A basic guide for empirical environmental social science

Michael Cox [1]

ABSTRACT. In this paper, I address a gap in the literature on environmental social science by providing a basic rubric for the conduct of empirical research in this interdisciplinary field. Current literature displays a healthy diversity of methods and techniques, but this has also been accompanied by a lack of consistency in the way in which research in this area is done. In part this can be seen as resulting from a lack in supporting texts that would provide a basis for this consistency. Although relevant methods texts do exist, these are not written with this type of research explicitly in mind, and so translating them to this field can be awkward. This paper is designed to fill this gap and enable more consistency in the conduct of empirical environmental social science. Basic types of research designs and methods are covered, as are criteria for evaluating these methods.

Key Words: *environmental social science; research design; research methods;*

INTRODUCTION

In this paper, I provide a guide for the conduct of empirical environmental social science (ESS), a type of social science that is closely related to the study of social-ecological systems and social-ecological research. It does so from a positivist perspective, emphasizing the collection of empirical data with the intent to uncover regularities across a set of observations. Castree et al. (2014:765) state that ESS "has two aims: (1) to study systematically the presuppositions, norms, perceptions, preferences, relations, regulations and institutions that together structure how humans value and use the non-human world; and (2) to identify and evaluate ways of altering human behaviour in light of one or more definitions of desirable or necessary ends. As part of this second aim, many environmental social scientists work with those effecting, or affected by, environmental change, rather than just conducting research on them."

ESS is a highly interdisciplinary and frequently participatory area of research that incorporates many scientific approaches, including institutional analysis, ecology, political ecology, geography, and anthropology, as well as the study of complex social-ecological systems. ESS is thus a highly diverse and interdisciplinary approach to science. This diversity and the exploratory nature of many ESS analyses, although providing important benefits to the communities of researchers involved, has left several gaps in the literature. One primary gap is that there are few standard instructional materials for this area of research. Several edited volumes and articles that synthesize ESS-based approaches do exist (Young et al. 2006, Moran 2010, Vacarro et al. 2010, Manfred et al. 2014), and the foundational literature on social science methods is substantial (e.g., see King et al. 1994, George and Bennett 2005, Bernard 2011). However, to-date there has been little integration between these two literatures to provide a synthetic rubric for the conduct of ESS. My aim is to fill this instructional gap by providing a research rubric that is tightly integrated with the ESS literature. Methodological concepts are presented with explicit reference to examples from ESS and are accompanied by discussions of the prevalence and importance of each concept within the ESS literature. Concepts and issues that are highly prevalent in the ESS literature are emphasized throughout.

I do not claim to provide an exhaustive treatment of ESS or of the different methods and methodologies that it encompasses. What this paper integrates is a particular subset of ESS with a particular set of methodological literature. It does not synthesize all ESS-related literature. Specifically, I do not cover quantitative social-ecological modeling and scenario-building techniques, which have been enormously important parts of the ESS literature, but have each received much attention elsewhere. My primary goal is to enable the beginning environmental social scientist to plan a research project and write a proposal that would guide such a project.

To support the text in the main body of this paper I have included several appendices. Key terms are defined in Appendix 1. For a more basic review of scientific terminology, e.g., concepts, variables, or theories, that is used throughout the paper and throughout ESS, the reader should refer to Appendix 2.

ELEMENTS OF A RESEARCH PROJECT

There are four primary elements of a research project that need to be developed and carried out to address a research question. These include (1) the research design, (2) the sampling methodology used to target observations, (3) measurement protocols used to collect data on the chosen observations, and (4) the analytical techniques applied to the collected data. For interested readers, Appendix 3 describes another important element, this being the extent to which the research project is primarily inductive vs. deductive. Additionally, Appendix 4 summarizes a real ESS research project in accordance with the terminology developed here to demonstrate how these concepts can be used to descriptively summarize any given project.

Research design

A research design is just that, a design that a research project employs to address its research questions. A taxonomy of research designs, or study types, is presented in the following list:

1. Experiment
   - 1a. True experiment
   - 1b. Quasi-experiment

2. Observational study
   - 2a. Natural experiment

[1] Environmental Studies Program, Dartmouth College

- 2b. Correlational study
- 2c. Case study
- 2d. Embedded case study

3. Synthetic study
- 3a. Literature review
- 3b. Case study meta-analysis
- 3c. Systematic review

The most common distinction among research designs is based on the extent to which they incorporate an experimental element (Bernard 2011). Essentially, an experiment is defined by its ability to isolate the effects of one or very few factors by controlling the variation of other variables across what are termed "control" and "treatment" groups. By randomly assigning observations to one of these two groups and then applying the treatment to the treatment group, the experimentalist minimizes the systematic difference between this group and the control group, other than the presence of the treatment, which can then be the only systematic explanation for differences in outcomes observed across the two groups. This increases their internal validity (to be discussed later) by limiting the possibility that alternative explanations could threaten causal inferences made in the subsequent analysis.

Within many scientific circles, there is a lot of rhetoric about the importance of experiments as the only fully valid way of establishing the importance of an independent variable. However, it is important to observe that their source of scientific strength is also the source of their primary weakness (Kauffman 2012). Experiments are comparatively poor at representing causal complexity and interaction effects. They are also usually conducted over relatively short time frames. As a result, they have comparatively low ecological validity, which, as I will discuss later, represents the extent to which we can generalize their results to real-world settings.

Experimentalists can incorporate additional complexity by conducting a factorial experiment, which essentially introduces additional independent variables, or treatments, into the design to examine interaction effects between these variables and the primary treatment variable. This quickly increases the number of distinct subgroups involved. For example, instead of having one control and one treatment group, an experiment with two independent factors as treatments would need to include one group for every possible combination (presence, absence) of each treatment, or four groups. Such designs can quickly become expensive and onerous, and it is unlikely that it will ever be feasible for them to mimic the complexity of real-world environments.

Recently, the field of development economics, which overlaps with ESS, has taken a distinctly experimental turn, most boldly embodied by the efforts of the researchers at the MIT Poverty Action Lab (http://www.povertyactionlab.org/). A primary outcome in this research has been the adoption of new agricultural technologies and knowledge, while common treatments include various financial devices and services. The field of ecology has also become highly experimental at small scales in the past several decades (Sagarin and Pauchard 2012).

The majority of analyses in ESS are nonexperimental. Those that are experimental tend to occur in the lab or in small field settings because large-scale experiments are usually untenable on account of resource limitations and the simple political infeasibility of the social changes that would be required. Such experimental ESS research is largely focused on the determinants of human behavior and cooperation because these affect environmental conditions (Ostrom 2006, Cox et al. 2009).

When true experiments are not feasible, researchers may turn to quasi-experiments, which are essentially experiments in which the assignment of observations to a treatment or control group is nonrandom. If neither true nor quasi-experiments are feasible, then the research can turn to an observational study. In such studies, "researchers do not attempt to change or manipulate variables to test their effects; they simply observe or measure things as they are in the 'natural,' unaltered world" (Remler and Ryzin 2011:355). Natural experiments are a cross of being experimental and observational studies. A natural experiment takes advantage of some naturally occurring treatment that is "applied" to one group, but not to another, highly similar, group. Such naturally occurring treatments can be environmental policies, or natural events such as floods and forest fires.

The final two observational categories (correlational study and case study) represent the great majority of data-oriented analysis in ESS. A correlational study involves a large enough number of observations to warrant a statistical or qualitative comparative analysis of many observations. Correlational studies are important in ESS for their ability to facilitate cross-observation generalizations, although for the most part their units of analysis are small, e.g., households, so generalizability across larger contexts is often still limited (Poteete et al. 2010). There are exceptions to this (see Gutiérrez et al. 2011).

In contrast, a case study (1) examines only one or few "cases" or observations of its primary unit of analysis, (2) involves multiple units of observation/data sources that it uses to draw inferences about each case, and (3) generally involves the measurement of more variables than observations, a situation that precludes statistical analysis in favor of qualitative analysis. ESS has traditionally involved many case studies (Lansing 1991, Ernst et al. 2013, Gilbert 2013, Kitamura and Clapp 2013). ESS case studies are generally distinctive for having some sort of social-ecological system as their unit of analysis, or the case that they examine.

When qualitative and quantitative approaches are combined the result is frequently referred to as a mixed methods approach. More specifically, a correlational analysis combined with a case study creates an embedded case study analysis (Yin 2014). The way in which these two research designs can be connected is based on the nested nature of reality: many entities exist in one-to-many relationships with each other because of the hierarchical nesting of social and biological life. Individuals lives within communities, which may exist within larger social units, and so on. Therefore an embedded case study may involve a qualitative analysis at the level of the case, but a researcher may collect data on enough observations of a nested unit of analysis to warrant a statistical analysis at this more disaggregated level. Ideally such analyses are complementary. For example, a researcher may compare four forests, but may do so by collecting tree-level data upon which he/she conducts a statistical analysis. If the researcher collects enough data at both levels (enough forests and trees in this case) he or she could conduct a multilevel statistical analysis at both

levels of organization. Embedded case studies are very common in ESS, with projects that target a particular social-ecological system examining this system by collecting data and conducting analyses of units embedded within this system (Acheson 1975, Vogt et al. 2006, Ayers and Kittinger 2014, Snorek et al. 2014).

Finally, there are synthetic studies that rely on aggregating secondary data and information from existing analyses, rather than collecting their own primary data. Primary data have yet to be collected by anyone, and need to be collected by the researcher. Secondary data are already present, at least in some form, although they may need to be reorganized, e.g., via some content analysis, in some way before the researcher can analyze them.

The most informal type of synthetic research design is simply a qualitative literature review that seeks to summarize the findings of a set of previous projects. A literature review is in fact frequently part of a larger project, but also can stand alone as a more hypothesis-driven type of research exercise (see Biggs et al. 2012). When it is systematized to enable a formal analysis of its own, it is referred more often as a case study meta-analysis or systematic review. In some fields the term meta-analysis refers specifically to the aggregation of multiple statistical studies, either by directly pooling quantitative data, or by pooling the results of multiple statistical analyses, e.g., pooling effect sizes. In ESS, it has tended more to mean the extraction of quantitative data from primarily qualitative case studies via a content analysis coding process (Rudel 2008). Such analyses generally either analyze the cases described in published studies (Cox et al. 2010, Evans et al. 2011, Cox 2014*a*) or use the studies themselves as their unit of analysis (Geist and Lambin 2002).

Systematic reviews initially became popular in public health-related fields (http://www.cochrane.org/), but have more recently spread to ESS-related disciplines (http://www.environmentalevidence.org/; Pullin and Stewart 2006, Bilotta et al. 2014). A systematic review involves the analysis of secondary data to explore the effects of a particular intervention, via either a qualitative narrative analysis or a statistical meta-analysis that combines quantitative data from the synthesized studies to calculate aggregate effect estimates of the intervention in question.

Sampling
A sampling strategy is generally the process of selecting a sample to study from a population of observations, say, trees in a forest. We collect a sample because we usually do not have the time or resources to collect data on an entire population of interest. The number of observations in a sample is referred to as the sample size. If we do collect an exhaustive sample of the entire population, this is called a census. Otherwise we want a representative nonexhaustive sample, or a sample with attributes that reflect the overall characteristics of the larger population. If we collect a biased sample for some reason, referred to as sampling bias, this means that we have somehow systematically identified or selected for one type of observation to sample over other types within the population. A biased sample is not representative of the larger population.

There are several dimensions that can be used to distinguish different types of sampling strategies and the sample types that result. First, a sample can be (1) random or nonrandom, and (2) stratified or nonstratified. Stratified samples can in turn by

proportional or nonproportional. These distinctions create the basic typology of sample types as shown in Table 1. A random sample is just that, random. These are seen as a gold standard of sorts because they are the best way to ensure that a sample is representative of the larger population.

Table 1. Sampling typology.

	Random	Nonrandom/purposive
Nonstratified	Simple random sample	Purposive sample
Stratified nonproportional	Random stratified sample	Nonrandom stratified sample
Stratified proportional	Random quota sample	Nonrandom quota sample

In practice, however, we frequently do not pursue a random sample for the following reasons: (1) because of the lack of a sampling frame, or a well-defined list of accessible observations, (2) the inability to collect data from randomly selected observations, (3) we are conducting a case study analysis of a small number of observations, and (4) we in fact do not want a representative sample. This final case occurs frequently when we are conducting key informant interviews, in which we target a select group of individuals and ask them about the system of which they are a part.

A sample can be stratified to account for large amounts of heterogeneity within the target population. This divides the population into strata along a dimension of particular theoretical importance, such as a demographic characteristic for a social study, or a biophysical gradient in an ecological one, and then collects a sample within each of these categories. Stratified samples can be stratified proportionally or nonproportionally. A proportional sample is one in which the numbers of observations within each strata are collected to be proportional to their presence in the larger population.

Purposive samples are nonrandom and nonstratified. These are guided by the researcher's judgment regarding what are the most appropriate observations to analyze with a sample. Common criteria for such judgments include (1) the expertise held by certain human respondents, in which case the sampling method is known as expert sampling, (2) the representativeness of the observations of the population, or conversely (3) the deviance of a particular case or observation from the population, if this deviance is to be explained.

Two types of samples that can be considered as subsets of purposive samples are (1) snowball samples and (2) convenience samples. A snowball sample is produced by a nonrandom sampling procedure in which initial observations are used to identify and access subsequent observations. This procedure is specific to human subject observations, who have the ability to identify other potential interview respondents. Given the importance of social networks, trust, and reputation in human interactions, this method is frequently the most, or only, feasible way to obtain access to many remote respondents. A convenience

sample is produced when the researcher selects observations almost entirely based on their availability, or the convenience of obtaining data from them.

Finally, a multistage sample is one in which an initial set of observations are selected, and then for these observations, multiple measurements are taken to ultimately produce many new observations per each original observation. This can happen in two basic ways: the original observations are selected, and then measurements are made for each observation at multiple periods in time to create what is known as panel data. Panel data enable both cross-sectional and longitudinal comparisons. The other way occurs when clusters of observations are first selected, and then a second sampling technique is applied to observations nested within these clusters. This is known as cluster sampling. A cluster could be a school, city, or forest, or parts of a forest, in which case each cluster is spatially explicit, and the nested units (say trees) would be physically nested within the cluster. In ecology this method is known as transect sampling, the clusters being called transects, and it is generally used to estimate the abundance and distribution of different species in an area. Cluster sampling and stratified sampling are similar, each involving an initial breakdown of the population based on some criterion. One main difference is that generally all strata are examined, while this is hardly ever the case for clusters.

Measurement
Measurement occurs in two steps (see Adcock and Collier 2001 for a more expansive discussion). These are depicted in Figure 1. In the first step, a concept is operationalized into a variable and assigned a particular range. In the second step, this variable is measured via a data collection method and/or instrument. One necessary part of operationalization is deciding on the level of measurement for a variable. The following presents a very common scale that is used to classify different levels of measurement. Note that I include qualitative as a category here.

Fig. 1. Measurement steps.
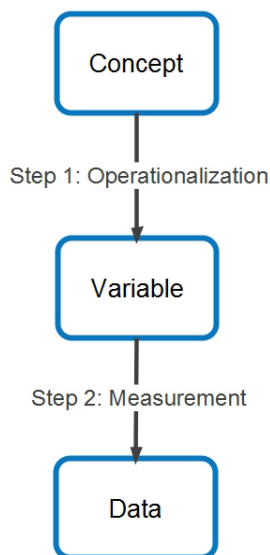
- Qualitative variable: No quantitative structure to the data. Allowable values are any text.

- Categorical variable: Divides possible values of a variable into discrete categories that cannot be meaningfully ordered (e.g., red, green).

- Ordinal variable: Divides possible values of a variable into discrete categories that can be meaningfully ordered (e.g., small, large). The absolute distance between values is not meaningful, even if there is a difference implied by the ordering.

- Interval/ratio variable: Interval variables can take on numeric values, the distances between which are consistently meaningful (e.g., 10pm is one hour after 9pm, which is also one hour after 8pm). Ratio variables are distinguished from interval by having a meaningful zero value. To be used meaningfully these variables must be assigned a unit of measurement (e.g., meters, years) that is being counted.

It is also important to note that there need not be a one-to-one relationship between a concept and a variable. In fact, it can be useful to have multiple variables for a given concept to ensure that the results of an analysis are not idiosyncratic to a particular type of operationalization. The process of checking one way of measuring a concept against others is referred to as triangulation (Yin 2014). ESS frequently involves both nonhuman observations and human observations, or respondents, which produce objective data vs. subjective data, each of which can be used to check the validity of the other. In the study of forest management, for example, both objective and subjective data have been commonly used to measure the concept of forest health. Nagendra and Ostrom (2011) show that in some cases these give comparable evaluations of forest condition.

Once concepts have been operationalized and observations sampled, the researcher needs to decide what measurement strategy he or she will employ to make measurements on the selected observations. The first set of such options involves soliciting responses from human subjects regarding written or spoken questions (see Fowler 2009 for a discussion of surveying techniques). Informal interviews are in fact not interviews the way most people think of the term, but are essentially informal conversations that a researcher has with human subjects during the course of, say, a fieldwork season. Unstructured interviews are also rather informal, but actually occur as interviews, and are understood as such by both the subject and the researcher. Focus groups are a bit like unstructured interviews but occur in group settings, with multiple respondents at once. Semistructured interviews involve the use of an interview guide to make sure that certain subjects are covered and certain questions are asked. Finally, structured interviews are done with the help of a written questionnaire that the researcher fills out as the subject responds to the questions it contains. Self-administered surveys are like questionnaires, but are filled out by the subjects without the presence of the researcher. As this list indicates, there is a range of formality a researcher might impose in the data collection process, and in the research design in general, and the researcher faces a trade-off between the goal of obtaining data that have a high probability of being consistent across observations and thus answering the original research questions if analyzed

appropriately, and the goal of adapting to changing circumstances as the project proceeds.

Next, Participant observation involves researchers becoming actively involved in a particular system to the extent that they gain knowledge of everyday and subtle causal complexities of that particular system which are hard to otherwise capture and can be difficult to generalize. It usually involves fieldwork in the study site and extensive note-taking. It has been most commonly practiced by anthropologists (see Bernard 2011), although Sagarin and Pauchard (2012) argue that a similar approach is critically important for the discipline of ecology, to complement more experimentally manipulative approaches.

Similarly, but with less direct involvement and engagement, direct observation can be conducted with any type of subject, and it involves the researcher directly observing the behavior of the subject, which is usually a live organism, if not a human. In comparison with participant observation this is generally seen as a more quantitative approach, with the aim frequently being to count frequencies of certain behaviors or otherwise make quantitative measurements based on what is being observed during a particular period of time (Guest et al. 2012).

Direct observation can be aided by technology, in which case the researcher is also using in-person instrumentation, which is the use of a technological device to record data about an environment, or to take samples from this environment. This technology could be a recorder or a video camera, or in the case of biophysical scientists it could be any range of data collection instruments. Finally, remote instrumentation likewise involves the use of some technology to collect the data, but does not involve or require the presence of the researcher.

Analysis

Qualitative vs. quantitative analysis

Analysis is the process of describing and then making inferences based on a set of data. To make an inference means to combine data with something else, say a set of assumptions or theories or more general knowledge, and draw a conclusion that goes beyond what the data themselves present. The most basic distinction we can make between different types of analysis is to classify them as either quantitative or qualitative. A quantitative analysis is mostly associated with correlational studies, and it involves the examination of relationships between quantitative (categorical, ordinal, and interval/ratio) variables, whereas a qualitative analysis does not deal with numerical data. A qualitative analysis [1] is generally done as part of a case study, and instead involves the construction of inferences from nonnumerical data sources.

It is possible to transform qualitative data into quantitative data via what is known as content analysis, which has been done by a fair number of ESS scholars (Delgado et al. 2009). There are many textbooks (Neuendorf 2002) and software packages available for this (Atlas.ti and NVivo). Importantly, the qualitative medium need not be in the form of text as a text variable, but can be any medium, such as videos or direct in-person observations. In addition to content analysis, quantitative data can be used to produce other quantitative data, in two basic ways: (1) transforming a variable at a "higher" level of measurement to a "lower" level of measurement, e.g., from an interval to an ordinal variable, and (2) calculating averages to summarize sets of observations and produce a new, more succinct, dataset.

Types of quantitative analysis

Statistical analysis is the primary analytical tool in many scientific fields, and it is probably the most commonly used quantitatively analytical tool in ESS and related disciplines (Agrawal and Yadama 1997, Hayes 2006, Lorent et al. 2009, Persha and Blomley 2009, K.C. 2013). Grounded in probability theory, statistical analysis primarily involves the application of calculations to a dataset to (1) describe a sample, (2) make statistical inferences about a population by constructing confidence intervals and conducting hypotheses tests, and (3) estimate the magnitude of associations between variables. Much of this is done via statistical modeling. There are numerous textbooks and online resources for introductory to advanced statistics. The most popular statistical software packages include SPSS, SAS, JMP, Stata, and R.

Network analysis, most commonly conducted by ESS researchers as social network analysis (SNA; Cohen et al. 2012), involves the conceptualization of a network of nodes connected by a series of links (see Borgatti et al. 2009 for a summary of the practice of SNA). In ESS, the nodes are most commonly social actors, such as resource users or managers, but this is not necessarily the case. The links connecting different nodes are generally identified based on their theoretical importance. Examples of links "include routine interactions among actors regarding environmental policy issues (Schneider et al. 2003), exchanges of information regarding a natural resource, fishing gear exchanges, and social support activities (Bodin and Crona 2008), and exchanging ideas and funds (Lauber et al. 2008)" (Cox 2014b:312). There are numerous social network software applications (e.g., Ucinet), and R actually includes several network analysis packages.

Within ESS, social network analysis is one of the fastest growing types of analysis (Prell and Bodin 2011, Isaac et al. 2014). However, a primary limitation of network analysis in ESS is the resource-intensity of the data collection required. This is because network analysis typically requires one to collect data from every node in a network, which is not cheap or easy, particularly if the nodes live in a remote community that is not very small.

Geographic information systems (GIS) and remote sensing (RS) refer to suites of hardware, software, and analytical techniques designed to collect and analyze spatial environmental data, that is, data describing observations with spatial coordinates. In fieldwork-based ESS these are frequently complemented by Global Position System (GPS) units to collect local spatial data (see Brondizio and Roy Chowdhury (2010) for a summary of the application of these technologies and techniques in ESS).

GIS operations are applied primarily to vector-based data, or data expressed as discrete points, lines, and polygons. The most popular GIS software package is ArcGIS. The term remote sensing most accurately refers to the remote detection of radiation that is reflected or emitted by different objects or land surfaces. This process produces images made up of cells, or pixels. RS operations are then applied to raster-based data, or images, primarily in the form of satellite imagery, e.g., via the Landsat program, and aerial photography. RS technologies and operations are used extensively by scholars studying land use and land cover change (Lambin et al. 2001, Vogt et al. 2006). Many environments and natural resources that cannot be easily monitored otherwise can frequently be feasibly measured with the aid of remote sensing technologies (Andersson et al. 2008).

Given the explosion in availability of spatial data, these tools are likely to only increase in importance in ESS. Several common RS applications include ERDAS IMAGINE, ENVI, and IDRISI.

Qualitative comparative analysis (QCA) is an approach that is designed to compare quantitative data on a medium to small number of observations without the normal emphasis on effect sizes and formal hypothesis testing that dominates the statistical approach. It places more of an emphasis on interaction effects and conditions of causal necessity and sufficiency than the normal statistical approach does. It has been championed primarily by Charles Ragin (1987, 2000).

QCA is essentially built on the notions of necessity and sufficiency combined with the causal logic established by John Stuart Mill. Without diving too much into the details, which can easily be found online, this logic enables us to examine Table 2 and conclude that, of the four aquaculture farms shown there, the outcome can be explained by the use of tilapia is the primary fish. This is because (1) it is the only potential cause that is present in all successful cases, meaning that no other cause is necessary, and it is sufficient, and (2) it is the only cause that is absent in all unsuccessful cases, meaning that no other cause is sufficient, and it is necessary. It is important to note that although supposedly not as rigorous as statistical analyses due to the small number of observations involved, the intuition here is used effectively by humans in all walks of life. Ragin's QCA method essentially elaborates on the logic proposed by Mill, and has been by several researchers studying social-ecological relationships (Rudel 2008, Basurto 2013, Pahl-Wostl and Knieper 2014).

**Table 2**. Sample data for a qualitative comparison.

Case	Fish used	Country	Size	Outcome
1	Tilapia	Indonesia	Large	Succeeded
2	Tilapia	Taiwan	Small	Succeeded
3	Salmon	Taiwan	Large	Failed
4	Salmon	Indonesia	Small	Failed

Types of qualitative analysis
Each of the previously mentioned analytical strategies relies on analyzing quantitative data, which are data that are produced ultimately by breaking a continuous world up into discrete chunks and counting and measuring features of these chunks so that they can be compared. Whenever this is done, some information and complexity is inevitably obscured or lost. Ideally, that gap is where purely qualitative analysis steps in.

There are several types of qualitative analysis that ESS scholars frequently conduct. These include (1) thick description, (2) narrative path analysis, (3) qualitative models that depict relationships of the sort shown in Figure 2, (4) congruence testing, and (5) inductive theorizing. Although I describe them as types of qualitative analysis, each is frequently integrated with quantitative data analysis and numerical methods (and vice versa). Each of these methods is also used frequently with the others.

Fig. 2. Types of variable relationships.
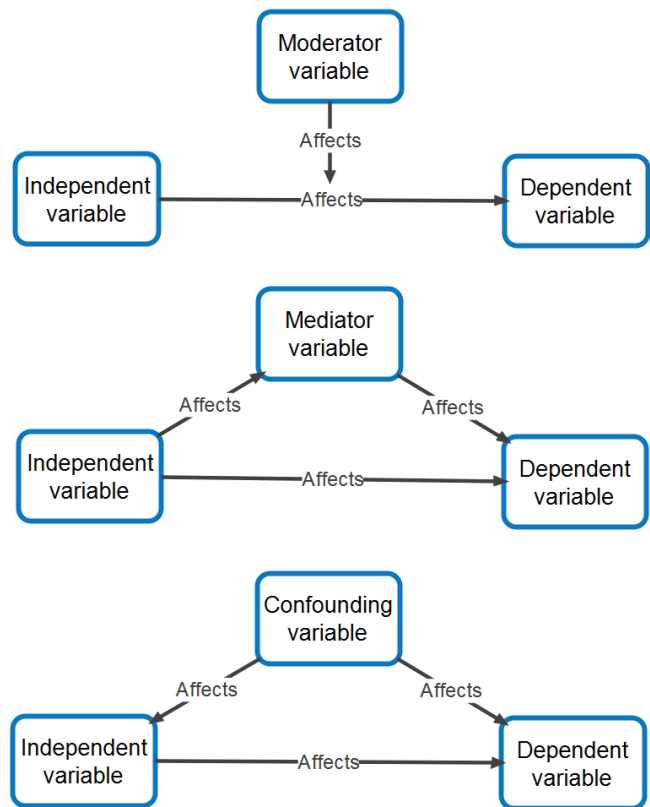
Thick description was popularized as an analytical strategy by Clifford Geertz (1973), who played a prominent role in the development of qualitative ESS (Geertz 1959, 1980). Thick description is characterized, and contrasted with thin description, by Denzin (1989:33) by having the following functions: "(1) It gives the context of an act; (2) it states the intentions and meanings that organize the action; (3) it traces the evolution and development of the act; (4) it presents the action as a text that can then be interpreted. A thin description simply reports facts, independent of intentions or the circumstances that surround an action." Some form of thick description is frequently a central part of an ESS case study.

Often what ESS scholars describe about a case is the historical development of the case's ecological or social components, and the relationships between these. The scientific field that uses this approach extensively has frequently been referred to as environmental history (Diamond 1997, 2005, Cronon 2011). Such studies exemplify the standard features of a case study, particularly in that they tend to rely on many sources of evidence to support the theories they promote.

Many concepts such as "path dependence" that are prevalent in ESS are explicitly temporal and historical. This is reflected in the emphasis that many case studies place on unpacking historical dynamics and, in turn, extrapolating from these to explore future scenarios for the paths that a system might take in the future.

These narrative path analyses frequently break out the history of a system into discrete chunks of time, and rely on a scientific framework to theoretically inform the characterization of these time periods and the relationships between them (Brown et al. 2013, Boonstra and de Boer 2014, Câmpeanu and Fazey 2014, Cody et al. 2015). In such analyses, the cumulative impacts of historical events and the resulting path-dependency of the current situation are often emphasized. In case study methods textbooks (see George and Bennett 2005), this has been referred to as process-tracing, and several ESS scholars have begun to implement this as an analytical strategy (Fleischman 2014).

More formally, a scientist can explore the structure of a system with a qualitative model. These go by many names, including access mapping of commodity chains (Ribot 1998), influence or logic models, arrow diagrams (Homer-Dixon 2010), path diagrams (Fleischman 2014), impact or linkage diagrams, and causal loop diagrams (Sendzimir et al. 2011). Qualitative models are generally represented as "box-and-arrow" diagrams, and the common denominator of each such model is that it breaks a system or process down into a set of constituent objects (boxes) and the directed relationships among these objects (arrows). Such models are in fact ubiquitous in the ESS literature as a way of developing an understanding of the system in question, particularly in case study analyses (Neudoerffer et al. 2005, Homer-Dixon 2010, Alberti et al. 2011, Österblom and Sumaila 2011, Fazey et al. 2011, Downing et al. 2014). The objects can take a variety of forms, and increasingly these objects are closely tied to elements of social-ecological systems identified in supporting social-ecological frameworks (see Villamayor-Tomas et al. 2014 and the associated special issue). The presentation of a qualitative model is most often accompanied by a narrative that describes the components and how they represent important dynamics within the target system or process. Such models are frequently developed collaboratively with local resource users and other participants to explore scenarios for future change in the target system (Marín et al. 2008, Delgado et al. 2009, Guimarães et al. 2013).

Qualitative work can also be explicitly comparative. Some scholars have qualitatively compared small numbers of cases informally, through simple nonquantitative comparisons that do not attempt to formally characterize the relationships among the variables. These may be done by making comparisons of the sort shown in Table 2, although not necessarily with the intent to establish necessity or sufficiency (Klooster 2000). Longitudinally and cross-sectionally comparative case studies in ESS are also frequently done by applying a formal scientific framework to different cases or to different time periods within one case (Pahl-Wostl et al. 2013, Barnett and Anderies 2014).

Some ESS case studies compare the findings of a case with one or more theories that describe expected findings in the form of a configuration of values for a set of variables (Fleischman 2014, Fleischman et al. 2014). This has been referred to as congruence testing (George and Bennett 2005) or "pattern matching" (Yin 2014). This is a method that examines the extent to which the main features of a case study are congruent with the hypotheses generated for this case by a theory. For some it has been standard wisdom that single cases cannot do much to confirm a theory (but may do much to contradict from a Popperian perspective). The extent to which a theory can be supported by a single case study depends in large part on whether or not the results of the case study can be explained by alternative theories or not. If the properties of a case are highly congruent with the expectations generated for it by a certain theory, and there are no alternative theories with which the case is also congruent, we may conclude that the results of the case are very unlikely except if they are explained by the theory, and this can be seen as strong evidence in support of the theory. As George and Bennett (2005:117) put it: "an explanation of a case is more convincing if it is more unique, or if the outcome it predicts could not have been expected from the best rival theory available." This is made possible by having well-established theories with many independent predictions for a case. The more predictions, the more points of congruence can potentially be established and the less likely it is that a confirmatory case could be explained without the theory in question. "This process relies on Bayesian logic-the more unique and unexpected the new evidence, the greater its corroborative power" (George and Bennett 2005:219).

Finally, the inverse of the deductive process of theory testing (see Appendix 2) is inductive theory-building, or simply theorizing. This is the essentially qualitative process of inferring generalized, usually causal, relationships from a set of data, be it quantitative or qualitative. In addition to being an important part of quantitative studies, this is frequently done through single and comparative case studies, with a general pattern being that an author introduces a theory, and then presents one or more cases that exemplify this theory (Holling and Meffe 1996, Scott 1998, Robbins 2000). A standard observation here is that the same data used to develop a theory inductively cannot be used to then test that theory through some congruence procedure as just discussed.

EVALUATIVE CRITERIA

Interest/practical importance

A research project needs to be of some interest or importance to a definable audience. Frequently this audience is a community of scientists that works on the same family of research questions with broadly similar methods. One of the most common ways for a research project to be of interest to such a community is for it to identify a particular gap in the knowledge that has been established by that community. Other communities can be considered as well, such as environmental practitioners or a client that has contracted with a researcher to conduct a project, such as the United States Agency for International Development.

Feasibility

Feasibility is simply the extent to which a research project can be effectively conducted given the resources available to the researcher(s) considering the project. These resources can include time, money, equipment, and human capital and skills. These resources can of course be developed and obtained for the purposes of a particular research project, but unrealistic assumptions regarding the ability to do this, particularly in a short amount of time, should be avoided.

Ethicality

The relationship between research and ethical considerations has a very long and sometimes checkered history. I will not delve very much into these issues here, except to say that any researcher should of course consider whether or not the research they

propose would unduly harm, or place at risk, vulnerable or disempowered populations.

Internal validity

Internal validity is the extent to which a causal inference made by an author correctly reflects a causal relationship between two or more variables. Two perspectives have predominated in discussions on the best method for understanding a causal relationship: one is the experimental/comparative perspective, and the other is the mechanistic/noncomparative perspective. The former views causation as being primarily detectible through ideally experimentally controlled comparisons, and secondarily through less controlled comparisons. Without comparing one observation with others, so the thinking goes, we have little leverage to gauge whether changing one variable will affect the value of others. This problem is often expressed in terms of counterfactuals. A counterfactual is an alternative scenario to which a realized scenario is compared to evaluate the significance of a causal factor that changes across the scenarios. Our interpretation of the causal significance of any factor depends critically on the most likely counterfactual to which the realized scenario is compared. In climate change and forest-related policies, this is expressed as the need to establish the additionality of a policy (Angelsen 2008). Additionality, most broadly, is the difference between the value of an outcome after the implementation of a policy and its value in a counterfactual scenario in which the policy is not enacted.

Establishing a counterfactual depends on finding observations that represent a primary scenario and others that reflect an alternative scenario, which, in a nonexperimental observational study, we can only do by making cross-sectional or longitudinal comparisons. For example, we might ask what the effects of a reforestation policy are on forest cover, and to do so longitudinally we would need to take into account the rates of reforestation or deforestation occurring prior to the implementation of the policy (this is referred to as establishing a baseline). The effect of the policy would be the difference between this rate and the rate sometime after the policy was implemented. To conclude, from the comparative perspective, we want to ask the following questions as ways to establish causal inferences:

1. Do the cause and effect covary with each other?

   · 1a. Is there a temporal comparison that can be made between repeated observations of the same units?

   · 1b. Is there a cross-sectional comparison that can be made between separate units?

2. Did the stipulated cause occur prior to the effect (temporal antecedence)?

From the noncomparative perspective, our primary concern is less about establishing counterfactuals, and more about whether or not we have a theoretical/mechanistic account that explains how or why such a covariation would occur. This perspective follows the aphorism that "correlation doesn't equal causation," and argues that we cannot infer causality without a notion of mechanism, or theory that explains how one factor actually affects another.

In addition to identifying sources of effective causal inference, the literature on social science research methods has established

possible threats to such inferences.[2] There are two main types of such threats. The first is that an alternative narrative or story could explain the patterns, or covariations, we find in the data. The second source is that the data might not tell the whole story, and that the effects of an independent variable (IV) on a dependent variable (DV) are not fully accounted for.

First, there are two main ways in which a covariation between two variables might not support an inference that one causes the other. The first such alternative explanation is something we have discussed before: endogeneity or the possibility to that the supposed DV in fact causes an IV to change, rather than, or in addition to, the other way round. The data themselves may not tell us which direction the causal arrow points between two variables.

The second such threat involves sources of spurious relationships, in which an association between two variables is not actually indicative of a causal relationship, or at least the strength of this relationship. Inferring an effect where there is none is an example of committing a type 1 error. A type 2 error would involve the opposite: inferring the lack of a relationship where there really is a relationship. The primary source of a spurious relationship is a confounding variable. Confounding occurs when an independent factor, A, is found to correlate with a dependent factor B, when in actuality both are caused by a third phenomenon, the confounding variable C. This is shown at the bottom of Figure 2. For example, we might find that communities that use private property rights to manage a resource are more productive than those that use common property. However, if in fact common property is an adaptation to high environmental and economic scarcity (Barbanell 2001), then it could be that a third variable, resource scarcity, explains the presence of both in some systems, and in so doing explains the association between property rights systems and productivity.

The final threats are all ways in which an intervention could affect areas outside of its target area, and thereby complicate causal inference (see also Lambin and Meyfroidt 2011). First, we have diffusion of treatments or spillover effects. The idea here is that these effects need to be taken into account to accurately summarize the effects of the intervention. Each of these terms reflects either a positive externality, in which the effects of the intervention on untreated units is positive, or a negative externality, in which the effects are negative.

Regarding positive effects, in the study of marine protected areas, it is often hypothesized that the benefits for biota within the protected areas might spill over, or positively benefit, surrounding areas (McClanahan and Mangi 2000). Another term used to describe this phenomenon is diffusion ("diffusion of treatment" or "technological diffusion"). This has most often been used to describe the potential for the effects of a technology to spread or otherwise benefit subjects other than those provided the technology. Historically this has been most commonly applied to the potential spread of agricultural technologies (Hayami 1974). For example, if we were to provide a group of farmers in Kenya with cell phones that they could use to obtain weather information, it is possible that their neighbors could also benefit from this by talking with them. Comparing a treatment and control group in this context could lead to us undervalue the benefits of the phones. From a policy perspective, it is probably desirable to encourage these positive externalities. However, if

unaccounted for analytically, they may cause a researcher to undervalue the positive effects of a policy.

Of course, new technologies and governance arrangements can have negative external effects as well. These could be simply negative externalities, such as when providing farmers with new pesticide technologies creates costs for other farmers who have not been provided them (Wilson and Tisdell 2001). Aside from this, the term leakage has become popular as a way of referring to situations in which the behavior or outcome that is forbidden in one area subsequently leaks out to other areas. Leakage has been widely discussed as a possible result of policies designed to mitigate climate change by preventing deforestation (Fahey et al. 2010). By preventing deforestation in one area, such policies might heighten incentives to cut down trees in other areas.

Finally, we have what Lambin and Meyfroidt (2011) refer to as a cascade effect, or an indirect effect that an intervention has on an outcome in areas other than that for which the intervention was intended. These are similar to spillover effects but are less direct. As a prominent example, the development of biofuels in developed countries has tended to incentivize agricultural expansion in developing countries via increased crop prices, thereby potentially negating much of the positive effects that increased biofuel production might have for climate change mitigation (Lapola et al. 2010).

External validity and generalization
External validity is the ability to generalize the findings of a study to other contexts, and could also be called inductive validity. There are two types I discuss: (1) generalizability from a sample to a population, and (2) generalizability from a population to other populations.

Each of these involves generalization from one set of observations to a larger set. The first of these has been the most frequently discussed in methods texts, and is the motivation behind the common desire to obtain a representative sample in moderate to large-n research projects. The intuition is that if we want to tell a scientific story about a population, and we can only really examine our sample, then we want to be able to generalize our findings about our sample to the population of interest.

The second type of external validity concerns whether or not the findings from a study can in fact be generalized to populations other than the primary population of interest. A study might take a sample of 300 Kenyan farmers out of several thousand farmers in a particular area, and if this is done randomly, or purposively in a way to achieve representativeness, then the findings of the study may be generalizable to the larger population of several thousand. However, we would still face the question of whether or not such findings would be generalizable to all Kenyan farmers, or all farmers in East Africa. Another example comes from work on toxicology, in which experiments on rats are frequently conducted to examine whether different substances are overly toxic. One possible criticism of this approach is the potential lack of generalizability from the population studied in these experiments to the population that we presumably care more about (humans).

Deductive validity and ecological validity
The next two types of validity deal with a combination of causal inference (internal validity) and generalization (external validity).

Each results at least in part from interaction effects that complicate causal inference. The first of these types is what I will call deductive validity. Rather than applying findings of a sample to a larger population or to other populations, deduction involves the application of findings regarding a sample to a smaller sample, or subgroup, of that sample. This process, rather than being inductive, is better described as a process of deduction, or applying general findings to a specific case.

Interaction effects are the primary threat to deductive validity. Referring back to Figure 2, if the IV is an intervention and the moderating variable is a distinguishing feature of two sets of cases in our sample, say, low development and high development countries, then we do not necessarily want to apply the same intervention to both areas of high and low development. Assuming that the policy would be effective in low-developed areas based on a sample-level positive effect could be deductively invalid, and doing so is sometimes referred to as committing the ecological fallacy.

Ecological validity is the extent to which the findings of a research project are generalizable to relatively uncontrolled contexts or environments. In ESS, generalizability to real-life contexts is invaluable, given the potential applicability of the work to the resolution of socio-environmental problems. A lack of ecological validity is of particular concern for experimental work, because experiments are defined as much as anything by the high degree of control that is exerted to isolate the effects of one particular factor on an outcome of interest. Because of this control, the findings from an experiment may not hold in a natural environment. For example, experiments are very popular in agricultural research, where highly similar plots of land are separated into control and treatments groups, and a treatment is applied to one group but not the other. This is done to accomplish the experimental goal of isolating the importance of one particular factor by ensuring that no other potential independent factors vary across the control and treatment groups.

This type of experimental control can come at the expense of being able to generalize these findings to much less controlled environments, where many variables may interact with those included in the experiment. Low generalizability also can result from the limited time-frame of an experiment. In such situations it can be difficult to establish permanence, or the extent to which relationships that hold over the short-run also hold in the long run after the initial intervention has run its course. For example, Hanna et al. (2012) find that when implemented in real-world settings over a substantial period of time, the benefits of improved cooking stoves in developing countries are not nearly as high as they are predicted to be by laboratory studies.

Measurement validity and reliability
Measurement validity refers to the accuracy of each of the steps depicted in Figure 1 (operationalization and measurement). The protocol by which a concept is operationalized as a variable determines how faithfully the resulting variable reflects what is meant by the theoretically important concept. In the second step, error may come from (1) the data collection instrument itself, whether it is a written questionnaire or a physical device that records environmental data, or (2) the use of the instrument.

The measurement of concepts in ESS is tricky given the nature of the concepts used. Young et al. (2006) comment on this issue,

particularly with respect to dependent variables such as sustainability and adaptive capacity, which are concepts that are demonstrated by a system over time. When measuring ecological outcomes, we have to acknowledge that an absolute value for one area may be a poor outcome for one area but a good one in another, depending on many environmental conditions. As a result, it is frequently necessary to develop indicators of ecological outcomes that are normalized by time, via previous measurements within the same system, or by ecologically comparative settings. The DV in this case becomes a comparison between different points in time in one system or between that system and similar systems.

Reliability generally refers to the consistency of the second step just described. Such measurements are generally taken with the aid of a data collection instrument and/or protocol. The implementation of this protocol is ideally highly consistent, or reliable, across distinct implementations. For research that involves many observations this is particularly important because without such reliability we cannot be sure that the measurement of variables on each observation is actually producing consistent and thus comparable data.

Statistical validity
This type of validity is only applicable to projects that conduct statistical analyses. It is sometimes confused with internal validity, but is rather distinct. Although internal validity deals with causal inference, statistical conclusion validity deals with conclusions that are made about the actual data. For example, we might examine whether two variables correlate with each other or not and conclude that they do. This is a conclusion whose statistical validity we should concern ourselves with. Whether or not this correlation actually implies a causal relationship between the two is a separate issue, and that is where internal validity comes in. Statistical conclusion validity is frequently determined by the extent to which important assumptions made by statistical methods and models are true. A critical part of a statistical analysis is to test for violations of any such assumptions and to correct for them if possible.

RELATIONSHIPS AMONG THE CRITERIA
The researcher should not view his or her task as maximizing every one of the criteria discussed above because there are some trade-offs that frequently must be dealt with. Most obviously, there is a trade-off between the feasibility of a project and many of the other criteria because improving a research design and implementation generally involves additional expenses. When considering a way to improve a research project along one of the criteria, researchers should consider what resources would be required and how demanding this would be on the resources available.

Another important relationship to consider is that between internal and ecological validity. One of the most popular discussions of this relationship describes it as a trade-off for the following reason: as we apply increasing amounts of control to minimize variation of the great majority of variables, we increase internal validity essentially by controlling for alternative explanations, but we decrease ecological validity. The more controlled a setting is, the less likely it is to be representative of a real-world setting (lowering ecological validity). Moreover, if we really only allow one independent factor to vary, then deductive validity may be threatened as well because we cannot explore the extent to which additional factors might interact with this variable to affect important outcomes. A factorial design can hypothetically account for all of this, but factorial designs can become very costly to implement if one tries to examine more than one or two interaction effects.

Other than this negative relationship, internal validity is generally increased when a researcher increases measurement, external, and statistical validity. The more external validity there is, the more representative a sample is of a population. If a causal inference is made with respect to the sample, external validity and internal validity become almost the same thing because we are then concerning ourselves with generalizing this causal inference to the population. Measurement validity supports internal validity in the sense that we cannot make valid causal inferences if we do not measure our concepts correctly. Finally, for statistical analyses, the causal inferences we make are based on statistical conclusions, which must be in compliance with the necessary statistical assumptions. When we are conducting inferential statistics, as we usually are, then there is also a similarly positive relationship between statistical and external validity.

CONCLUSIONS
It has not been my intent here to promote one particular way of conducting ESS. As mentioned in the introduction, one of the hallmarks of this field is its interdisciplinarity. This is widely accepted as a strength, based on the premise that the complex systems that ESS scholars analyze require the application of multiple methods to be scientifically understood. At the same time, I believe that some of the concepts and issues presented here are faced by many if not most ESS scholars. Such scholars must decide, for example, how much to emphasize deductive vs. inductive approaches in their research and consider the political as well as practical implications of this decision. They need to consider what scientific values (internal validity, external validity) they strive for in their work, and whether they face trade-offs between these. A highly diverse ESS should be compatible with a systematic recording and documenting of that diversity, which is one of the goals that a rubric like this may achieve. [1] For those readers interested in qualitative and case-study-based research, see George and Bennett (2005) and Tong et al. (2007).
[2] See Shadish et al. (2002) for a list that is oriented more toward an experimentalist perspective with individuals as the primary unit of analysis.

Responses to this article can be read online at:
http://www.ecologyandsociety.org/issues/responses.
php/7400

LITERATURE CITED

Acheson, J. M. 1975. The lobster fiefs: economic and ecological effects of territoriality in the Maine lobster industry. *Human Ecology* 3(3):183-207. http://dx.doi.org/10.1007/BF01531640

Adcock, R., and D. Collier. 2001. Measurement validity: a shared standard for qualitative and quantitative research. *American Political Science Review* 95(03):529-546. http://dx.doi.org/10.1017/S0003055401003100

Agrawal, A. 2001. Common property institutions and sustainable governance of resources. *World Development* 29(10):1649-1672. http://dx.doi.org/10.1016/S0305-750X(01)00063-8

Agrawal, A., and G. N. Yadama. 1997. How do local institutions mediate market and population pressures on resources? Forest *Panchayats* in Kumaon, India. *Development and Change* 28 (3):435-465. http://dx.doi.org/10.1111/1467-7660.00050

Alberti, M., H. Asbjornsen, L. A. Baker, N. Brozovic, L. E. Drinkwater, S. A. Drzyzga, C. A. Jantz, J. Fragoso, D. S. Holland, T. (Tim) A. Kohler, J. (Jack) Liu, W. J. McConnell, H. D. G. Maschner, J. D. A. Millington, M. Monticino, G. Podestá, R. G. Pontius, C. L. Redman, N. J. Reo, D. Sailor, and G. Urquhart. 2011. Research on coupled human and natural systems (CHANS): approach, challenges, and strategies. *Bulletin of the Ecological Society of America* 92(2):218-228. http://dx.doi.org/10.1890/0012-9623-92.2.218

Anderies, J. M., M. A. Janssen, and E. Ostrom. 2004. A framework to analyze the robustness of social-ecological systems from an institutional perspective. *Ecology and Society* 9(1): 18. [online] URL: http://www.ecologyandsociety.org/vol9/iss1/art18/.

Andersson, K., T. P. Evans, and K. R. Richards. 2008. National forest carbon inventories: policy needs and assessment capacity. *Climatic Change* 93(1-2):69-101. http://dx.doi.org/10.1007/s10584-008-9526-6

Angelsen, A. 2008. *Moving ahead with REDD: issues, options and implications*. CIFOR, Bogor, Indonesia.

Augerot, X., and C. L. Smith. 2010. Comparative resilience in five North Pacific regional salmon fisheries. *Ecology and Society* 15 (2): 3. [online] URL: http://www.ecologyandsociety.org/vol15/iss2/art3/

Ayers, A. L., and J. N. Kittinger. 2014. Emergence of co-management governance for Hawai'i coral reef fisheries. *Global Environmental Change* 28:251-262. http://dx.doi.org/10.1016/j.gloenvcha.2014.07.006

Barbanell, E. 2001. *Common-property arrangements and scarce resources: water in the American West*. Praeger, Westport, Connecticut, USA.

Barnett, A. J., and J. M. Anderies. 2014. Weak feedbacks, governance mismatches, and the robustness of social-ecological systems: an analysis of the Southwest Nova Scotia lobster fishery with comparison to Maine. *Ecology and Society* 19(4): 39. http://dx.doi.org/10.5751/ES-06714-190439

Basurto, X. 2013. Linking multi-level governance to local common-pool resource theory using fuzzy-set qualitative comparative analysis: insights from twenty years of biodiversity conservation in Costa Rica. *Global Environmental Change* 23 (3):573-587. http://dx.doi.org/10.1016/j.gloenvcha.2013.02.011

Berkes, F., T. P. Hughes, R. S. Steneck, J. A. Wilson, D. R. Bellwood, B. Crona, C. Folke, L. H. Gunderson, H. M. Leslie, J. Norberg, M. Nyström, P. Olsson, H. Österblom, M. Scheffer, and B. Worm. 2006. Globalization, roving bandits, and marine resources. *Science* 311(5767):1557-1558. http://dx.doi.org/10.1126/science.1122804

Bernard, H. R. 2011. *Research methods in anthropology*. Altamira, Plymouth, UK.

Biggs, R., M. Schlüter, D. Biggs, E. L. Bohensky, S. BurnSilver, G. Cundill, V. Dakos, T. M. Daw, L. S. Evans, K. Kotschy, A. M. Leitch, C. Meek, A. Quinlan, C. Raudsepp-Hearne, M. D. Robards, M. L. Schoon, L. Schultz, and P. C. West. 2012. Toward principles for enhancing the resilience of ecosystem services. *Annual Review of Environment and Resources* 37:421-448. http://dx.doi.org/10.1146/annurev-environ-051211-123836

Bilotta, G. S., A. M. Milner, and I. Boyd. 2014. On the use of systematic reviews to inform environmental policies. *Environmental Science & Policy* 42:67-77. http://dx.doi.org/10.1016/j.envsci.2014.05.010

Binder, C. R., J. Hinkel, P. W. G. Bots, and C. Pahl-Wostl. 2013. Comparison of frameworks for analyzing social-ecological systems. *Ecology and Society* 18(4): 26. http://dx.doi.org/10.5751/ES-05551-180426

Bodin, Ö., and B. I. Crona. 2008. Management of natural resources at the community level: exploring the role of social capital and leadership in a rural fishing community. *World Development* 36(12):2763-2779. http://dx.doi.org/10.1016/j.worlddev.2007.12.002

Boonstra, W. J., and F. W. de Boer. 2014. The historical dynamics of social-ecological traps. *Ambio* 43(3):260-274. http://dx.doi.org/10.1007/s13280-013-0419-1

Borgatti, S. P., A. Mehra, D. J. Brass, and G. Labianca. 2009. Network analysis in the social sciences. *Science* 323 (5916):892-895. http://dx.doi.org/10.1126/science.1165821

Brondizio, E. S., and R. Roy Chowdhury. 2010. Spatiotemporal methodologies in environmental anthropology: geographic information systems, remote sensing, landscape changes, and local knowledge. Pages 266-298 *in* I. Vaccaro, E. A. Smith, and S. Aswani, editors. *Environmental social sciences: methods and research design*. Cambridge University Press, Cambridge, UK. http://dx.doi.org/10.1017/CBO9780511760242.014

Brown, R. R., M. A. Farrelly, and D. A. Loorbach. 2013. Actors working the institutions in sustainability transitions: the case of Melbourne's stormwater management. *Global Environmental Change* 23(4):701-718. http://dx.doi.org/10.1016/j.gloenvcha.2013.02.013

Câmpeanu, C. N., and I. Fazey. 2014. Adaptation and pathways of change and response: a case study from Eastern Europe. *Global Environmental Change* 28:351–367. http://dx.doi.org/10.1016/j.gloenvcha.2014.04.010

Castree, N., W. M. Adams, J. Barry, D. Brockington, B. Büscher, E. Corbera, D. Demeritt, R. Duffy, U. Felt, K. Neves, P. Newell, L. Pellizzoni, K. Rigby, P. Robbins, L. Robin, D. B. Rose, A. Ross, D. Schlosberg, S. Sörlin, P. West, M. Whitehead, and B. Wynne. 2014. Changing the intellectual climate. *Nature Climate Change* 4(9):763-768. http://dx.doi.org/10.1038/nclimate2339

Chambers, R. 1994. The origins and practice of participatory rural appraisal. *World Development* 22(7):953-969. http://dx.doi.org/10.1016/0305-750X(94)90141-4

Cinner, J. E., T. R. McClanahan, N. A. J. Graham, T. M. Daw, J. Maina, S. M. Stead, A. Wamukota, K. Brown, and Ö. Bodin. 2012. Vulnerability of coastal communities to key impacts of climate change on coral reef fisheries. *Global Environmental Change* 22(1):12-20. http://dx.doi.org/10.1016/j.gloenvcha.2011.09.018

Cody, K. C., S. M. Smith, M. Cox, and K. Andersson. 2015. Emergence of collective action in a groundwater commons: irrigators in the San Luis Valley of Colorado. *Society & Natural Resources*. http://dx.doi.org/10.1080/08941920.2014.970736

Cohen, P. J., L. S. Evans, and M. Mills. 2012. Social networks supporting governance of coastal ecosystems in Solomon Islands. *Conservation Letters* 5(5):376-386. http://dx.doi.org/10.1111/j.1755-263X.2012.00255.x

Cox, M. 2008. Balancing accuracy and meaning in common-pool resource theory. *Ecology and Society* 13(2): 44. [online] URL: http://www.ecologyandsociety.org/vol13/iss2/art44/

Cox, M. 2014*a*. The SESMAD project. *International Journal of the Commons* 8(2).

Cox, M. 2014*b*. Applying a social-ecological system framework to the study of the Taos Valley irrigation system. *Human Ecology* 42(2):311-324. http://dx.doi.org/10.1007/s10745-014-9651-y

Cox, M. 2014*c*. Modern disturbances to a long-lasting community-based resource management system: the Taos Valley acequias. *Global Environmental Change* 24:213-222. http://dx.doi.org/10.1016/j.gloenvcha.2013.12.006

Cox, M., G. Arnold, and S. Villamayor Tomás. 2010. A review of design principles for community-based natural resource management. *Ecology and Society* 15(4): 38. [online] URL: http://www.ecologyandsociety.org/vol15/iss4/art38/

Cox, J. C., E. Ostrom, J. M. Walker, J. Castillo, E. Coleman, R. Holahan, M. Schoon, and B. Steed. 2009. Trust in private and common property experiments. *Southern Economic Journal* :957-975.

Cox, M., and J. Ross. 2011. Robustness and vulnerability of community irrigation systems: the case of the Taos Valley acequias. *Journal of Environmental Economics and Management* 61(3):254-266. http://dx.doi.org/10.1016/j.jeem.2010.10.004

Crate, S. A. 2006. *Cows, kin, and globalization: an ethnography of sustainability.* Altamira, Lanham, Maryland, USA.

Cronon, W. 2011. *Changes in the land: Indians, colonists, and the ecology of New England*. Macmillan, London, UK.

Delgado, L. E., V. H. Marín, P. L. Bachmann, and M. Torres-Gomez. 2009. Conceptual models for ecosystem management through the participation of local social actors: the Río Cruces wetland conflict. *Ecology and Society* 14(1): 50. [online] URL: http://www.ecologyandsociety.org/vol14/iss1/art50/

Denzin, N. K. 1989. *Interpretive biography.* Sage, Newbury Park, California, USA. http://dx.doi.org/10.4135/9781412984584

Diamond, J. 1997. *Guns, germs, and steel: the fates of human societies.* W. W. Norton & Company, New York, New York, USA.

Diamond, J. 2005. *Collapse: how societies choose to fail or succeed*. Viking, New York, New York, USA.

Downing, A. S., E. H. van Nes, J. S. Balirwa, J. Beuving, P. O. J. Bwathondi, L. J. Chapman, I. J. M. Cornelissen, I. G. Cowx, K. P. C. Goudswaard, R. E. Hecky, et al. 2014. Coupled human and natural system dynamics as key to the sustainability of Lake Victoria's ecosystem services. *Ecology and Society* 19(4): 31. http://dx.doi.org/10.5751/ES-06965-190431

Ernst, B., J. Chamorro, P. Manríquez, J. M. L. Orensanz, A. M. Parma, J. Porobic, and C. Román. 2013. Sustainability of the Juan Fernández lobster fishery (Chile) and the perils of generic science-based prescriptions. *Global Environmental Change* 23 (6):1381-1392. http://dx.doi.org/10.1016/j.gloenvcha.2013.08.002

Evans, L., N. Cherrett, and D. Pemsl. 2011. Assessing the impact of fisheries co-management interventions in developing countries: a meta-analysis. *Journal of Environmental Management* 92(8):1938-1949. http://dx.doi.org/10.1016/j.jenvman.2011.03.010

Fahey, T. J., P. B. Woodbury, J. J. Battles, C. L. Goodale, S. P. Hamburg, S. V. Ollinger, and C. W. Woodall. 2010. Forest carbon storage: ecology, management, and policy. *Frontiers in Ecology and the Environment* 8(5):245-252. http://dx.doi.org/10.1890/080169

Fazey, I., N. Pettorelli, J. Kenter, D. Wagatora, and D. Schuett. 2011. Maladaptive trajectories of change in Makira, Solomon Islands. *Global Environmental Change* 21(4):1275-1289. http://dx.doi.org/10.1016/j.gloenvcha.2011.07.006

Fleischman, F. D. 2014. Why do foresters plant trees? Testing theories of bureaucratic decision-making in central India. *World Development* 62:62-74. http://dx.doi.org/10.1016/j.worlddev.2014.05.008

Fleischman, F. D., B. Loken, G. A. Garcia-Lopez, and S. Villamayor-Tomas. 2014. Evaluating the utility of common-pool resource theory for understanding forest governance and outcomes in Indonesia between 1965 and 2012. *International Journal of the Commons* 8(2).

Fowler, F. J. 2009. *Survey research methods*. Sage, London, UK.

Geertz, C. 1959. Form and variation in Balinese village structure. *American Anthropologist* 61(6):991-1012. http://dx.doi.org/10.1525/aa.1959.61.6.02a00060

Geertz, C. 1973. *The interpretation of cultures: selected essays.* Basic Books, New York, New York, USA.

Geertz, C. 1980. Organization of the Balinese subak. Pages 70-90 *in* E. W. Coward, editor. *Irrigation and agricultural development in Asia: perspectives from the social sciences.* Cornell University Press, Ithaca, New York, USA.

Geist, H. J., and E. F. Lambin. 2002. Proximate causes and underlying driving forces of tropical deforestation. *Bioscience* 52 (2):143-150. http://dx.doi.org/10.1641/0006-3568(2002)052[0143:PCAUDF]2.0.CO;2

George, A. L., and A. Bennett. 2005. *Case studies and theory development in the social sciences.* MIT Press, Cambridge, Massachusetts, USA.

Gilbert, H. 2013. 'Bedouin overgrazing' and conservation politics: challenging ideas of pastoral destruction in South Sinai. *Biological Conservation* 160:59-69. http://dx.doi.org/10.1016/j.biocon.2012.12.022

Guest, G., E. E. Namey, and M. L. Mitchell. 2012. *Collecting qualitative data: a field manual for applied research.* Sage, Thousand Oaks, California, USA.

Guimarães, M. H., J. Ballé-Béganton, D. Bailly, A. Newton, T. Boski, and T. Dentinho. 2013. Transdisciplinary conceptual modeling of a social-ecological system: a case study application in Terceira Island, Azores. *Ecosystem Services* 3:e22-e31. http://dx.doi.org/10.1016/j.ecoser.2012.12.007

Gunderson, L., and C. S. Holling. 2002. *Panarchy: understanding transformations in human and natural systems.* Island Press, Washington, D.C., USA.

Gutiérrez, N. L., R. Hilborn, and O. Defeo. 2011. Leadership, social capital and incentives promote successful fisheries. *Nature* 470:386-389. http://dx.doi.org/10.1038/nature09689

Hanna, R., E. Duflo, and M. Greenstone. 2012. *Up in smoke: the influence of household behavior on the long-run impact of improved cooking stoves.* NBER Working Paper No. 18033. National Bureau of Economic Research, Cambridge, Massachusetts, USA. http://dx.doi.org/10.2139/ssrn.2039004

Hayami, Y. 1974. Conditions for the diffusion of agricultural technology: an Asian perspective. *Journal of Economic History* 34(01):131-148. http://dx.doi.org/10.1017/S0022050700079675

Hayes, T. M. 2006. Parks, people, and forest protection: an institutional assessment of the effectiveness of protected areas. *World Development* 34(12):2064-2075. http://dx.doi.org/10.1016/j.worlddev.2006.03.002

Holling, C. S., and G. K. Meffe. 1996. Command and control and the pathology of natural resource management. *Conservation Biology* 10(2):328–337. http://dx.doi.org/10.1046/j.1523-1739.1996.10020328.x

Homer-Dixon, T. F. 2010. *Environment, scarcity, and violence.* Princeton University Press, Princeton, New Jersey, USA. http://dx.doi.org/10.1515/9781400822997

Isaac, M. E., L. C. N. Anglaaere, D. S. Akoto, and E. Dawoe. 2014. Migrant farmers as information brokers: agroecosystem management in the transition zone of Ghana. *Ecology and Society* 19(2): 56. http://dx.doi.org/10.5751/ES-06589-190256

Janssen, M. A., and E. Ostrom. 2006. Empirically based, agent-based models. *Ecology and Society* 11(2): 37. [online] URL: http://www.ecologyandsociety.org/vol11/iss2/art37/

Kauffman, S. 2012. Random clinical trials are not the only answer to identifying effective therapies. *13.7: Cosmos And Culture.* NPR, Washington, D.C., USA. [online] URL: http://www.npr.org/blogs/13.7/2012/10/01/162084487/random-clinical-trials-are-not-the-only-answer-to-identifying-effective-therapie

Keskitalo, E. C. H., C. Sandström, M. Tysiachniouk, and J. Johansson. 2009. Local consequences of applying international norms: differences in the application of forest certification in northern Sweden, northern Finland, and northwest Russia. *Ecology and Society* 14(2): 1. [online] URL: http://www.ecologyandsociety.org/vol14/iss2/art1/

K.C., S. 2013. Community vulnerability to floods and landslides in Nepal. *Ecology and Society* 18(1): 8. http://dx.doi.org/10.5751/ES-05095-180108

King, G., R. O. Keohane, and S. Verba. 1994. *Designing social inquiry: scientific inference in qualitative research.* Princeton University Press, Princeton, New Jersey, USA.

Kirsten, J. 2009. *Institutional economics perspectives on African agricultural development.* International Food Policy Research Institute, Washington, D.C., USA.

Kitamura, K., and R. A. Clapp. 2013. Common property protected areas: community control in forest conservation. *Land Use Policy* 34:204-212. http://dx.doi.org/10.1016/j.landusepol.2013.03.008

Klooster, D. 2000. Institutional choice, community, and struggle: a case study of forest co-management in Mexico. *World Development* 28(1):1-20. http://dx.doi.org/10.1016/S0305-750X(99)00108-4

Lambin, E. F., and P. Meyfroidt. 2011. Global land use change, economic globalization, and the looming land scarcity. *Proceedings of the National Academy of Sciences of the United States of America* 108(9):3465-3472. http://dx.doi.org/10.1073/pnas.1100480108

Lambin, E. F., B. L. Turner, H. J. Geist, S. B. Agbola, A. Angelsen, J. W. Bruce, O. T. Coomes, R. Dirzo, G. Fischer, C. Folke, P. S. George, K. Homewood, J. Imbernon, R. Leemans, X. Li, E. F. Moran, M. Mortimore, P. S. Ramakrishnan, J. F. Richards, H. Skånes, W. Steffen, G. D. Stone, U. Svedin, T. A. Veldkamp, C. Vogel, and J. Xu. 2001. The causes of land-use and land-cover change: moving beyond the myths. *Global Environmental Change* 11(4):261-269. http://dx.doi.org/10.1016/S0959-3780(01)00007-3

Lansing, S. 1991. *Priests and programmers: technologies of power in the engineered landscape of Bali.* Princeton University Press, Princeton, New Jersey, USA.

Lapola, D. M., R. Schaldach, J. Alcamo, A. Bondeau, J. Koch, C. Koelking, and J. A. Priess. 2010. Indirect land-use changes can overcome carbon savings from biofuels in Brazil. *Proceedings of the National Academy of Sciences of the United States of America* 107(8):3388-3393. http://dx.doi.org/10.1073/pnas.0907318107

Lauber, T. B., D. J. Decker, and B. A. Knuth. 2008. Social networks in community-based natural resource management. *Environmental Management* 42(4):677-687. http://dx.doi.org/10.1007/s00267-008-9181-8

Lorent, H., R. Sonnenschein, G. M. Tsiourlis, P. Hostert, and E. Lambin. 2009. Livestock subsidies and rangeland degradation in central Crete. *Ecology and Society* 14(2): 41. [online] URL: http://www.ecologyandsociety.org/vol14/iss2/art41/

Manfred, M. J., J. J. Vaske, A. Rechkemmer, and E. A. Duke. 2104. *Understanding society and natural resources.* Springer, New York, New York, USA. http://dx.doi.org/10.1007/978-94-017-8959-2

Marín, V. H., L. E. Delgado, and P. Bachmann. 2008. Conceptual PHES-system models of the Aysén watershed and fjord (Southern

Chile): testing a brainstorming strategy. *Journal of Environmental Management* 88(4):1109-1118. http://dx.doi.org/10.1016/j.jenvman.2007.05.012

May, R., and A. McLean. 2007. *Theoretical ecology: principles and applications*. Third edition. Oxford University Press, Oxford, UK.

McClanahan, T. R., and S. Mangi. 2000. Spillover of exploitable fishes from a marine park and its effect on the adjacent fishery. *Ecological Applications* 10(6):1792-1805. http://dx.doi.org/10.1890/1051-0761(2000)010[1792:SOEFFA]2.0.CO;2

Moran, E. F. 2010. *Environmental social science: human - environment interactions and sustainability*. Wiley-Blackwell, Malden, Massachusetts, USA.

Nagendra, H., and E. Ostrom. 2011. The challenge of forest diagnostics. *Ecology and Society* 16(2): 20. [online] URL: http://www.ecologyandsociety.org/vol16/iss2/art20/

Neudoerffer, R. C., D. Waltner-Toews, J. J. Kay, D. D. Joshi, and M. S. Tamang. 2005. A diagrammatic approach to understanding complex eco-social interactions in Kathmandu, Nepal. *Ecology and Society* 10(2): 12. [online] URL: http://www.ecologyandsociety.org/vol10/iss2/art12/

Neuendorf, K. A. 2002. *The content analysis guidebook*. Sage, Thousand Oaks, California, USA.

Osbahr, H., C. Twyman, W. N. Adger, and D. S. G. Thomas. 2010. Evaluating successful livelihood adaptation to climate variability and change in southern Africa. *Ecology and Society* 15(2): 27. [online] URL: http://www.ecologyandsociety.org/vol15/iss2/art27/

Österblom, H., and U. R. Sumaila. 2011. Toothfish crises, actor diversity and the emergence of compliance mechanisms in the Southern Ocean. *Global Environmental Change* 21(3):972-982. http://dx.doi.org/10.1016/j.gloenvcha.2011.04.013

Ostrom, E. 1990. *Governing the commons: the evolution of institutions for collective action*. Cambridge University Press, Cambridge, UK. http://dx.doi.org/10.1017/CBO9780511807763

Ostrom, E. 2005. *Understanding institutional diversity*. Princeton University Press, Princeton, New Jersey, USA.

Ostrom, E. 2006. The value-added of laboratory experiments for the study of institutions and common-pool resources. *Journal of Economic Behavior & Organization* 61(2):149-163. http://dx.doi.org/10.1016/j.jebo.2005.02.008

Ostrom, E. 2007. A diagnostic approach for going beyond panaceas. *Proceedings of the National Academy of Sciences of the United States of America* 104(39):15181-15187. http://dx.doi.org/10.1073/pnas.0702288104

Pahl-Wostl, C., G. Becker, C. Knieper, and J. Sendzimir. 2013. How multilevel societal learning processes facilitate transformative change: a comparative case study analysis on flood management. *Ecology and Society* 18(4): 58. http://dx.doi.org/10.5751/ES-05779-180458

Pahl-Wostl, C., and C. Knieper. 2014. The capacity of water governance to deal with the climate change adaptation challenge: using fuzzy set qualitative comparative analysis to distinguish

between polycentric, fragmented and centralized regimes. *Global Environmental Change* 29:139-154. http://dx.doi.org/10.1016/j.gloenvcha.2014.09.003

Parker, D. C., S. M. Manson, M. A. Janssen, M. J. Hoffmann, and P. Deadman. 2003. Multi-agent systems for the simulation of land-use and land-cover change: a review. *Annals of the Association of American Geographers* 93(2):314–337. http://dx.doi.org/10.1111/1467-8306.9302004

Persha, L., and T. Blomley. 2009. Management decentralization and montane forest conditions in Tanzania. *Conservation Biology* 23(6):1485-1496. http://dx.doi.org/10.1111/j.1523-1739.2009.01276.x

Poteete, A. R., M. A. Janssen, and E. Ostrom. 2010. *Working together: collective action, the commons, and multiple methods in practice*. Princeton University Press, Princeton, New Jersey, USA. http://dx.doi.org/10.1515/9781400835157

Prell, C., and Ö. Bodin. 2011. *Social Networks and Natural Resource Management: Uncovering the social fabric of environmental governance*. Cambridge University Press, New York, New York, USA.

Pullin, A. S., and G. B. Stewart. 2006. Guidelines for systematic review in conservation and environmental management. *Conservation Biology* 20(6):1647-1656. http://dx.doi.org/10.1111/j.1523-1739.2006.00485.x

Ragin, C. C. 1987. *The comparative method: moving beyond qualitative and quantitative strategies*. University of California Press, Berkeley and Los Angeles, California, USA.

Ragin, C. C. 2000. *Fuzzy set social science*. Chicago University Press, Chicago, Illinois, USA.

Remler, D. K., and G. G. Van Ryzin. 2011. *Research methods in practice: strategies for description and causation*. Sage, Thousand Oaks, California, USA.

Ribot, J. C. 1998. Theorizing access: forest profits along Senegal's charcoal commodity chain. *Development and Change* 29 (2):307-341. http://dx.doi.org/10.1111/1467-7660.00080

Robbins, P. 2000. The rotten institution: corruption in natural resource management. *Political Geography* 19(4):423-443. http://dx.doi.org/10.1016/S0962-6298(99)00087-6

Rudel, T. K. 2008. Meta-analyses of case studies: a method for studying regional and global environmental change. *Global Environmental Change* 18(1):18-25. http://dx.doi.org/10.1016/j.gloenvcha.2007.06.001

Sagarin, R., and A. Pauchard. 2012. *Observation and ecology: broadening the scope of science to understand a complex world*. Island Press, Washington, D.C., USA. http://dx.doi.org/10.5822/978-1-61091-230-3

Scheffer, M., S. Carpenter, J. A. Foley, C. Folke, and B. Walker. 2001. Catastrophic shifts in ecosystems. *Nature* 413:591-596. http://dx.doi.org/10.1038/35098000

Schlager, E. 2007. A comparison of frameworks, theories, and models of policy processes. Pages 293-319 *in* P. A. Sabatier, editor. *Theories of the policy process*. Westview Press, Boulder, Colorado, USA.

Schneider, M., J. Scholz, M. Lubell, D. Mindruta, and M. Edwardsen. 2003. Building consensual institutions: networks and the National Estuary Program. *American Journal of Political Science* 47:143–158. http://dx.doi.org/10.1111/1540-5907.00010

Scott, J. C. 1998. *Seeing like a state: how certain schemes to improve the human condition have failed*. Yale University Press, New Haven, Connecticut, USA.

Sendzimir, J., C. P. Reij, and P. Magnuszewski. 2011. Rebuilding resilience in the Sahel: regreening in the Maradi and Zinder Regions of Niger. *Ecology and Society* 16(3): 1. http://dx.doi.org/10.5751/ES-04198-160301

Shadish, C., T. Cook, and D. T. Campbell. 2002. *Experimental and quasi-experimental designs for generalized causal inference*. Houghton-Mifflin, Boston, Massachusetts, USA.

Snorek, J., F. G. Renaud, and J. Kloos. 2014. Divergent adaptation to climate variability: a case study of pastoral and agricultural societies in Niger. *Global Environmental Change* 29:371–386. http://dx.doi.org/10.1016/j.gloenvcha.2014.06.014

Stoffle, R. W., D. B. Halmo, T. W. Wagner, and J. J. Luczkovich. 1994. Reefs from space: satellite imagery, marine ecology, and ethnography in the Dominican Republic. *Human Ecology* 22 (3):355–378. http://dx.doi.org/10.1007/BF02168857

Tong, A., P. Sainsbury, and J. Craig. 2007. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *International Journal for Quality in Health Care* 19(6):349-357. http://dx.doi.org/10.1093/intqhc/mzm042

Vaccaro, I., E. A. Smith, and S. Aswani. 2010. *Environmental social sciences: methods and research design*. Cambridge University Press, Cambridge, UK. http://dx.doi.org/10.1017/CBO9780511760242

Villamayor-Tomas, S., F. D. Fleischman, I. P. Ibarra, A. Thiel, and F. van Laerhoven. 2014. From Sandoz to salmon: conceptualizing resource and institutional dynamics in the Rhine watershed through the SES framework. *International Journal of the Commons* 8(2):361-395.

Vogt, N. D., A. Y. Banana, W. Gombya-Ssembajjwe, and J. Bahati. 2006. Understanding the stability of forest reserve boundaries in the West Mengo region of Uganda. *Ecology and Society* 11(1): 38. [online] URL: http://www.ecologyandsociety.org/vol11/iss1/art38/

Walker, B. H., S. R. Carpenter, J. Rockstrom, A.-S. Crépin, and G. D. Peterson. 2012. Drivers, "slow" variables, "fast" variables, shocks, and resilience. *Ecology and Society* 17(3): 30. http://dx.doi.org/10.5751/ES-05063-170330

Wilson, C., and C. Tisdell. 2001. Why farmers continue to use pesticides despite environmental, health and sustainability costs. *Ecological Economics* 39(3):449-462. http://dx.doi.org/10.1016/S0921-8009(01)00238-5

Yin, R. K. 2014. *Case study research: design and methods*. Sage, Thousand Oaks, California, USA.

Young, O. 2002. *The institutional dimensions of environmental change: fit, interplay, and scale*. MIT Press, Cambridge, Massachusetts, USA.

Young, O. R., E. F. Lambin, F. Alcock, H. Haberl, S. I. Karlsson, W. J. McConnell, T. Myint, C. Pahl-Wostl, C. Polsky, P. S. Ramakrishnan, H. Shroeder, M. Scouvart, and P. H. Verburg. 2006. A portfolio approach to analyzing complex human-environment interactions: institutions and land change. *Ecology and Society* 11(2): 31. [online] URL: http://www.ecologyandsociety.org/vol11/iss2/art31/

Appendix 1. Definitions of important terms.

Additionality
The difference between the value of an outcome after the implementation of a policy, and its value in a counterfactual scenario in which the policy is not enacted.

Analysis
A step in the research process that involves describing and then making inferences based on a set of data.

Baseline
A measurement of an outcome made prior to an intervention, which is then compared with measurements made after the intervention in order to evaluate the effects of the intervention.

Cascade effect
The indirect effects that an intervention has on an outcome in areas other than that for which the intervention was intended.

Case study
An observational study that involves the in-depth study of a single, or relatively few, observations.

Case study meta-analysis
A quantitative synthetic study that uses content analysis to systematically codes findings from previous case studies in order to compare them with a new dataset.

Categorical variable
A variable with a range of possible values that cannot be meaningfully ordered.

Census
An exhaustive sample that contains every member of the population of interest.

Cluster sample
A sample collected by first dividing a population into (frequently geographically defined) clusters, selecting a sample of these clusters, and then selecting samples of observations within each cluster.

Concept
A noun or adjective of theoretical/scientific significance that has been given a specific definition by a research community.

Congruence testing
A deductive qualitative analytical strategy of testing a theory by comparing the observational expectations it generates for a case with the observations made on this case.

Confounding variable
A variable that is associated with both a dependent variable and an independent variable and is thus a threat to causal inferences made regarding these two variables.

Content analysis
The process of coding qualitative information and data to produce quantitative data.

Convenience sample
A sample of observations that is selected primarily or exclusively based on the accessibility of these observations and the convenience of accessing them.

Correlational study
An observational study that involves the (usually statistically) comparative analysis of a large number of observations.

Covariation
A relationship between two variables in which an increase in one is associated with an increase or decrease in the other.

Counterfactual
An alternative scenario to which a realized scenario is compared in order to evaluate the significance of a causal factor that changes across the scenarios.

Cross-sectional comparison
A comparison of multiple observations, each of which is a measurement of a different entity at the same point in time.

Deduction
The process of developing testable hypotheses as the observational implications of a theory, and testing these hypotheses, and thus the theory, with empirical data.

Deductive validity
The accuracy with which a general principle or theory is applied to a specific case or context.

Dependent variable
A variable that is viewed as an outcome to be explained.

Direct observation
Measurement strategy in which the researcher directly observes the subject of observation, and either directly (via video or audio) records the subject, or records the values of qualitative or quantitative variables describing the subject.

Ecological fallacy
Inaccurately assuming that the characteristics of a population or group are representative of subgroups within that population or group.

Ecological validity
The accuracy with which findings from a highly controlled project, usually an experiment, can be generalized to more complex, real-world environments.

Embedded case study
A study that combines a case study of one unit of analysis, as well as a correlational study of a unit of analysis nested within the cases study.

Endogeneity
A situation in which a supposed dependent variable causes an independent variable to change.

Environmental History
A largely qualitative and historical approach to analysis that examines the historical relationship between humans and their natural environment.

Ethicality
The extent to which a research project upholds important ethical standards.

Ethnography
A highly inductive, fieldwork-based approach to social science that focuses on developing a thick understanding of a particular culture.

Experiment
A study that randomly assigns observations to control and treatment groups in order to isolate the effects of the treatment on the treatment group.

Expert sampling
A type of purposive sampling in which human respondents are identified based on their expertise in a subject matter relevant to the research project.

External validity
The extent to which characteristics of a sample can be generalized to describe the larger population or other related population (also called internal validity).

Factorial experiment
An experiment with more than one treatment, and subsequently multiple distinct subgroups to reflect each possible combination of the presence or absence of each treatment.

Feasibility
The extent to which a research project can be feasibly accomplished given the resource constraints facing a research team.

Focus group
A data collection strategy that involves a researcher holding a meeting with multiple respondents at once in order to engage with them and observe their interactions.

Framework
A set of concepts along with a set of statements describing their relationships.

Geographic information systems
Hardware, software and analytical operations designed to collect, process and analyze (primarily vector-based) spatial data.

Hypothesis
An observational implication of a theory. A statement that describes an empirically observable pattern that would be expected if a theory were correct.

Independent effect
An effect that an independent variable has on a dependent variable irrespective of the values of other variables.

Independent variable
A variable that is viewed as a cause of an outcome.

Induction
The formation of general principles or theories based on patterns or regularities found in a set of data.

Inductive theory-building
The process of inferring generalized relationships among a set of variables based on the (qualitative or quantitative) analysis of a particular case or set of observations.

Inference
A conclusion that explains a set of data by combining the data with something else, such as prior knowledge, a theory or model, or set of assumptions.

Informal interview
Occurs when researcher talks informally with subjects without any structured way of guiding the discussion or recording data.

In-person instrumentation
The use of a technological device to record data about an environment, or to take samples from this environment.

Interaction effect
The effect that two independent variables have on a dependent variable based on a non-additive interaction between them.

Internal validity
The validity of an inference connecting two or more variables in a causal relationship.

Interval/ratio variable
A variable with a range of possible values that includes a set of numeric values that can be compared in absolute terms.

Leakage
A process in which forbidding certain behaviors or outcomes in one jurisdiction creates incentives for these activities to spread elsewhere. This complicates causal inference.

Linear relationship
A relationship between an independent variable and a dependent variable that doesn't change in nature or magnitude across the range of either variable.

Longitudinal comparison
A comparison among multiple observations, each of which is a measurement of the same entity at distinct points in time.

Measurement validity
Quality of a variable based on (1) the fidelity of this variable to the concept it operationalizes, and (2) the accuracy with which this variable is measured to produce data.

Mediator variable
A variable that serves as a proximate cause of a dependent variable by mediating the effects of an underlying independent variable on this dependent variable.

Mid-range theory
A theory that is relevant for an identifiable set of cases or type of case, but not all cases.

Model
A theory, or set of theories, expressed in a formal language (e.g. with graphs, mathematics).

Moderator variable
A variable that affects the strength of a relationship between two other variables, thus producing an interaction effect.

Multistage sample
A sample that is obtained via two selection procedures rather than just one.

Narrative path analysis
A qualitative analytical method that explores the historical path that a system has taken and how this path affects the system's current situation and the prospects for future dynamics (rigidities, adaptations).

Natural experiment
An observational study that mimics an experiment, in which some observations are naturally exposed to a condition, while other highly similar observations are not.

Necessity
A condition in which an independent variable is required for a dependent variable to be present.

Negative feedback
The self-dampening effect of a relationship between two variables in which the first positively affects the other, and the other negatively affects the first.

Negative relationship
A relationship in which an increase in an independent variable causes a decrease in a dependent variable.

Network analysis
Analysis of a system as a network of nodes connected by links. Involves estimation of network-level and node-level properties and attempts to associate them with important outcomes.

Nonlinear relationship
A relationship between an independent variable and a dependent variable that changes in nature or magnitude across a particular threshold within the range of each variable.

Objective data
Data obtained without the involvement of human respondents, produced by the use of some physical data collection device or direct observation.

Observation
An instance of a unit of analysis or observation (e.g. a resource user) that is measured and analyzed.

Observational study
A non-experimental study that does not involve any active intervention or control on the part of the researcher.

Operationalization
The process of constructing a variable out of a concept by assigning it a range along which it can vary.

Ordinal variable
A variable with a range of possible values that can be meaningful ordered but cannot be compared in absolute terms.

Participant observation
A data collection strategy in which the researcher becomes an active participant in a study system in order to understand complex, day-to-day nuances in that system which otherwise may be hard to predict or understand.

Participatory rural appraisal (PRA)
A set of techniques that build on rapid rural appraisal by formally incorporating community stakeholder input into the appraisal process.

Permanence
The ability of a policy intervention to maintain its effects over time, particularly after the intervention itself has ceased.

Population
A set of observations of theoretical interest, about which the researcher wants to tell a scientifically rigorous story.

Positive feedback
The self-reinforcing effect of a relationship between two variables in which each has the same effect on the other.

Positive relationship
A relationship in which an increase in an independent variable causes an increase in a dependent variable.

Practical importance
The extent to which a research question and associated research project have some practical importance to society and the environment.

Primary data
Novel data collected by a researcher for a specific project.

Proportional sample
A stratified sample that has numbers of observations in each strata that reflect their proportions within the larger population.

Proximate cause
A cause that is in some way closest to an outcome of interest.

Purposive sample
A non-random sample that is obtained by purposively selecting observations from a population. Usually used in small-n research.

Qualitative analysis
An analysis of non-numerical data, usually either via content analysis to create quantitative data, or inferences made via direct observation and experience.

Qualitative comparative analysis
Estimation of the necessity and sufficiency of combinations of factors to produce an outcome.

Qualitative literature review
A non-quantitative synthetic study that summarizes findings from a particular research program or discipline.

Qualitative modeling
The process of developing a qualitative model of a system that divides it up into constituent components and describes the relationships between them, without quantification.

Qualitative variable
A variable that can take on any text value.

Quantitative analysis
An analysis that examines the associations among quantitative (categorical, ordinal, and interval/ratio) variables.

Quasi-experiment
An experiment in which the assignment of observations to control and treatment groups is non-random.

Random sample
A sample that is obtained by randomly selecting observations from a population. It is usually, but not necessarily, representative of that population.

Rapid rural appraisal (RRA)
A set of multidisciplinary techniques primarily conducted by development professionals to expediently collect data in rural areas by balancing between formal surveys and completely unstructured interview approaches.

Reliability
The consistency with which variables are measured across data collectors.

Remote instrumentation
Measurement method in which a researcher remotely manages a data collection technology that records features about the subject of observation.

Remote sensing
Hardware, software and analytical operations designed to collect, process and analyze (primarily raster-based) spatial data.

Representative sample
A sample that is representative of the population from which it was selected. This is needed in order to generalize findings regarding the sample to the larger population.

Research design
The type of study (e.g. correlational, experimental) that a research project employs to address its research questions.

Research question
A question of some practical and scientific interest that may be descriptive or causal, in which case it generally asks what factors affect an important outcome.

Sample
A subset of observations from a population that is selected for measurement and analysis.

Sample size
The number of observations within a sample.

Sampling bias
A bias that results from a sampling strategy that systematically selects for particular types of observations more than others, leading to a non-representative sample.

Sampling frame
A list of observations within a population that are available for sampling.

Sampling strategy
The process of selecting a sample of observations from a population of theoretical interest.

Secondary data
Data collected previously for different research projects that are used by a researcher to address a new question.

Self-administered survey
Measurement strategy in which human subjects fill out questionnaires without the presence of the researcher.

Semi-structured interview
An interview in which the researcher conducts an interview with the help of an interview guide.

Snowball sample
A sample that is obtained by identifying subsequent observations based on interactions with a preliminary set of observations.

Spillover effect
The process whereby an intervention applied to one set of observations ends up affecting other observations, complicating causal inference. Also called diffusion.

Spurious relationship
A correlation or association between two variables that does not actually indicate a causal relationship between these variables.

Statistical analysis
Calculations applied to quantitative data to descriptively summarize a sample and to make inferences about the population from which the sample is drawn.

Statistical validity
The validity of statistical inferences based on the satisfaction of important supporting statistical assumptions.

Stratified sample
A sample obtained by dividing a population into subgroups and sampling within each one of those subgroups.

Structured interview
Occurs when a researcher conducts an interview with the help of a written questionnaire that is filled out as the subject responds to the questions it contains.

Subjective data
Data obtained by eliciting the perceptions of human respondents.

Sufficiency
A condition in which the presence of an independent variable guarantees the presence of a dependent variable.

Synthetic study
A study that relies exclusively, or at least primarily, on secondary data to synthesize findings from a set of existing research projects, results, or case studies.

Systematic review
A synthetic study that qualitatively and/or quantitatively analyzes secondary data from a set of studies in order to examine the effects of an intervention.

Theory
A statement that describes (1) a (causal) relationship between two or more concepts, and (2) a mechanism by which this relationship occurs.

Triangulation
Comparing the results of different methods in order to check whether they produce similar or different results.

Type 1 error
Inferring that a pattern or cause exists when it does not.

Type 2 error
Inferring that a pattern or cause does not exist when it does.

Underlying cause
A cause that affects an outcome via a proximate cause.

Unit of analysis
The object of an analysis. A category about which a research question is posed, instances of which are frequently compared to test hypotheses.

Unit of measurement
The unit, such as meters or pounds, which is used to count or measure an interval or ratio-level variable.

Unit of observation
A data source that is used to obtain information about a unit of analysis. Can also be a unit of analysis.

Unstructured interview
Occurs when a researcher conducts an interview without any supporting materials.

Variable
An operationalized concept that has been assigned a level of measurement and range of possible values.

Appendix 2. Review of scientific terminology.

Basic terms

The design of any research project starts by developing a **research question**, because the value of everything done subsequently depends on its ability to help the researcher address this question. The following is a list of sample research questions:

1. To what extent does technological change ***crowd out behavioral change*** and what balance between technological and behavioral change is appropriate under different circumstances? For example, do recycling efforts lead us to reduce and reuse (the other of the three R's) less?

2. What factors about the current generation in the United States explain why it is ***apathetic*** with respect to particular environmental issues?

3. Can a small-scale biodiesel company that sources its oil primarily from street-food vendors make a profit in a New Delhi neighborhood? If so, what factors would affect the ***financial viability*** of this company?

4. What are the different ***environmental and economic impacts*** of shade-grown vs. sun-grown coffee plantations in Guatemala?

5. Is the cultivation of Jatropha curcas for biofuel production an ***economically advantageous*** option for small-scale farmers in Haiti? If so, what factors enable this outcome?

6. How ***effective*** will a payment for ecosystem services (PES) scheme be in the community of Boca Pariamanu in Madre de Dios, Peru?

7. How can remote sensing and geographic information systems technologies ***lower the transaction costs*** involved in monitoring outcomes of market-based watershed and water quality protection schemes in the Northern Forest region of New England?

8. What is the relationship between the presence of oil refinery sites, fuel-burning plants and ***county-level Medicaid expenditures on asthma***?

9. What makes for an ***effective eco-label*** on clothing?

10. What drives ***tropical deforestation***?

Frequently the formulation of a research question will establish several other elements of a research project. This can include the primary outcome that the researcher wants to explain (such as changes in levels of forest cover). Each of the questions listed above contains an outcome of interest, expressed at various levels of specificity. The terms most closely associated with the outcome to be explained are bolded and italicized in each question. Research questions also frequently imply or establish a **unit of analysis** that exhibits this outcome (e.g. forests) and potentially a particular population of this unit (a set of forests). For example, a researcher could ask, what types of institutional arrangements and property rights regimes help to sustain forest cover? Sustained forest cover is the outcome identified here, and it is an outcome that different forests may exhibit to varying degrees.

Research questions generally inquire about relationships among **concepts**. By "concept" I mean what Adcock and Collier (2001) refer to as a "systematized concept", or one that has been given a fairly specific definition by a particular researcher or research group. Concepts in ESS include both biophysical concepts, such as "forest cover change" as just described, and social concepts, such as "resource dependence" or the extent to which a user group depends on a natural resource. Many social concepts are very intangible and difficult to measure with much precision. **Variables** are like concepts but are assigned a well-defined range along which they can vary (such as high, medium, or low). Variables are essentially what we turn concepts into in order to measure them. The process of turning a concept into a variable is called **operationalization.**

Statements that describe (1) a causal relationship between two or more concepts, and (2) the mechanism by which this occurs are called **theories**. Theories are generally derived from **induction**, or the process of forming generalizations based on patterns found in a set of observations. Several authors (Young 2002; Cox 2008) have argued that the most desirable theories in ESS are **mid-range theories**. These are theories that are not so specific that they are overfit to, or only relevant for, a particular case or dataset, but also are not so general as to be broadly but only superficially applicable to many types of cases. One example of a mid-range theory is what Berkes et al. (2006) refer to as "roving banditry." This describes the tendency of highly mobile resource users to serially deplete a set of resources when they are not dependent on any particular one of them. It has been mostly discussed in the context of fisheries management, due to the highly mobile nature of the resource and the consequent mobility of many fishing actors.

Next, a **hypothesis** is sometimes used to refer to a theory for which there is little evidentiary support. I prefer to use the term to refer to the observable patterns we would expect to find in our data if a theory were true. Essentially a hypothesis as an observational implication of a theory. Ideally we can unpack multiple observational implications of a particular theory so that we can test the theory in multiple ways via a process known as **deduction**. Many texts describe deduction and induction as being entirely distinct steps of the scientific process, occurring iteratively or in a sequence. But in practice they frequently occur at the same time: deductive hypothesis testing via statistical analysis may find unexpected patterns in the data, and the process of data collection, even when it is not explicitly guided by a set of hypotheses, is frequently "theory-laden."

Variables also are found in **frameworks** and **models**. Ostrom (2005) has discussed the difference between these, as well as their relationship to theories, at length. She states: "Frameworks…provide the most general set of variables that should be used to analyze all types of settings relevant for the framework. Frameworks…attempt to identify the universal elements that any relevant theory would need to include." More briefly, Schlager (2007, 294) states that the primary goal of a scientific framework is to "provide theories with the general classes of variables that are necessary to explain phenomena."

A scientific framework is a way of organizing the phenomena under investigation into broad categories, subdividing a rather continuous world into discrete chunks in order to analyze the relationships among these chunks. For example, in ecology, scientists frequently make a basic distinction between autotrophs and heterotrophs to organize their analyses. For institutional economists, concepts such as transaction costs, incentives, information, and rationality are equally important as a way of organizing their view of the world (see Ostrom 2005). In some cases frameworks are formalized and presented in a cohesive package in a particular published work. Indeed, there are many formal frameworks in the ESS literature, including Ostrom's (2007) diagnostic social-ecological framework and the Robustness framework (Anderies et al. 2004). Binder

et al. (2013) recently presented a summary and comparison of numerous social-ecological frameworks.

Models in ESS are similar to theories, but are more precisely formalized. A model is essentially a set of one or more (frequently mathematically) formalized theoretical statements, each of which is related to the others to describe how a system works. Within ESS, agent-based models have become quite popular in the last several years (see Parker et al. 2003; Janssen and Ostrom 2006). In some fields, notably both economics and ecology, the term "theory" ("theoretical ecology" or "economic theory") actually refers to the practice of formalizing theories through the art of mathematical modeling (May and McLean 2007).

Next, we have **units of analysis** and **units of observation**, which are two terms that are easily confused. A unit of analysis for a research project is the category or unit about which the researcher is trying to answer questions. A research question presented at the beginning of a paper will often explicitly mention, or ask a question about, such a unit. Common units of analysis in ESS include individuals, households and communities which are involved in environmental management (Ostrom 1990; Agrawal 2001), and/or affected by large-scale environmental change (Osbahr et al. 2010; Cinner et al. 2012), as well as larger-scale environmental policies, governance systems, and their associated ecological jurisdictions (Keskitalo et al. 2009; Augerot and Smith 2010).

We don't actually examine a unit of analysis in a research project. Rather we study instances of a unit of analysis, or **observations**, and the unit of analysis is the category to which these observations each belong**.** For example, we might want to understand outcomes for a set of trees in a forest. If so, the category of "tree" is our main unit of analysis, and we will probably try to compare different individual trees to look for patterns across these trees. There are two types of such comparisons that can be made: **cross-sectional comparisons** and **longitudinal comparisons.** In a cross-sectional comparison, the observations are different entities (e.g. trees) at the same point in time. In a longitudinal comparison, we compare the same entity at multiple points in time, say each year. So in this case our observations wouldn't be different trees, but different years for one particular tree. If we have panel data, which involves multiple observations of each entity over time, we can conduct both types of comparisons at once. Here our observations would be tree-years.

While a unit of analysis is what we analyze and observations are what we compare, a unit of observation is what we observe. They may well be the same thing as a unit of analysis: we might directly observe trees in order to collect the data needed to compare them. But the two don't have to be the same thing. In order to infer the value of variables describing our observations, we may rely on multiple data sources, or units of observation. For example, if we are trying to compare towns, we probably won't directly observe them, but instead may rely on key informant interviews with town citizens about them. Here the town is the unit of analysis, but we "observe" key individuals who provide us with information about the towns. Additionally, a unit of observation should not be confused with a method of observation. There are multiple ways in which we might try to observe a tree or a forest (e.g. directly with our eyes or through remotely sensed images).

Types of relationships among variables

A defining characteristic of ESS is the emphasis on multiple types of relationships among variables. Figure A2.1 demonstrates several of these graphically. Each graph in this figure shows a relationship between several variables by plotting a hypothetical set of observations along two dimensions, X and Y. We are usually concerned with finding patterns of **covariation** between variables, in which a

change in one variable causes a change in another variable. Each variable in such a relationship can be thought of as a cause, or **independent variable** (IV), or an effect, or a **dependent variable** (DV). Variables per se are not dependent or independent, but may be used in a given analysis in one or both ways: as outcomes to be explained, or as factors that affect outcomes. Or, if the research is purely descriptive and not causal, they can be thought of as neither.

An IV can covary positively or negatively with a DV. A **positive relationship** (Figure A2.1 A) means that an increase in the IV causes the DV to increase as well (slope is positive), and a **negative relationship** (Figure A2.1 B) indicates that an increase in the IV causes a decrease in the DV (slope is negative). If this slope is relatively constant over the range of both variables, then the relationship is **linear** (A2.1 A and A2.1 1B are both roughly linear). If the slope of the relationship between two variables changes at some threshold, say going from positive to negative, then the relationship is **nonlinear** (Figure A2.1 C). The nature of a relationship frequently changes fundamentally as a threshold is crossed.

One source of non-linearity is **endogeneity**, which describes a situation in which a supposed DV in fact causes an IV to change. This may simply be reverse causality, where the supposed IV does not affect the DV, or it may be a case in which two variables are mutually affected by each other, either in a negative relationship, which produces **negative feedback**, or in a positive relationship, which produces positive, self-reinforcing **positive feedback**.

Positive feedbacks, as a source of nonlinearity, are particularly important to recognize, as they create the conditions for a range of behaviors in social and ecological systems, including resilience, path dependence, technological lock-in, and hysteresis (Gunderson and Holling 2002). Each term here broadly reflects the tendency of systems to self-reinforce themselves along a particular social or ecological path, sometimes in spite of a shift in the efforts of decision-makers in those systems. For example, Scheffer et al. (2001) describe a shift from grasslands to deserts that has occurred in many parts of the world. Once grass species disappear, the conditions that had facilitated their persistence, which they themselves enabled, disappear as well. So a desert-like condition may persist, even if human actors remove the initial cause of the transformation (say by removing livestock that had grazed on the grass).

Related to nonlinearity are the concepts of **necessity** and **sufficiency**. These can be understood via the following logical arguments:

If X is necessary for Y, then:
(1) If X is absent, then Y must be absent
(2) If Y is present, then X must be present

If X is sufficient for Y, then:
(1) If Y is absent, then X must be absent
(2) If X is present, then Y must be present

One simple example of necessity in ecology comes from Leibig's of the minimum, which plays a prominent role in agricultural science, and states that plant growth is constrained by the most limited nutrient, implying a necessity of each of a set of nutrients, and a lack of fungibility among them. Some have argued that the roles of distinct institutional arrangements and processes (e.g. property rights, monitoring and enforcement) on environmental and development outcomes are similarly necessary and non-fungible (Kirsten 2009). Within the ESS literature, these concepts are most

closely related with the method of qualitative comparative analysis as promoted by Charles Ragin (1987, 2000), which will be discussed later.

In addition to assuming linearity, scientists frequently simplify their view of the world by thinking primarily of **independent effects**, or the effects that an IV has on a DV, irrespective of changes in any other variables. In contrast, an **interaction effect**, as shown in figure 1 and figure A2.1 D, occurs when two or more IVs interact to affect a DV. This occurs when a **moderator variable** affects the nature or magnitude of the relationship between an IV and a DV. For example, the effects of acid rain on soils depend in large part on the buffering capacity of those soils: the more buffering capacity there is, the less acidic the soil is made by a given amount of rain. In figure A2.1 D, variables Z is the moderator variable that affects the relationship between variables X and Y. Interaction effects are the reason why the most responsible answer to an environmental policy question is usually "it depends." The role of research in this context is to unpack the ways in which "it depends" in a generalizable way, to produce mid-range theories.

A similar, but distinct, phenomenon occurs when the effects of an IV on a DV are mediated by a **mediator variable**. This is also shown in figure 1. For example, in a Dominican fishing community I have worked in, we found that members of the local fishing association tended to catch certain types of fish, and these types were significantly different from non-members. They also fished much closer to shore than non-members. What we ultimately found, however, is that this effect was mediated by the fact that members fished without compressors, and the use of fishing technology played the dominant large role in determining where they fished and what fish they caught. Gear type served as a mediating variable in this case.

The process of mediation in turn relates to the distinction between **proximate causes** and **underlying causes** or drivers, which has been used extensively in the literature on land use and land cover change. A proximate cause is most directly connected to an outcome of interest. An underlying cause is what explains or produces the proximate cause. Underlying drivers affect outcomes via a proximate, mediating variable. Geist and Lambin (2002), for example, identify a mix of political and economic underlying causes of agricultural expansion, which is in turn an important proximate cause of deforestation in many countries. This distinction is basically a way of tracing back a path of processes that lead to an outcome. Underlying causes frequently change more slowly, and are more difficult to change, than are proximate causes. But it can be difficult to change an outcome of interest by proximate causes alone. A related distinction made by many researchers is between slow variables and fast variables, with slow variables, such as soil properties (e.g. phosphorous content) serving as a context for more quickly-moving variables, such as crop production. Walker et al. (2012) comment that fast-moving variables are more frequently the objects of management, just as proximate causes are more easily governed.

Figure A2.1: Types of variable relationships
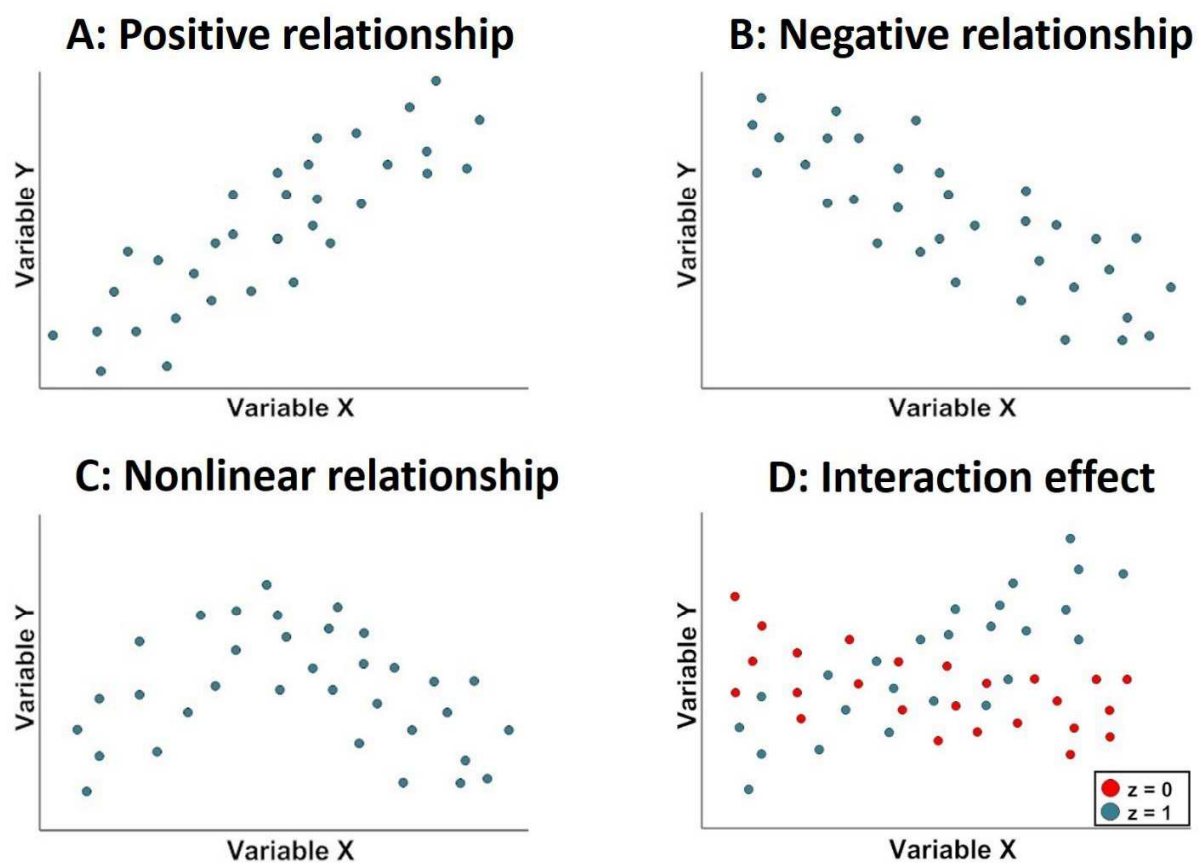
Appendix 3. Balancing induction and deduction.

Induction and deduction are mentioned in appendix 2 as steps in the scientific process. They also represent distinct perspectives within the ESS community, and two extremes of a dimension, along which it is helpful to locate any given research project. Broadly speaking, an empirically deductive research project seeks to apply established theories to new cases. It is highly hypothesis-driven, and has well-specified expectations regarding the patterns it expects to find in the data. It generally will have less room to adapt to unexpected events and new information as the project proceeds. An inductive research project is the opposite: it tends to not be guided by a set of hypotheses, and is more exploratory in nature, attempting to establish new theories from the bottom up. **Ethnography** has become firmly established as a very highly inductive, fieldwork-based approach to social science generally, as well as ESS (Stoffle et al. 1994; Crate 2006).

Deductive vs. inductive approaches to research are sometimes associated with quantitative vs. qualitative research, respectively. While there is some truth to this insofar as quantitative measurement probably presupposes at least some theoretical expectations, I believe this association is also frequently inaccurate. A qualitative case study can be highly deductive, for example, if it is approached with a well-specified set of hypotheses, each derived from a particular theory that it is aimed at critically testing. Additionally, a quantitative analysis can be highly exploratory, such as is the case with several multivariate techniques (e.g. cluster analysis). Within the history of ESS there is a lively (and unfortunately sometimes mutually dismissive) debate about the merits of each of these perspectives. However, it is important to note that no research project is entirely inductive or deductive. Rather, the decision of the researcher is to how heavily tilt their project towards one approach or another.

The approaches of **rapid rural appraisal** (RRA) and subsequently **participatory rural appraisal** (PRA) represent steps that many scholars have taken to increase the inductive nature of ESS: see Chambers (1994) for a seminal discussion. RRA emphasizes semi-structured interviews that combine a certain amount of structure and flexibility in the data collection process. RRA practitioners have developed a suite of techniques to conduct empirical fieldwork, including transect walks and seasonal calendars. PRA takes the approach further to formally incorporate the communities being examined into the research design process.

RRA and PRA represent a perception that, for some time, development-oriented ESS research was overly deductive, and thus failed to incorporate the perspectives of the rural populations that were the subject of much ESS research. Instead, it was only the perspective of the researchers that was seen to matter, a situation that has ethically ambiguous political implications. This issue derives from the fact that ESS is inescapably normative: in conducting ESS we must decide what is socially and environmentally important (and to whom) and what is not. The concern of inductive-oriented scholars has been that deductive research left this decision entirely up to the researchers themselves, without allowing communities to contribute their own perspectives and to guide important aspects of research design and implementation. Highly inductive, and particularly participatory, ESS is characterized in part by endogenizing the design of important research elements into the research process itself, allowing interactions with community members to steer much the work. In sum: while it is important to maintain a deductive perspective in order to ensure that the research conducted is replicable, generalizable, and avoids overly ad hoc theorizing, the researcher should be aware of the unfortunate history this perspective has enabled and the undesirable power dynamics that have been involved in its implementation (see Scott 1998).

Appendix 4. Sample summary and description of a research project.

This appendix summarizes a sample research project in accordance to the research elements described in the paper. The first half presents a summary, and the second half unpacks this summary.

Summary:

Name: Resilience and vulnerability in a set of community-based irrigation systems

Background: This research project examines a set of community-based irrigation systems known as acequias in the Taos Valley of New Mexico. An acequia is a community of irrigating farmers led by a mayordomo and three commissioners, and which uses earthen ditches to convey water to the fields of each farmer. The acequias in Northern New Mexico and Southern Colorado are the descendants of Spanish colonists who began settling in what is now the Southwestern United States in the 17th century.

Research questions:

(1) Which social and biophysical features of the acequias communities have historically enabled the farmers to sustain the collective action needed to persist as community-based irrigation systems over time?
(2) Are the acequias resilient or vulnerable to the variety of novel disturbances that are affecting them in the modern era?

Research design: Embedded case study, including a case study of Taos acequias and statistical analyses of individual acequias and their members.

Units of analysis and analytical strategy for each:

Unit of analysis	Analytical strategy
Taos Valley	Qualitative modeling
Acequias	Statistical analysis
Acequia group-years	Statistical analysis
Acequia members	Statistical analysis

Unit of observation, associated units of analysis, and sampling methodology for each:

Unit of analysis	Unit of observation	Sampling method
TV, Acequias	Acequia leaders via interviews	Purposive and snowball
TV, Acequias	Court testimony documents	Exhaustive
TV, Acequias	Hydrographic surveys	Exhaustive
TV, Acequias	Repartimentio documents	Exhaustive
TV, Acequias	NRCS SSURGO soil data	Exhaustive
TV, Acequias	Taos County assessor data	Exhaustive
TV, Group-years, Acequias	USGS stream-gauge data	Purposive
TV, Group-years, Acequias	Landsat imagery	Purposive
TV, Acequia members	Acequia members via survey	Convenience

TV = Taos Valley

Discussion:

This project has focused on a set of community-based irrigation systems known as acequias that have resided in the Taos valley of New Mexico for several hundred years (see figure A4.1). There are two research questions that have guided this project. The first one asks about the conditions that have enabled the farmers to maintain the collective-action needed to sustain themselves in the face of drought and resource scarcity for several hundred years. Collective-action is a very common outcome examined by ESS scholars, as it plays a large role in affecting whether an environmental commons is maintained or not. The second question focuses on the resilience, or vulnerability, of the acequias to a suite of modern disturbances that have largely resulted from increased economic growth and integration in their region. Many resource using communities that have persisted for centuries are struggling to adapt to modern economic conditions.

The research design adopted for this project is identified as an embedded case study. The first table summarizes the units of analysis that are involved in this project and the analytical strategies applied to each. At the top we have the case itself, which includes all 51 acequias in Taos valley. At this level the analytical strategy applied is the development of a qualitative model describing of the patterns of interactions among various groups of farmers both within and between acequia communities (Cox 2014b). This qualitative model involves an application of the social-ecological framework by Ostrom (2007). Additionally, the project involved a statistical analysis to examine differences across acequias in a vegetation index measured by a series of Landsat images over time (Cox and Ross 2011). Finally, Cox (2014c) examined data from a survey of 107 acequia members, and conducted a statistical analysis of groups of acequias over time. For this final analysis, seven acequia groups were constructed based on the main rivers that fed each set of acequias in the valley, and the unit of analysis is in fact an acequia group-year, with measurements taken at each year for seven groups of acequias.

The final table relates each of the units of analysis from the previous table to the units of observation/data sources that were used to measure important features of these units, and the sampling methods for these sources. Without delving into each of the data sources, it is important to note that, for each unit of analysis, multiple data sources are used. This is common in ESS. Combining the information in the two tables reveals a basic structure that relates the most important elements of a research project: a given project has (1) one or more research questions, (2) a research design, (3) one or more units of analysis, each of which is analyzed with an (4) analytical strategy, and of each which is measured with (5) one or more data sources, each of which in turn is obtained through a particular (6) sampling method.

Figure A4.1: Project study site