abiotic environment (light, temperature, water, etc.) and the biotic environment, often with communities of two or a very few species. Some communities are randomly assigned to receive the treatment whereas other communities do not (the control).

With FEs, one selects outdoor sites so as to initially have the same values of unregulated variables to the extent possible; they are generally adjacent to one another (this will be important later) However, there is no further regulation of the environment. Some

communities are randomly assigned to receive the treatment (e.g., removal or introduction of a species) whereas other communities do not (the control).

There are two types of NEs: Natural Trajectory Experiments (NTEs) and Natural Snapshot Experiments (NSEs). With NTEs, comparisons of the *same* community at various times before, during, and after a witnessed field perturbation are performed. The environment is not regulated. Before the perturbation, the site serves as the control; after the perturbation, it can be considered to have received the “treatment.” So, the “sites” are matched except to the extent they change over time. With NSEs, comparisons of *different* communities assumed to have reached a quasi-steady state with respect to the perturbing variable are performed. Again, the environment not regulated. Sites are matched to the extent possible; however, sites are often distant in space. Sites having experienced a certain perturbation are considered to have received the treatment; sites that have not are considered the control.

According to Diamond, the three types of experiment differ in their merits; there are *tradeoffs*. With respect to regulation of independent variables and site matching (both of which give confidence in the causal inference), $LE > FE > NTE > NSE$. With respect to maximum spatial and temporal scale, $NE > FE > LE$ (for practical reasons). With respect to the range of species and perturbations that can be studied, $NSE > NTE > FE > LE$ (for both practical and ethical reasons). With respect to realism: $NE > FE > LE$. Finally, with respect to the ability to follow the causal trajectory: $LE, FE, \text{ and } NTE$, but not NSE .

Causal process tracing (CPT)

But what is *causal process tracing*? Clear definitions are a bit hard to come by; this is the best I have found:¹

Process tracing... is an analytic tool for drawing descriptive and causal inferences *from diagnostic pieces of evidence*— often understood as part of a temporal sequence of events or phenomena (Collier 2011, 824; emphasis added).

As a tool of causal inference, process tracing focuses on the unfolding of events or situations *over time*. Yet grasping this unfolding is impossible if one cannot adequately describe an event or situation at *one point in time*... To characterize a process, we must be able to characterize key steps in the process, which in turn permits good analysis of change and sequence (Collier 2011, 824; emphasis in original).

¹ See also Crasnow 2012, Morgan 2012.

However, contra my original assumption, there does not seem to be one specific *CPT method*. Many kinds of data can be used in CPT, both qualitative and quantitative. Interviews, written records, etc., are all legitimate. Examples of CPT are as disparate as Semmelweis's discovery of the cause of puerperal fever, understanding the results of the 2000 U.S. Presidential election (Bush vs. Gore), and Sherlock Holmes's detective work. Thus, perhaps the general class that includes LEs, FEs, NEs, and CPTs can be considered *cause-finding practices*. If so, what can these different ways of finding causes tell us? Are CPTs used outside of the social sciences?

Let's turn to the case of the Kaibab for partial answers.

The Canonical Case of the Kaibab

In part inspired by his study of the Kaibab Plateau in Arizona in the 1920s, Aldo Leopold wrote:

"I have lived to see state after state extirpate its wolves. I have watched the face of many a newly wolfless mountain, and seen the south-facing slopes wrinkle with a maze of new deer trails. I have seen every edible bush and seedling browsed, first to anaemic desuetude, and then to death. I have seen every edible tree defoliated to the height of a saddlehorn. Such a mountain looks as if someone had given God a new pruning shears, and forbidden Him all other exercise. In the end the starved bones of the hoped-for deer herd, dead of its own too-much, bleach with the bones of the dead sage, or molder under the high-lined junipers" (Leopold 1949, from "Thinking Like a Mountain")

Less poetically, Leopold wrote:

"We have found no record of a deer irruption in North America antedating the removal of deer predators. Those parts of the continent which still retain the native predators have reported no irruptions. This circumstantial evidence supports the surmise that removal of predators predisposes a deer herd to irruptive behavior" (Leopold 1943).

The Kaibab would come to be seen as the exemplar of the dangers of predator removal and the exemplar of a *trophic cascade*, found in many ecology textbooks (Young 2002). A trophic cascade is "[w]hen the presence of top trophic-level predators significantly affects herbivores (the next lower trophic level), and this interaction alters or influences vegetation (e.g., species composition, age structure, or spatial distribution)" (Ripple and

Beschta 2005). A loss of predators leads to an increase in herbivores which leads to a change in vegetation.

But the canonical case of the Kaibab was debunked by Graeme Caughley in 1970:

“Little can be gleaned from the original records beyond the suggestion that the population began a decline sometime in the period 1924-1930, and that this decline was probably preceded by a period of increase. Any further conclusion is speculative.”

Caughley points to a *reduction in sheep and cattle* as a more likely cause. Then, thirty-six years later, Dan Binkley and colleagues debunked the debunker:

“We conclude that Caughley’s (1970) hypothesis about the reduction of livestock/ deer competition as a driver of the irruption is refuted...The evidence for deer irruptions following periods of reduced predation was consistent for both the 1920s and the 1940s, supporting the idea that predation limits the density of low deer populations, and food limits deer populations (and the absence of aspen recruitment) at high populations.”

In responding to Caughley (1970), Binkley et al. (2006) focused primarily (but not exclusively) on the herbivore/ vegetation part of the trophic cascade, i.e., the claim that an increase in herbivores leads to a change in vegetation. More specifically, they examined the question of whether a loss of wolves and cougars led to increase in deer which led to a decrease in aspen (deer eat young aspen shoots, preventing aspen regeneration).

In what follows, I reinterpret their study as consisting of three parts:

1. a natural trajectory experiment (NTE)
2. two natural snapshot experiments (NSEs)
3. causal process tracing (CPT)

Part 1: Binkley et al’s Natural Trajectory Experiment

The NTE examined the numbers of quaking aspen of different ages, via a study of the aspens present on the Kaibab today, using tree diameter as a proxy for age. Remember that for an NTE, one performs “comparisons of the *same* community at various times before, during, and after a witnessed field perturbation” - here, a study of the aspen population *through time*.

Binkley et al. found that the age structure of aspen on the Kaibab *generally* followed a typical pattern for all-aged forests, with number of trees decreasing exponentially with age. But there were exceptions to this pattern: periods when there were *more* aspen than expected (1877-1886 and 1967-1992) and periods when there were *fewer* aspen than expected (1913-1937, *especially between 1923-1927*, and 1953-1962). Note that low aspen numbers correlate with periods of purported deer irruptions (the perturbation) in the 1920s and 1950s.

Part 2: Binkley et al's Natural Snapshot Experiment

The NSEs examined the ages of quaking aspen present on the Kaibab today, comparing aspen within a fenced-in area (no deer) *to* aspen outside that area (deer) as well as comparing aspen within an area protected by dogs (no deer) to aspen outside that area (deer). Remember that with an NSE: “comparisons of *different* communities assumed to have reached a quasi-steady state with respect to the perturbing variable” – here, the perturbing variable is the deer. However, unlike most NSEs, the “control” and “treatment” sites were adjacent, similar to an FE in the matching of the sites.

Binkley et al. found that “The only successful aspen recruitment during this period was found in areas protected from deer by fences or dogs.”

Combining the NTE and the NSE

The NTE and the NSEs each provide causally relevant evidence. The NTE reveals that increased numbers of deer and decreased numbers of aspen are correlated *through time*. The NSEs give FE-quality evidence that increased numbers of deer are a causal factor for decreased aspen recruitment (again, the only significant aspen recruitment from the 1920s is found where in places where there were no deer, adjacent to locations with deer but no aspen).

So, by combining NSEs with an NTE, Binkley et al. strengthen their causal inferences; there is evidence that increased deer numbers caused decreased aspen numbers through time. But what about other possible causes? That's where the CPT comes in.

Part 3: Binkley et al.'s Causal Process Tracing

The CPT was a reconstruction of the late 19th-20th century timeline of the Kaibab, looking for correlations between other possible causal factors and aspen recruitment. Remember that CPTs use multiple sources of information, including qualitative sources,

to infer the temporal sequence of events or phenomena.

Their sources included:

- Comparative ring-width chronologies for some of the older aspens and pines (to help determine the past effects of climate)
- Interviews with Dennis Lund (formerly with the Kaibab National Forest) and John Goodwin (Arizona Fish and Game Department)
- Published and unpublished records from the Arizona Fish and Game Department
- Published reports from USDA Forest Services
- Secondary sources from peer reviewed journals such as *Journal of Climate*.

Binkley et al. found that increased aspen numbers (which might have led to deer irruption) were not significantly correlated with climate, decrease in livestock (sheep and cattle)² but that they were correlated with fire suppression. They found that decreased aspen numbers (which might have resulted from deer irruption) were not significantly correlated with climate, fire, increase in livestock, logging but were correlated with increased deer numbers. And they found that increased deer numbers were correlated with decreased human hunting of deer, increased human hunting of predators, and fire suppression.

These findings challenge Caughley's (1970) hypothesis that removal of livestock caused the deer irruption and support Leopold's hypotheses concerning hunting and fire suppression.

Combining NTE, NSE, and CPT

By adding the CPT analysis to the NTE and the NSEs, Binkley et al mitigate the shortcomings of the natural experiments. In natural experiments, independent variables cannot be regulated; Binkley et al's CPT rules out other plausible causes of the low aspen numbers in the 1920s and gives some weight to predators over livestock as the cause of the deer irruption. Moreover, the CPT rules these out as causes not just at the time in question but through time (late 19th century -> 20th century). In sum, Binkley et al. leverage the strengths of NEs (realism, large spatial/ temporal scale) with fewer weaknesses.

Combining Cause-Finding Practices

² In these cases, Binkley et al. (2006) cited sources to correct Caughley's (1970) numbers.

Diamond claims: “Ecologists, like scientists in many other fields, can profit by applying *different methodologies to the same system*” (1986, 21; emphasis added). Among other reasons, we achieve a more complete understanding of the system by using different experimental methods because each methodology yields some information that is inaccessible to the others.

The analysis I have given here supports Diamond’s claim; moreover, the addition of a CPT analysis yields even more information than just LEs, FEs, and NEs would have.

Other possible roles for CPT in conjunction with other cause-finding practices include filling in stages between causes, revealing the intervening process (Millstein 2013) and challenging the conclusions of other cause-finding practices (Brady 2010).

Conclusions

I’ve argued that Binkley et al. make use of a CPT to settle questions of causality on the Kaibab plateau in the 1920s. CPT has been typically deployed in the social sciences; this raises the question of how common CPTs are in ecology and other natural sciences. Notable here is that Binkley et al. are examining a *particular historical case*.

Crasnow argues that “specific pieces of evidence produced through process tracing are useful as evidence for singular [“token”] causation” (2012, 665; emphasis added).

Binkley et al.’s conclusions fit this general characterization:

- *In the Kaibab in the 1920s*, there was a deer irruption which depleted aspen
- *In the Kaibab in the 1920s*, the loss of predators caused a deer irruption.

The authors also hint at their case in combination with other cases as evidence for trophic cascades; in other words, they believe that their case contributes to general (“type”) causation as well:

“Our results combine with other case studies (for example, Gasaway and others 1992; Krebs and others 2003; Ripple and Beschta 2003, 2004) to indicate that top-down control of food webs is probably not unusual in terrestrial ecosystems” (Binkley et al. 2006, 240).

I’ve further argued that Binkley et al. use CPT in conjunction with an NTE and two NSEs. This shows that Leopold may have been right about trophic cascade in the Kaibab in the 1920s, i.e., that there are good (albeit defeasible) reasons to think that a loss of predators (together with fire suppression) led to a deer irruption which decreased aspen recruitment. Using the different cause-finding practices in

combination can strengthen causal inferences and mitigate the shortcomings that each practice has.

Acknowledgements: Thanks to Chris Young for helpful discussion and for his detailed analysis of the Kaibab in his book, *In the Absence of Predators: Conservation and Controversy on the Kaibab Plateau*, to the Griesemer / Millstein Lab for reading and discussing some of the background material on Diamond's work and causal process tracing, which formed the inspiration for this talk, and to the audience at ISHPSSB 2013 for comments on an earlier version of this talk that involved guppies instead of deer, wolves, and cougars.