
Types of Empirical Studies

Introduction

Many of the things music scholars do is empirical — in the sense that the activities involve observation. Examples might include deciphering manuscripts, studying a score, or observing a gamelan rehearsal. What makes “empirical research” different from the tasks that humanities scholars typical engage in?

A distinction can be made between *formal* and *informal* observation. An informal observation might simply be noticing that there are more women than men on the dance floor of a discotheque. Another informal observation might involve recognizing that a printed score switches from using the abbreviated “sfz” marking to the verbose “sforzando” in the final pages. Formal observation is distinguished by some sort of prior procedure. The researcher *plans* to observe specific things. For example, a researcher might view videos of two performers with the aim of determining which performer moves more while playing. Or a researcher might search through old newspapers from November 1884 in order to determine if any reviews of a concert were published. What makes something Empirical Research is that the observations are planned in advance.

At least seven types of empirical studies can be distinguished: (1) reconnaissance study, (2) descriptive study, (3) measurement study, (4) correlational study, (5) experimental study, (6) meta-study, and (7) modeling study.

1. Reconnaissance Study

A reconnaissance study is a preliminary or exploratory investigation intended to gain familiarity with a new field or phenomenon. Charles Darwin spent five years sailing around the world on the *Beagle*: the purpose was exploration and reconnaissance. The goal was to see what plants and animals existed in different parts of the world.

An ethnomusicologist might travel to a remote village simply to expose her/himself to the culture and music. A historical musicologist might browse through the uncatalogued contents of an archive box, simply to see what is there. A music sociologist might interview some pre-teens about their musical tastes, simple to become aware of their concerns and enthusiasms.

Reconnaissance studies are not hypothesis-driven and do not involve any quantitative

measurement. Reconnaissance studies are common when a researcher begins work on a novel topic or discovers a new phenomenon. The principal purpose of the reconnaissance study is to alert the researcher to new possibilities. The research endeavors to cast a wide net, trying to observe anything that might be potentially informative. Reconnaissance studies may involve some sort of collecting, such as recording music or assembling biological samples.

2. Descriptive Study

Descriptive studies attempt to document and *interpret* some phenomenon. Like the reconnaissance study, descriptive studies are not hypothesis-driven and do not involve any quantitative measurement. The principal purpose of the descriptive study is to understand some phenomenon through detailed description and interpretation. When the phenomenon pertains to people, descriptive studies are usually referred to as *ethnographic studies*. When the phenomenon is a musical work, descriptive studies commonly take the form of *analyses*. When the phenomenon pertains to understanding the past, descriptive studies are commonly called *historical studies*. There are innumerable descriptive methods, including qualitative interviews and participant-observation methods.

Description - Thick and Thin

In ethnography, describing a cultural phenomenon involves some sort of fieldwork in which one establishes rapport with the members of a community of interest, selects local informants or research collaborators, and begins recording various observations, commonly in a field diary. A distinction is made between *thin description* and *thick description* (Denzin, 1989). “Thick description” expands this enterprise, mostly by emphasizing the *context* of various phenomena, interpreting the behaviors in terms of intentions and meanings, and tracing the historical changes of development of the activity. In simple terms, thin description emphasizes what might be regarded as factual reporting, whereas thick description emphasizes how the phenomenon is interpreted or understood within a culture (Geertz, 1973).

Note that descriptive ethnographic studies carried out in the field are often referred to as *field studies*.

3. Measurement Study

Like the reconnaissance and descriptive studies, the measurement study is not hypothesis-driven. However, the researcher engages in some quantitative activity. That is, the researcher counts or measures something.

When a paleoanthropologist discovers the skull of a long dead human ancestor, the first order of business is to describe the skull by reporting a series of detailed measurements. Publishing a detailed description is useful, even if the anthropologist has no opinion or interpretation to offer, and no theory or hypothesis to test.

Measurement studies often present so-called *descriptive statistics*. A descriptive statistic is a measurement without any accompanying interpretive claim. Examples of descriptive statistics: the average age of marriage for women in Libya is 29 years; the average European folksong is 52

notes in length. Measurement studies and descriptive statistics may invite a “so what?” response. Their value usually lies in later studies that make use of the published measurements.

The principal purpose of the measurement study is to assemble quantitative descriptors of some phenomenon, with a minimum of interpretation. A famous historical example of a measurement study is the classic work of the Danish astronomer, Tycho Brahe (1546-1601). Brahe built the most sophisticated telescope of his era and devised ways to measure the positions of celestial bodies with excellent precision. His measurements of the positions of the stars and planets were far more accurate than earlier measures. His work formed the basis for the later theories of Johannes Kepler, who determined that the planets moved in elliptical orbits. Brahe himself did not discover the elliptical orbits of the planets, but his careful measurements provided an essential precursor to that discovery.

4. Correlational Study

Correlational studies aim to identify linkages or relationships between things. We say two things are correlated when there is some sort of connection or association between them. For example, music in the minor mode tends to be slower in tempo than music in the major mode. Although there are exceptions to this, in general, there is a correlation between mode and tempo.

The correlational study involves collecting at least two different sets of measurements, and determining whether there is any relationship between the two sets. For example, a survey of middle-school students might find that more female than male students play flute, while more male than female students play trombone. That is, we might find a correlation between gender and instrumentation. Correlational studies cannot be used to identify causation. The study itself gives us no idea of why particular associations might exist.

A common type of correlational study is the *survey* (although many surveys are descriptive or measurementive rather than correlational). For example, a survey might reveal that people with high incomes are more likely to prefer jazz than country music, or that social conservatives are less likely to enjoy sad music. Once again, correlational studies say nothing about causation: they simply suggest that certain relationships exist.

Exploratory Correlational Study

Correlational studies may or may not be hypothesis-driven. In many cases, the researcher has an explicit interest in testing whether a proposed relationship exists. In other cases, the researcher has no prior hypothesis to test and may be looking to see if anything correlates with a concept of interest. When the study is motivated by an *a priori* hypothesis, the principal purpose of the correlational study is to test an idea by inviting failure. When no prior hypothesis is being tested, the study is referred to as an *exploratory correlational study*. When the study is motivated by a prior hypothesis the study is sometimes referred to as an *experimental correlational study*.

5. Experimental Study

A study is “experimental” when the researcher *manipulates* some aspect of the world. For example, an experimenter may expose listeners to musical excerpts that vary from sad to happy

and observe the effect of the different moods on, say, memory. The property that is manipulated by the researcher (in this case mood) is referred to as the *independent variable*. The property that is observed by the researcher (in this case memory recall) is referred to as the *dependent variable* (or dependent measure).

Experimental studies are nearly always hypothesis-driven. That is, the researcher makes a prediction about the effect of manipulating the independent variable on the dependent variable. When hypothesis-driven, the experiment is referred to as a *true experiment*. All experiments involve some sort of measurement. Experiments may involve more than one independent variables and more than one dependent variables. The principal purpose of the experimental study is to test an idea by inviting failure.

Of all the different kinds of studies, the experimental study is the most highly regarded by empirical researchers. There is a reason for this. The experimental study is the only type of study that allows the researcher to say something about causation.

Exploratory Experiment

An *Exploratory Experiment* involves manipulation and measurement, but the manipulation is not motivated by some prior theory, hypothesis or conjecture. For example, a researcher might play traditional Japanese and Andean pop music to naive Western listeners while making a series of measurements, such as heart-rate, respiration, body temperature, and observable behavior, etc. The researcher may have no idea what to expect. That is, no prediction was made. Nevertheless, having collected the data, the researcher might then carry out statistical tests to see whether a significant increase in body temperature resulted. Notice that the study involves manipulation of the world (playing different kinds of the music), but the study is not motivated by some prior theory or idea.

6. Meta Study

A meta-study is a “study of studies.” It is typically done when a large number of studies have been carried out related to some problem. For example, many studies have been carried out related to whether television violence promotes violent behavior in viewers. Some of the studies seem to show a link, whereas other studies seem to show no link. In a meta-analysis, the researchers identify all of the pertinent studies. They then evaluate the quality of each study, including the quality of the samples used, the number of participants, the quality of the stimuli, the extensiveness of the controls, and other factors. Poor studies are simply discarded if they fail to achieve the minimum quality criteria established by the researchers. Then the researchers combine together all of the good studies, and do a statistical analysis on the aggregate data. The principal purpose of the meta-study is to determine whether all of the studies pertaining to some topic ultimately tell a coherent story.

7. Modeling Study

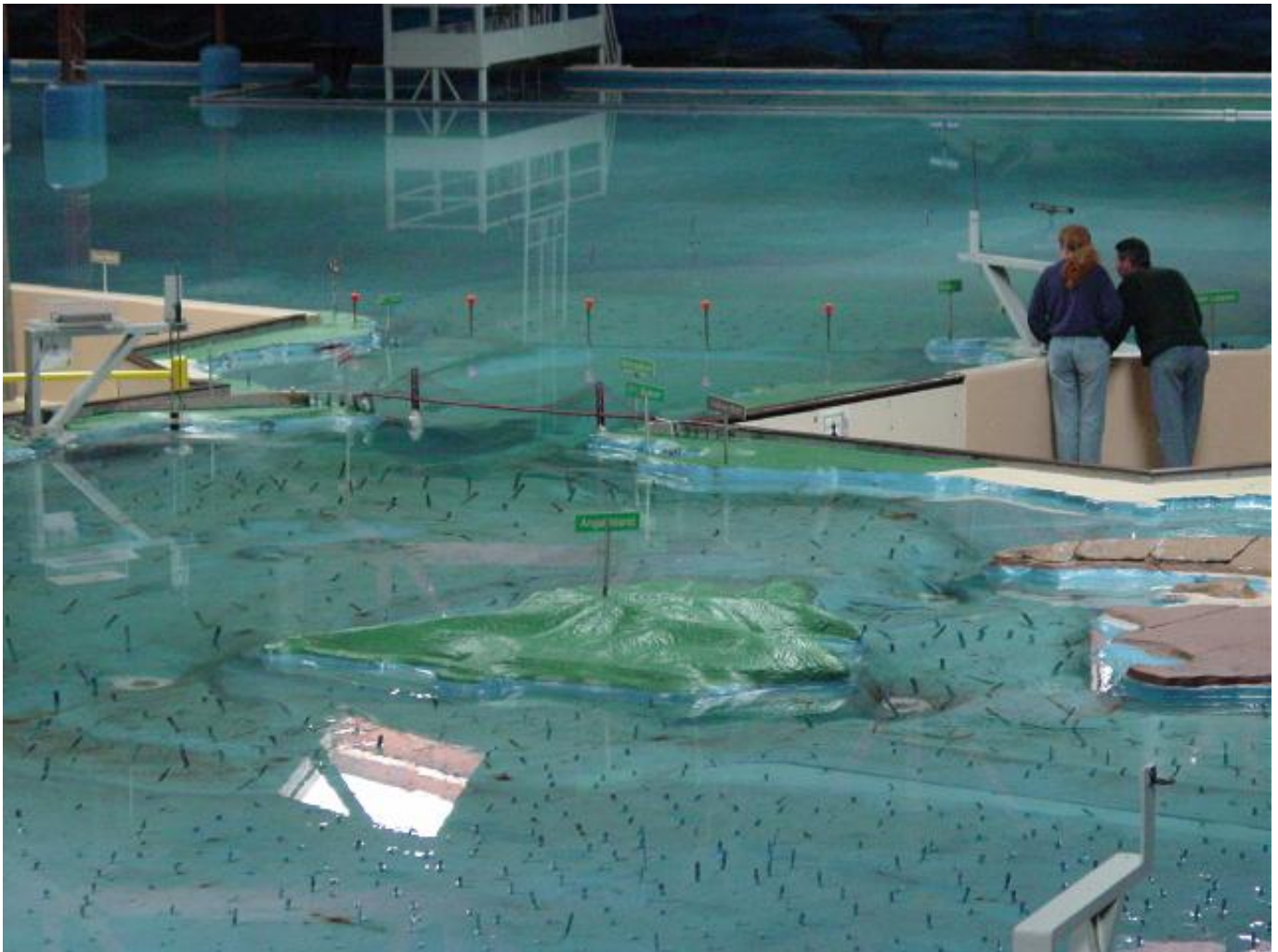
Theories can often be implemented as models. An example of a *physical model* is a large model of San Francisco Bay built by the U.S. Army Corps of Engineers. The actual bay is 100 km long. The model is 1 km in length so the scale is 1 meter = 100 meters. Models are useful for testing

hypotheses that are impossible (or unethical) to test in reality. For example, how long will it take an oil spill in Oakland to reach the mouth of the Sacramento River?

Physical models are rather rare. More commonly, models are rendered as computer programs. There are a number of commercial software applications designed explicitly to help a researcher build a model. Models might be used, for example, to predict the spread in popularity of South Korean pop music introduced into North Korea. Or a model might be used to predict a listener's musical preferences based on past listening habits.

An advantage of models is that they can be used to investigate “what-if” scenarios. The researcher can change some of the initial states, and then see what happens when the model is set in motion. Running a model with a set of initial conditions is referred to as a *simulation*.

The principal purpose of a modeling study is to build a model that has some predictive value related to some phenomenon. In general, few models have been created related to musical phenomenon. Models are more common in disciplines that have a mature research base to build on.



Part of a 1 km-long physical model of San Francisco Bay. (The Golden Gate Bridge can be seen near the center of the photo.)

Pilot Study

Whatever type of study one uses, it is often useful to begin with a sort of “practice” study — known as a *pilot study*. A pilot study is carried out simply as a way of testing the research procedure. Pilot studies usually involve relatively small numbers of participants or small sample sizes. Pilot studies can prove very useful by exposing various unanticipated problems that help the researcher fine-tune an ensuing main study. The principal purpose of a pilot study is to determine whether the full-fledged study is feasible and to reveal possible problems with the research design.

Mixed Studies

It is not uncommon for a research article to report several different studies often mixing different types of studies within a single report. The research might begin by reporting an exploratory study. The preliminary results may inspire the authors to formulate a theory, from which a hypothesis is generated. The article might then go on to report a correlational study or an experiment whose purpose is to explicitly test the hypothesis. Frequently, two or more experiments are reported, with each successive experiment aimed at testing a different refinement of the initial hypothesis.

Natural Experiments

A special kind of experiment is the so-called *natural experiment*. A natural experiment relies on a manipulation of the real world that occurs without the intervention of the researcher. In the field of climatology, the best-known example of a natural experiment relates to the influence of commercial aircraft on heating of the earth’s surface. Everyday, thousands of aircraft fly — creating vapor trails that often produce “linear clouds.” What effect do these vapor trails have on the earth’s temperature? On the one hand, the clouds partially block the sun and reflect light back into space — suggesting that their presence cools the earth. On the other hand, the clouds tend to insulate the earth — reflecting heat radiating from the earth back toward the earth. So what effect do vapor trails have on the earth’s temperature?

The effect (if any) of vapor trails is difficult to observe against the constantly fluctuating general weather patterns. Vapor trails are relatively small, often temporary, and winds in the upper atmosphere blow them so they don’t hover over the same point on the earth. It is almost impossible to measure the effect of a single vapor trail. Ideally, it would be helpful to carry out an experiment. Imagine if an experimenter could manipulate the world — banning all air traffic on one day, and then having the skies filled with aircraft the next day? This would allow the calculation of the average surface temperature over a very large geographical area, and so allow the researcher to test whether the vapor trails tend to cool or warm the earth. This is the sort of experiment meteorologists dream about, but obviously it is impractical.

In the aftermath of the terrorist attacks of September 11, 2001, the US Federal Aviation Authority (FAA) shut down all air traffic across the United States for three days. Although this event was a tragedy of the first order, it offered an extraordinary opportunity for climate researchers and meteorologists to measure the effect of vapor trails on the surface temperature of the earth. Scientist David Travis averaged the daily highs and lows for some 4,000 locations

across the U.S. and compared temperatures during the FAA ban, with data when the planes were flying. This natural experiment allowed researchers to determine that vapor trails raise nighttime temperatures and lower daytime temperatures. That is, at night, the presence of vapor trails reduces heat-loss, but during the day they block sunlight. Moreover, the natural experiment also established that the main overall effect is to lower daytime temperatures. Contrary to the views of some climatologists, vapor trails appear to have a net cooling effect on the earth. The effect is about 1 degree C.

In this case, the manipulation sought by researchers occurred due to other (non-research) circumstances. From time-to-time, it is possible to carry out such natural experiments.

Exploratory versus Testing Studies

The single most important distinction between types of studies is whether a theory *precedes* data collection (a priori) or whether a theory *follows after* data collection (post hoc). Notice that only the *a priori* study invites failure. That is, only the *a priori* study partakes of the rhetoric of prophecy.

Methodologists commonly classify all non-*a priori* studies as *exploratory studies*. This includes reconnaissance studies, descriptive studies, measurement studies, and exploratory experiments. Since exploratory studies don't test prior hypotheses, they are regarded as methodologically weaker. Exploratory studies are quantitative (measurement) studies that do not involve *a priori* hypothesis testing. After the data are collected, tests may be carried out, but these tests are done after the researcher has examined the data and tried to make some sense of the phenomenon. After-the-fact hypotheses are referred to as *post hoc* hypotheses.

There are two common types of exploratory studies.

Most correlational studies also involve testing some hypothesis. Nevertheless, data might be assembled and relationships sought without any over-arching idea of what to expect. For example, a researcher might relate weather data with play-lists from online music services. Once again, the researcher may have no idea or conjecture. Perhaps warm sunny days are correlated with greater listening to dance music; or perhaps poor weather is correlated with greater classical listening. The researcher has some data, and engages in what methodologists informally call a "fishing expedition" to determine whether there are any significant correlations in the data. This type of study is an *Exploratory Correlational Study*.

Disguising Exploratory Studies as Non-Exploratory Studies

A grave methodological error occurs when a researcher engages in an exploratory study, but presents the work as though it were a true experiment or hypothesis-testing correlational study. Notice that this implies that the researcher had the idea before seeing the data and so gives the audience or readers the impression that the statistical tests invited failure. However, since the test was carried out only after the data were seen, there is no chance for the researcher to be wrong. Presenting *post hoc* hypotheses as though they are *a priori* hypotheses is morally wrong in the same way that making-up data is morally wrong.

Taxonomy

In general, it is helpful to classify studies according to four criteria:

1. Does the researcher simply describe and observe? Or does the researcher offer an interpretation or *explanation*? (That is, does the study offer an explanatory theory?)
2. Does the researcher make a *prediction*? (That is, does the method invite failure?)
3. Does the researcher *manipulate* the world? (That is, does the method allow the researcher to infer causality?) And
4. Does the researcher count or measure something? (That is, is the research qualitative or *quantitative*?)

Reconnaissance and *Descriptive studies* involve no manipulation, no hypothesis test, and no measurement. They differ in whether the researcher *interprets* the observed phenomenon. *Measurement studies* involve no manipulation and no hypothesis test — they simply report some sort quantitative measures. An *Experimental study* involves manipulation, testing of a hypothesis, and (necessarily) some measurement. An *Exploratory Experiment* involves manipulation and measurement, but no hypothesis testing. A *Correlational study* involves no manipulation, but it does involve measurement. Usually, correlational studies also involve the testing of a hypothesis. If a correlational study involves no hypothesis, then it is an *Exploratory Correlational study*. (What researchers commonly call an “Exploratory study” is either an *Exploratory Experiment* or an *Exploratory Correlational study*.)

	Explanatory Theory	Hypothesis Tested	Infer Causality	Quantitative
Reconnaissance				
Descriptive	<i>post hoc</i>			
Measurement				✓
Correlational	<i>a priori</i>	✓		✓
Exploratory Correlational	<i>post hoc</i>			✓
Exploratory Experimental	<i>post hoc</i>		✓	✓
(True) Experimental	<i>a priori</i>	✓	✓	✓
Meta Study	<i>a priori</i>	✓	(✓)	✓

References

Norman Denzin (1989). *The Research Act: A Theoretical Introduction to Sociological Methods*. Englewood Cliffs, NJ: Prentice-Hall.

Clifford Geertz (1973). *The Interpretation of Cultures*. New York: Basic.