C H A P T E R   5

From Concepts to Data

CHAPTER PREVIEW

What's It All About?

Chapter 5 takes us from the assumptive and theoretical foundation of a research problem to an overview of the methodological tools that would be available to the analyst in metric or interpretive research and scholarship. It then provides the bases for evaluating the quality of that research and scholarship in a discussion of reliability, accuracy, precision, and validity in metric research as well as coherence, resonance, and vraisemblance in interpretive scholarship.

What Are the Major Topics?

Research practice starts with problems, but problems appear out of the assumptions we make about the domain of media, texts, audiences, culture, society, and the individual. Even as researchers, we are inextricably bound to the ecology of media. How we understand our own experience, particularly our agency to manage our role in that ecology, affects our problem development as well.

   Research practitioners also work inside some theoretical framework, level of approach, conceptualization, or overarching epistemology that shapes both problem development and methodological choices.

   A distinction is drawn between constructs and concepts. Concepts are the ideas from which we form theory. Constructs are these ideas as located within some location of theory and methodology. Concepts can move across theory domains but don't readily lend themselves to measurement or analysis. Constructs are less flexible but give good specificity for measurement and analysis.

   The process of engaging the empirical from either metric or interpretive standpoints requires separate sets of tools and practices. In the metric, the set composes the means of quantification and measurement. In the interpretive, the study of people is conducted through participant observation, and the study of texts is conducted through close reading, coding, and an examination of the textual warrants.

At the end of each of the metric and interpretive methods, the criteria of trustworthy work in those methods are discussed. Metric arguments depend on a level of precision that produces reliable and accurate measurements that are valid representations of the construct under analysis. Trustworthy interpretive arguments require narrative that hangs together in a coherent fashion, resonates with the context of performance, and exhibits vraisemblance as a plausible analysis.

What Special Terms Are Used?

Agency	Axiom	Cognitivism
Critical issue theories	Determinism	False Consciousness

Figure 5.1   Enigmatic Data

3	8	5	7	2	14
2	6	7	7	1	15
1	8	7	7	1	15
6	4	4	9	2	15
2	8	3	6	2	11
2	9	5	6	1	12
2	7	5	5	2	12
3	9	4	8	2	14
6	4	6	9	3	18
5	8	6	5	2	13
5	9	5	5	5	15
5	10	6	4	4	14
4	7	6	4	5	15
5	8	6	5	3	14
2	9	6	3	2	11
6	8	4	2	7	13
4	6	5	4	5	14
5	10	5	2	8	15
8	4	8	6	10	24
2	5	5	5	8	18
5	7	4	2	9	15
4	5	4	1	9	14
6	4	7	5	9	21
6	5	7	4	9	20
3	8	4	3	8	15
6	1	8	9	3	20
9	3	7	9	2	18
7	3	7	9	3	19
9	4	8	8	3	19
8	4	6	9	2	17
8	6	7	7	3	17
10	2	8	9	4	21
5	1	5	9	1	15
10	2	7	10	3	20
8	3	6	7	4	17
7	5	7	6	3	16
5	3	5	3	8	16
8	6	7	6	6	19
7	5	8	7	5	20

THE FOUNDATIONS OF ARGUMENT IN RESEARCH

This chapter is about the materials and tools we use to build arguments. Research would be much easier if we could simply collect the data, run a few summary calculations and some tests for significance, and say, "See, there it is." But just as taking delivery of a load of bricks does not build the wall, so, too, do data have to be used in an argument to reach some conclusion or action step.

We make this claim for two reasons: First, data make sense only when viewed as embedded in some provenance of theory and protocol. It's a silly example, but Figure 5.1 shows what data look like when stripped of this provenance.

While there is some information that can be gleaned by careful inspection, the numbers make little sense until you add in what they are supposed to mean (measure), who (or what) they represent, and how they were collected. The construction of that provenance is the result of human action, not of some objective process.

The second reason for the necessity of an argument is that research is a human practice. Because humans act based on both what they believe to be true and what they believe to be right, data are shot through with uncertainty and necessarily contain error both suspected and unknown.

The result is that when I present data to you, I also have to convince you that the data are trustworthy and appropriate for you to base some subsequent action on. That task is made a lot easier if you and I believe in the same principles. If you believe all numerical measures are junk or that ethnographic field notes are just a fancy diary or that hybrid studies are poaching and preaching, then a study from any of those positions is going to have a difficult time gaining your acceptance. One role of methods textbooks is to show the value of methodology and why it can be trusted in a manner appropriate to its limitations. That work is often done in an "of course it's true" sort of way. I'm hoping to be a bit more cautious, but we'll see, and you'll be the judge.

ASSUMPTIONS PRIOR TO DATA

We have to make some pretty complex assumptions before we can begin to collect data of any sort, and they start well before the first number or field note. We might call these the epistemological foundations of the research problem or, more simply, what has to be believed to be true for a particular problem to be studied. There seem to be three constellations of assumptions that we have to address: The first involves the media-text-audience configuration; the second is the way we configure the relationship among culture, society, and the individual; and the third is the evidence, warrants, and claims configurations—or, more simply, the logic-in-use that supports the research approach. We deal with the third set throughout the book and specifically in the whole of this chapter. Let's complete the triptych by considering a few of the different starting places for media/texts/audiences and culture/society/individuals.

Assumptions About Media/Texts/Audiences

Media

Assumptions about the media play across the devices and the industries. We are certainly aware of assumptions (common in critical theory, for example) about the media as a common conspiracy or under the control of a single set of interests. We may reject that assumption, but we ordinarily think of a medium as a singular entity. Our assumptions about television, for example, often preserve that medium as it was in the 1970s—a limited-production oligarchy, with vertical control of content, and few consumer choices. Thirty years ago, the majority of viewers were limited by their over-the-air service. Now fewer than 15% receive their television off of a broadcast antenna. Yet we still read of calls for regulation (and draconian penalties) that are based only on that small percentage within a much larger system that may soon undergo yet another era of radical change.

How would we better describe television today? Is YouTube television? Is Hulu television? Is the basketball game on my smartphone the same television as the game in my home theater? Am I watching television when I skip the commercials? With a DVR, I can watch a three-hour football game in under an hour. Is that television? Is "television" so diverse that it is no longer a medium? I've used television here, but I think you could plug

in any of our contemporary media—newspapers, radio, film (how long do you think motion pictures will actually be physically on film?)[1]—into the analysis.

Texts

Assumptions determine what we consider to be a text. Is the text the same as the content? Or is the text the interpretation that results in the confluence of content, audience, and culture? And we can push the assumptive level down to ask what counts as content. Is the content of this sentence the words on the page (does layout matter)? The words on the page as presented in a textbook? In a textbook used in a course? In a course in which the reader is enrolled? You get the idea. The questions are "Where are the boundaries?" and "What are the differences that occur across those boundaries?" As a researcher, the most accessible boundary is the words on the page, but that might not be the most effective boundary. Certainly, if I am doing a critical or cultural analysis, that can't be the boundary.

Audiences

The third set of assumptions in this constellation deals with the audience. In his interview with Jon Stewart on *The Daily Show*, Bruce Springsteen noted that when he plays before an audience, he is really playing before multiple audiences.[2] Some are there to hear their favorite hits; some are there for his message; some are there for the party. The only thing many have in common is that they are there. If diversity of purpose is the key to understanding an audience of 50,000 in an arena, how much more of a role does it play in an audience of 30 million for a reality TV program?

Nonetheless, we routinely read about *the audience* in the singular for large, heterogeneous groups of people whose singular commonality may be that a television set is tuned to a particular channel. Yes, it is *the* audience in that sense, but the language masks the wide variations across motive, attention, interest, viewing practices, and environment. The classification of audience or not audience may make no other useful distinction.

We have other visions of media audiences dancing in our head: They concern how smart, how vulnerable, how insightful, and how gullible the people are who form them. The elitist legacy of our field looms large in these assumptions. There is a solid assumptive vein that characterizes the audience for popular culture as something less; that glorifies old media over new; and that considers the under-20 group as foolish and vulnerable (rather than celebratory and imaginative), the under-30 group as shallow and vain (rather than altruistic and passionate), and the over-60 group as disconnected and declining (rather than wise and discerning).

The greatest effect of these assumptions about media/texts/audiences occurs when we connect the dots. A vision of a medium connects to an understanding of its content that is delivered to some sort of an audience. One of the most important questions that are answered in the resultant conceptual figure is whether I can look at a text and declare what

---

[1]Developing and distributing a product on a long plastic strip that has to be mechanically driven past a shutter and a light source at 24 frames per second does seem to be a bit costly given the growing ability of the digital alternative.

[2]http://www.thedailyshow.com/watch/thu-march-19-2009/bruce-springsteen—interview. Accessed April 15, 2011.

it means for others. This assumption is basic to all effects research and to research in the design of messages in general. An alternative assumption is that the meaning of a text is a probability rather than a certainty. The text will mean X a certain percentage of the time it is engaged or to a certain percentage of the audience that engages it. The former suggests that meaning X might be governed by the conditions of engagement, and the latter indicates that meaning X is based on the individuals who engage it.

The assumption that meaning is the same for all gets marked by our interpersonal experience of usually achieving a successful communication. Our experience supports the notion that we can be generally successful in accomplishing our communication goals in face-to-face communication. What we fail to track are the processes of message modification, supervision of interpretation, and confirmation of understanding that are part of interpersonal communication. All of those processes are missing in asynchronous, noninteractive, industry-produced texts. (Few know better than an author in a classroom of how often meanings for the same content differ.)

In the research argument, assumptions are necessarily implied and not stated. We have to start the argument somewhere, and everything that supports that starting place is assumed to be true (at least for the argument). The effects scholar who is concerned about children and violence in the media has a particular set of assumptions about children as an audience, about what is violence and its representation, about the literal character of content, about the relationship between the child and the viewing experience, about the properties and character of social aggression, and about the proper way to measure all of it. Those assumptions permit the interpretations of the data. Change the assumptions, and the data change their implications. My argument here is that such is true of all research. And I believe in the assumptions about the world that make that argument possible.

Assumptions About Culture/Society/Individuals

The abstractions of culture, society, and the individual require us to act from a set of assumptions that allow the logic of analysis to proceed. The following sections consider some of those commonly made.

Definitions of Culture

Wikipedia tells me that culture has a variety of definitions, the shadings of which help distinguish different academic disciplines. For example, the entry author claims that in American anthropology, culture "most commonly refers to the universal human capacity to classify and encode their experiences symbolically, and [to] communicate symbolically encoded experiences socially."[3] At a more disciplinary level, Anderson and Englehardt (2001, p. 57) define culture as a set of seven systems: the semiotic (the system of meaning in language and action), the epistemic (the system of truth making), the ethical (the system of right and wrong), the aesthetic (the system of beauty), the economic (the system of value exchange), the political (the system of allocation), and the social (the system of self, other, and relationships). All culture is composed of these systems, and all cultures have some

---

[3]http://en.wikipedia.org/wiki/Culture. Accessed March 26, 2009.

component that connects to each. Anderson and Englehardt would hold that a society is the particular structurations (configurations of actions, structures, and resources) that provide for the expression of those systems in a particular time and place.

Relationship Among Culture, Society, and the Individual

Concepts like culture, society, and the individual have multiple definitions because they are abstractions that are put in the service of some intention. What appears to be important in media studies is not so much the particular definition, but the relationship among the elements. For the critical-cultural scholar and often for the ethnographer, the order is culture, society, and individual in that culture provides for the society and the society provides for the individual. Consequently, one's analysis starts at the cultural level rather than the individual level. The individual is simply the route we have for understanding the cultural.

Psychology in the United States has traditionally taken the opposite view, in that the individual provides for both society and culture and, in turn, is modified by each. We study the individual because the individual is the fundamental building block of culture and society.

How does this difference make a difference? I think we can see the difference by looking at effects research in comparison with cultural analysis. Effects research depends on the assumption of literal referentiality of content (meaning of content is fixed, and we all know what it means) and on the assumption that an individual can be changed from one predictive state to another by exposure to that content. The effect, then, occurs when the probabilities and rates of one performance (rather than another) have changed for that individual. The individual is, in short, one thing before exposure and another after exposure.

A cultural analysis would look at the relationship between media content and the individual and start by saying that both are embedded in a common field of understanding where only the particular arrangement of elements can be new. (Even now, I know what the weather will be on the news tonight; it's just the numbers that change.) Anyone who is culturally competent (see box, left) already knows the story about sex, violence, drug use, AIDS, and so on; the surprise is how you will tell it this time. So content doesn't change things; it lays down another layer on an already existing multilayered platform of understanding. Consequently, watching sexually inflected popular television will not lead one to more sex. There is no position of innocence for any of us. The possibility of more sex is already there. Our cultural values provide both for the program and for the possibility.

Note that the data we collect from individuals about the relationship between watching sexually inflected television programs and participating in more sexual activity might look exactly the same, but the interpretation that we would give them would be substantially different. In the one, content changes a person's beliefs about sexual behavior and leads to greater promiscuity.

The Concept of *Competent*

Competent (and *competence*), as used in theory, refers to the quality of being recognizable as being wholly formed, correctly done, capable, and the like. Theorists use the term to exclude the incomplete, the deviations of error, the random, and the unrecognizable from the scope of their explanations.

In the other, sex causes sex, and culture is the basis of understanding both how content can be considered sexually inflected and how a person could be considered more promiscuous.

Agency

There is one other element that is involved in this difference and that entails the issue of agency. Agency is an individual's ability to do otherwise. From a strict cognitivist point of view, an individual has no agency. There is no central faculty of the will that can invent a different set of probabilities from what the existing cognitive structures provide. Our action is not by choice but by the confluence of circumstances that bring one set of probabilities into play over another. If we could analyze the situation well enough, we could see how each action is determined by the particular conditions at the moment of its performance.

Few analysts choose to maintain that strict a position, but the concept of agency remains a problem for much of social science but particularly for the cognitivist and the effects scholar. Understand that if I really can "just say no," then an effect cannot be driven by content. It becomes a choice. There have been two traditional solutions to the problem. One is to focus on respondent groups with supposed diminished capacity of the will (reason) such as children or, sadly, in our history whomever is designated as the Other. The second is to hold that one can "say no," but he or she does not have to "say yes." Media influence can be blocked, but it takes effort to do so. In the absence of that effort, media can effectively influence their audiences. Either or both of these positions form the basis of most media literacy programs and other interventionist strategies.

The early work of critical theory and cultural studies suffered from the same conflict between agency and determinism. Early critical theorists were diviners—people who were inspired to see things that others could not. The critical theorist could see the influence of the dominant interests by their control of media; the masses were under the rule of "false consciousness"—a self-deceiving state of mind that allowed the media to work their influence. This position was pretty much undermined by the appearance of the postmodern critique, which essentially asked, "What makes you so special?" Stuart Hall and the Birmingham School coined the term *cultural dupe* to express the cartoonish figure that inhabited early critical theory and introduced the concept of resistance to the dominant readings, thereby recouping some form of agency for us all (see Edgar & Sedgwick, 2005).

Regardless of epistemological standard, it is difficult for the media scholar and student alike to grant agency or reasoned action to the audience. Part of that resistance comes from the critical impulse that is central to science and scholarship to take things as they seem and reveal what they actually are to an otherwise unknowing reader. We scholars do have advanced technical skills that generate higher-level insights. (If we don't, then we are frauds.) It is easy to look down from that higher ground.

Summary of Assumptions

An assumption is a statement that is necessary for us to believe as true in order to conduct some line of reasoning. In logic we call it an axiom. Any enterprise in scholarship or research has to start with some set of assumptions. (One assumption of the postmodern critique is that there is no solid ground—there is no epistemological standpoint that is not

supported by assumptions.) The assumption set is the ultimate vulnerability of any argument; refuse the assumption, and the argument is moot. I believe it is in each of our own self-interests to reveal as much as we can about the assumptions we hold concerning all of the elements in media research. It will help us understand the lines of research we accept and those we don't.

You might consider writing out the answers to a set of questions about assumptions just to clarify your own thinking. For example, consider these question-and-answer sets: What is the starting point for an analysis of the consequences of media engagement? Possible answers might be culture, society, the discursive domain, the individual, content, the content-medium combination, or the interaction between content and the individual. What best describes the conditions under which content is effectively encountered? Simple exposure, attentive presence, or active engagement, as activated in an interpretive process? Are the boundaries of a medium defined in its technology, the combination of technology and the forms and conventions of content, the intentions and narrative forms of the text regardless of delivery system, its constitution in society and culture, or its history because media are indistinguishable in this era of convergence? Do media in their industrial texts produce a consistent set of cultural values, offer a heterogeneous and disconnected set of values, or ignore values except as economically important?

There are dozens of questions like these to be asked (notice that I have asked one question each for consequences, practices, properties, and value), and the answers you develop will guide you toward a particular theoretical and methodological path. It's not a take-out menu, however. The answers have to form a coherent and noncontradictory foundation for research. The pedagogical impulse is, of course, to direct the reader to an appropriate set of selections. This can be effective reproductive teaching, but it does not support critical thinking.

POSTASSUMPTIVE THEORY

All research problems are embedded in some theory whether explicitly stated or not. Theory develops on the foundation of assumptions or irreducible axiomatic beliefs. Theory starts when we assemble a more or less coherent base to consider the implications in some domain. There are a very large number of theories or theory-like claims in mediated communication. In the next sections, we spend some time with the variations.

Theory Types

Given the variety of positions that can form the foundations for theory, it is not surprising that there is not a singular domain of analysis that we could call media theory. Media theory is a much divided territory with border skirmishes throughout. The major states are (a) cognitivism, an inwardly looking set, which houses theories of messages and effects that are adapted to metric methods; (b) social action, an outwardly looking approach in which we find interactionist and interpretivist theories that are adapted to ethnographic methods; (c) psychoanalytical-semiotic theories that are concerned with the mechanisms of understanding and meaning and are drawn to discursive methods; (d) critical issue theories such

as Marxism or feminism or race studies with their focus on class, gender, and race or ethnicity as the center of concern in media analysis and a general affinity for any method but metric; and finally (e) critical-cultural theories that are an eclectic membership that borrows extensively from literary and rhetorical criticism as well as critical issue and psycho-analytical-semiotic theories and generally uses discursive methods.

Levels of Approach

But, of course, we are not done because each of these domains can be addressed at the psychological, sociological, and/or cultural level. The psychological is concerned with the formations and behaviors of the individual, the sociological addresses the functions and performances of societies, and the cultural targets the understandings that bind individuals to a society and across nation-states. One could build a tidy table and fill in the cells with particular named theories such as agenda setting at the intersection of the cognitive and the psychological.

Initiating Concepts

Unfortunately, even that complexity is not enough because we can further parse theories according to their starting place in the nexus of media, content, message, audience, and performance. If one starts with media, one goes into McLuhanesque theories of hot and cold media or into technological determinism. If one starts with content, the move is to typologies such as formalism, master contracts, genre theory, and forms and conventions. If message is the starting place, one moves to effects and cultivation theories. Starting with the audience (or auditor or spectator—one audience member) leads to a field called audience studies or to uses and gratifications theories. Finally, starting with performance (the culturally recognizable things that people do) leads to social action theories in which media and their messages are part of a larger ecology in which we act out our lives.

Epistemic Divisions

And finally we have the epistemic divisions of Cartesianism, modernism, and postmodernism. Cartesianism (after Rene Descartes) holds that reason is the foundation of true knowledge, modernism would look to objective science as that foundation, and postmodernism would argue that human performance is the source of human knowledge, discarding both capital-*T* truth and objectivity. Theories and methods can be distributed along these divisions. Postmodern ethnography, for example, is quite different from modernist (or material) ethnography.

It may be self-serving, but I believe that it is impossible to master this complexity. The result is that just like the communities of methodologists we met in Chapter 1, individual theorists tend to gravitate toward some intersection and hang around with others on the same corner. We are a disparate community, and I'm not willing to say that any group has strong answers to the questions we pursue. Nothing in the theory of 10 years ago would have predicted the appearance of Twitter (nor can we now predict the causes of its

inevitable disappearance). Consequently, as scholars, we take up a theory knowing that it has limits, flaws, and many alternatives. That doesn't stop us from a passionate and extended relationship with a set of ideas that would flow from our theory of choice, however.

CONCEPTS AND CONSTRUCTS

Assumptions provide the foundations for theory, and theory provides the concepts and constructs that populate our research problems. The terms *concepts* and *constructs* are often used interchangeably, but there is some value in maintaining a difference. Concepts can be viewed as ideas that populate theories—the "agenda" in agenda-setting theory is a concept of our thinking-speaking-performing menu. Constructs can be seen as ideas that populate methodologies—an agenda as a construct is a predictive cognitive state that can be measured on a scale in metric methods or is an emergent understanding of performance in ethnographic ones. Constructs are concepts encountered in some context of method.

Concepts can travel across theoretical domains. I can take the idea of media producing a common knowledge base (an agenda) on which a social interaction might develop (giving us something to talk about) and leave behind both the negative implications of a lessened ability because of media influence and the need for a cognitive state limiting us to that agenda. Exporting this liberated concept of agenda, I can explore how those of us into ballroom dancing (have I revealed too much?) can use *Dancing with the Stars* as the basis of conversation without constraining our evaluations of the dance or the show.

It is much more difficult to manipulate constructs because they are anchored into a particular location in the complex matrix we described in theory. Take the idea of message effectiveness. Traditionally, this concept is anchored in a transmission model of communication (e.g., who says what to whom with what effect). The transmission model depends on us being able to specify the "who," "what," "whom," and "effect" unproblematically. Traditional message effectiveness depends on the "what" moving from the "who" to the "whom" in a manner that maintains the properties of the message. The analyst wants to be able to specify the properties of the message upon construction and prior to delivery and then observe those same properties in the receiver. Message effectiveness in this example is a construct because it depends on the delivery of a set of sender-developed intentions and properties through a message constructed for that purpose to a receiver capable of reproducing those intentions and properties. It specifies the requirements of its measurement in great detail.

What happens if we pull effectiveness out of the transmission model? Is the construct of effectiveness robust enough to survive as a concept in a different epistemological location? For me, the fundament of effectiveness is consequence. Something happens as a result of something else. I don't have to start the analysis of what happens with the intentions of the sender, however. I could start with the intentions of the receiver (as uses and gratifications theory would direct us to do) or with the culturally influenced interpretations produced under the supervision of an interpretive community (as social action theory might suggest). What would message effectiveness look like under those analytical conditions?

Let's try one more example: the third-person effect. The third-person effect is an interesting amalgam of cognitivism and psychoanalysis. It holds that in an agenda-setting/cultivation analysis/media effects sort of way, we are aware of the consequences of engaging media (increased violence, body image, stereotyping), but instead of attributing those consequences to ourselves or our own, we project those consequences on others and particularly on the Other. What happens when we pull it out of its effects location? What if we do a mash-up with it and the spiral of silence (the idea that the more different we perceive our own opinions to be, the less likely we are to express them)? Or what happens if we see the third-person effect as a coping mechanism to deal with the disconnect between what we are told are the effects of media and what we see are the effects of media?

In finding answers for these questions, the evidence for the third-person effect remains the same, but the argument we can construct—the knowledge that we gain—is quite different. That is the value of concepts over constructs. Concepts can be manipulated without a lot of theory or method baggage. It is also their weakness. We have a number of communication theories that are conceptually exciting—chaos theory (for dealing with the self-organization of new media), systems theory (for dealing with media ecologies), and structuration theory (for understanding media texts as resources) are three that come to mind—but at the same time we have not been able to develop the constructs that will allow a robust program of research to develop. They remain ideas without research practices.

The quantitative-qualitative debates of 30 years ago (with pockets of resistance still) were part of the process by which interpretive empiricism moved itself from an idea to a research practice. We now have conventionalized criteria for what good auto, interpretive, and critical ethnography as well as close textual readings ought to look like (even if there could be far more evidence of the actual practice). Probably the strongest evidence we have for the existence of constructs in a field is the availability of computer software to provide an analytical framework. Metric studies are supported by SPSS, SAS, and SATA as full-featured data analytical schemes, and interpretive studies are supported by ATLAS.ti, Framework, and NVivo 9 as text-based (multimedia) analytical methods. (As a side note, media effects scholars should consider investigating the effect of computer software on the research process.)

ENGAGEMENT, EVIDENCE, CLAIM, AND TRUSTWORTHINESS

We will split the discussion under this heading into metric and interpretive, but I want to point out at the beginning that the problems that are faced in metric and interpretive empiricism are relatively parallel, albeit under different names and procedures. Each has to engage the empirical whether it is considered objectively material or semiotically constituted. Each has to capture that engagement in some form of evidence. Each has to demonstrate the consequential value of that evidence in a claim about the empirical produced in the research argument. And each has to demonstrate the trustworthiness of the claim.

We engage the empirical in our research protocols, which will be the subject of the next two sections and constitute argument in the processes of publication that we will only brush against in this book. Here we look at the formation of evidence and the criteria of trustworthiness.

Metric Evidence

In the metric domain, evidence is generated through the quantitative measurement of variables, which in turn produces the data that move through the statistical analyses of the protocol. In the following sections, we take up variables and their constitution in measurement, quantification, data, and instruments.

Figure 5.2 Netflix Ratings Legend

Ratings Legend

★ ★ ★ ★ ★
Loved It

★ ★ ★ ★ ☆
Really Liked It

★ ★ ★ ☆ ☆
Liked It

★ ★ ☆ ☆ ☆
Didn't Like It

★ ☆ ☆ ☆ ☆
Hated It

⊘ Not Interested

Not interested in this movie

No Opinion

I don't know enough about this movie to have an opinion

Source: http://www.netflix.com/

Variables

A construct is an idea that has been positioned at some point in the theory-method matrix, and a variable is a construct that can take different values in the measurement processes of a particular study. (Constructs that don't vary within a study are either ignored or called constants.) Variables become visible in the measurement process. The fact that a variable can be measured is evidence of its existence, and the reliable differences in the variations of values are used as evidence of the performance of the variable under the conditions of the study.

Measurement

Variables exist as ideas; measurements exist as processes or operations. The connection between the two is managed by the operational definition. The operational definition is a philosophically contentious but common practice used to complete the connection between theory and data. Again, it is easy to think of it as a line that starts with a concept, which then is located in the matrix as a research construct and more specifically as a problem variable, and then makes its operationally defined appearance in some measurement device and practice. Netflix, for example, asks you to rate the movie you just watched. It gives you a 5-point rating scale with two "disliked it" choices and three "liked it" choices ranging from "Hated it" to "Loved it" (see Figure 5.2).

That is their operational definition of some like-dislike construct. Of course, it is marketing and not research, but the practice is the same.

Professional researchers would criticize this operational definition because it does not appear

to make a good connection to the like-dislike construct. It is a single-item scale evaluating a complex text. It appears to deliberately bias the answers toward liking, and the distances between intervals do not appear to be evenly spaced. We would say that it lacks construct validity. It does not make a good connection to what we expect to be the idea of liking a movie. Netflix may be just delighted with it as a marketing tool, however.

Quantification

Measurement processes connect to theory through operationalism, produce data as their product, and, in metric studies, produce those data in numerical form through quantification. Quantification is the process by which we attach the properties of the object or individual under study to the properties of numbers. Numbers have four characteristics: uniqueness, order, distance, and relative value. These characteristics get used to produce different kinds of data. We call the four types of data nominal, ordinal, interval, and ratio data.

Nominal Scales. Nominal data make use of the uniqueness of each number. Each number is its own set, and its set does not overlap with any other set. Figure 5.3 shows five circles, each representing the boundaries of a set and each with a separate number inside. They are floating in figure space in no particular arrangement because we have no information as to arrangement. The utility they have for us is that each one is unique. If the number is of one value, it cannot be of another. These are called nominal data. We use nominal scales when the properties being measured fall into unique categories: true or false, yes or no, male or female. More technically, we use the unique name of each number (hence *nominal*) to identify a unique property of a class of entities.

We use it improperly when, say, that yes or no could also mean maybe. In that case, both the yes and no choices might also contain some instances of maybe. As a result, they are not truly independent of one another. We would say that this circumstance is a source of measurement error. To correct this error, we could add a third unique category called maybe. Does the researcher know when he or she makes this kind of error? Generally not, and most ordinarily not until after the fact of data collection when it is discovered that the variable doesn't "act" as predicted.

Nominal scales allow us to count things—give the number of things classified as, say, 9—and to know that the count of 9s does not affect the count of, say, 7s. They are independent
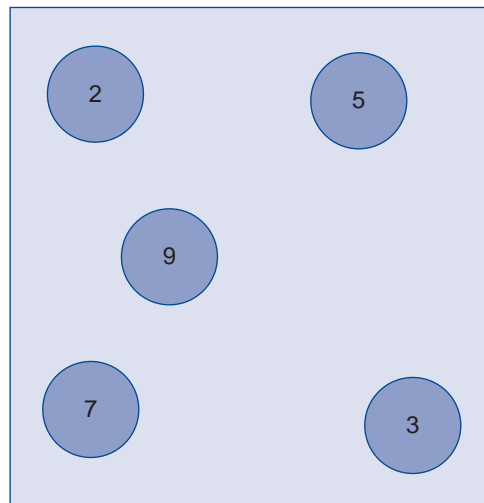

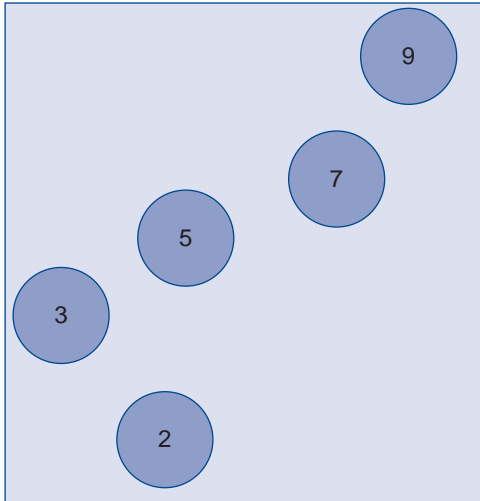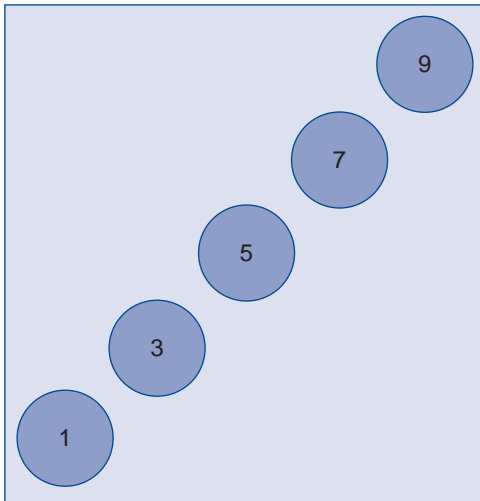**Figure 5.3**   Nominal Scale

Figure 5.4   Ordinal Scale

categories. The average that reflects the most common value in a set of categories is the mode (the largest class), and the inferential statistics in use are chi-square and tests of proportion.

Ordinal Scales. The second property of numbers is that they can be arranged in order of greater and lesser as I have done in Figure 5.4. When used in measurement, this *ordinal scale* indicates more or less of something. Each point on the scale represents a unique value of whatever is being measured, and that value is ordered in relation to the other values on the scale. The way we identify the ordinal value is different from the cardinal number terms. In the figure, we would call the number 9 first, 7 second, and 2 fifth or last. An ordinal scale makes no claim about the distance between the points on the scale; it claims only that they increase or decrease.

In the case of the Netflix scale, the dimension in play is the more or less "likability" of a film. The Netflix scale is a good example of what should be treated as an ordinal scale. No attempt (in my reading of it) has been made to make the descriptors of the points equally spaced in the cognitive state of liking the film, but clearly the values move up.

Ordinal values allow for all the analysis that nominal values do—one can always move down in the assumptions of properties. In addition, there are tests of the median (the halfway point of occurrences) and of ordinal relationships.
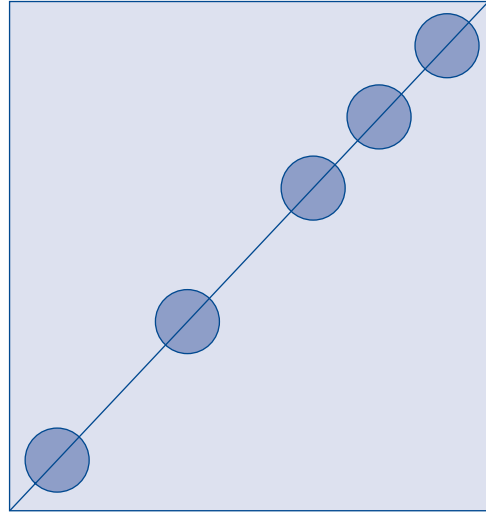
Figure 5.5   Interval Scale

Interval Scales. The third property of numbers takes up the distance (the interval) between the numbers. In not-so-simple terms, the principle holds that one can unitize a scale of numbers such that the relationship between any contiguous pair of points on the scale is the same. If one has a scale of integers ranging from 1 to 10, then the distance from 1 to 2 is the same as the distance between 9 and 10. Figure 5.5 shows a scale in which the points are two units

apart, but the relationship between 1 and 3 is the same as the relationship between 7 and 9. This sort of scale is called an *interval scale*. Each point on the scale represents a unique value, the values are ordered along some dimension of more or less, and when the points are displayed in equal intervals, the distance between the points is the same. If we were to consider the Netflix scale an interval scale, it certainly would not be an equal interval scale. I think it would look like Figure 5.6 with discontinuities on the dislike side. I would much prefer a 7-point scale as in Figure 5.7 with missing units (in dark blue) added in.[4]

YouTube, by the way, has a similar problem with its Poor/Nothing Special/ Worth Watching/Pretty Cool/Awesome! scale. Again there are two negatives to three positives, and how does one go from Nothing Special to Worth Watching in one step?

Why does all this matter anyway? Good measurement is the foundation of good results. The goal of quantification is to use the properties of numbers to represent the properties of the object under study. Both Netflix and YouTube inflate their ratings (which may the source of multiple disappointments), perhaps because getting people to watch something is how they make their money. If a similar error is accepted into the scholarly archive, it tends to reproduce itself. If it becomes authoritative, it cements the error in as others repeat it. In consulting, the consequences are often much more immediate when recommendations fail to produce the results expected.

Figure 5.6 Interval Scale With Discontinuities

Figure 5.7 Seven-Point Interval Scale

---
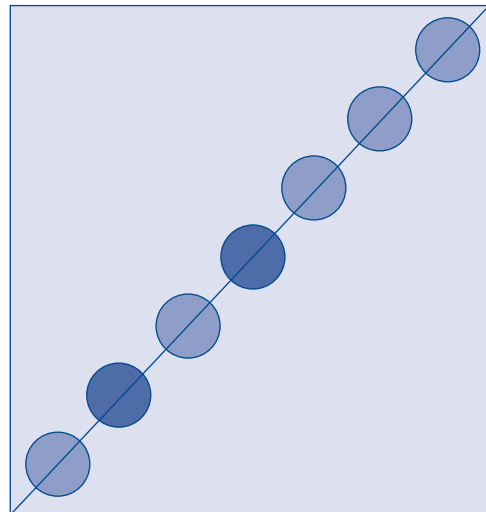
[4]Or perhaps, better yet, a slider across the values from 1 to 10.

Equal (appearing) interval scales have been more or less standardized in metric research. Quite often the scale is composed of a declarative sentence followed by an agree or disagree response element. For example:

I like reading about the properties of numbers.						
⎯	⎯	⎯	⎯	⎯	⎯	⎯
Agree strongly	Agree	Agree slightly	Neither	Disagree slightly	Disagree	Disagree strongly

Another approach is to anchor the endpoints and let each respondent divide the intervals equally according to his or her own evaluation of the relative distances. This approach is called *ipsative measurement*, and it avoids the semantic problems of labeling the intervals, but it also means that your value of 5 may not be the same as my value of 5. Consequently, we cannot claim that a movie that was rated a 4 is twice as good as a movie that was rated a 2. We could say it was two units better, however, without any notion of the size of the units.

Despite the uncertainties and the high potential for error in using interval scaling, it is the measurement of choice. It is preferred because this form of quantification not only allows us to count and order entities, but we can also apply nearly the whole range of descriptive and inferential analyses. More than enough will be said about that later.

Ratio Scales. The fourth property of numbers is the zero. Zero is a relatively late arrival in the pantheon of numbers developed by Hindu mathematicians in the second century of the Common Era. Zero is both a placeholder and the representation of the absence of some property. When there is a natural absence of a property, we can take ratios. A *ratio scale* depends on the presence of a natural zero value. Age, height, weight, income, and the number of magazines received in a home are measures with natural zeros. Note that all of these measures have some material character about them. They exist in front of the eyes.

The great advantage of ratio scales is that measurements can be referenced to some known value. According to the Association of Magazine Media (www.magazine.org), the average number of issues read per month (subscription multiplied by appearance per month) is a little over 11. I read 16 (3 weekly magazines and 4 monthlies), not counting academic journals. My issue index would be 1.45 (16 divided by 11), indicating that I read 45% more magazine issues than average. That sort of claim would not be justified with an interval scale.

There are problems with ratio scales, of course. Most of them have to do with very small numbers or with a shifting base. There was a time when my daughter was .04 times my age. Now she is 0.65 times my age. (If we plot those points out, she will be older than I in a short while.) Of course, the number of years between us has remained the same, but our relative rates of change have not been the same. Her age doubled 5 times during the time it took for mine to double once. (If you remember your algebra, you can figure our

respective ages.) Growth rates, relative growth rates, and changes in probabilities are all susceptible to these errors in interpretation because with small numbers, small changes can make large differences in reported values. There are other issues such as the question of scale: If you take 4 magazines and I take 2, you have twice as many. If you take 400 magazines and I take 200, you have twice as many, and we both have a serious problem. Consequently, we have to exercise some care.

Measurements that have a natural zero allow us to use all the properties of numbers to represent the properties of the objects measured. We can count them, order them, space them, and consider the values relative to one another. I'm a big fan.

Instruments

The last element in our list is (instruments). I've put the term in parentheses because it is most often left out of the discussion. Like servants in a 19th-century British household, instrumentation is the invisible but necessary element that makes the whole enterprise possible.

As little as three decades ago, we might have been excused this oversight because the majority of instrumentation was "paper and pencil." Such is no longer the case, and in fact there is little reason to use such primitive methods. True to form, there will be little discussion here because it would overwhelm this section (more later on, however). Nonetheless, instrumentation is not transparent. It is face-to-face or mediated communication with all the complications that communication entails.

Metric Trustworthiness: Reliability, Precision, Accuracy, and Validity

Because the common assumptions of metric measurement include the notion that whatever is being measured has something of the characteristics of a material object, the questions of the reliability, precision, accuracy, and validity of the measurement come into play. A few definitions: Reliability is the ability to repeatedly get the same values from the measurement of the same thing. If my desk measures 72 inches at Time 1, it should measure 72 inches at Time 2. If it doesn't, there are two theoretical possibilities: Either one of my measurements is in error, or the object has changed in dimension. More on this later.

Precision is the level at which a measurement is both reliable and accurate. U.S. government standards hold that a tape measure has to be accurate to plus or minus one sixteenth of an inch over 6 feet (72 inches). That means that any measurement between 71 and fifteen sixteenths of an inch and 72 and one sixteenth of an inch is the same within the tolerance standards, but don't try to build furniture at that level of precision (use the same tape measure throughout).

Accuracy is the ability of a measurement to reproduce some standard. If I build a 72-inch desk consistently using a tape measure that is one inch short, my measurements will be reliable and very precise (I'm bragging), but inaccurate. And validity is measuring what you claim you are measuring. If I measure the top of my desk by pulling a calibrated tape measure over the top of the books and papers piled upon it, my measurements will be reliable, precise, and accurate, but they will not be of the top of my desk.

Now we need to walk these ideas over to a 7-point Likert scale measuring a socially constructed, nonmaterial, composite variable that resides in some communication context

held in place by human enactment. As the sentence implies, this is a difficult task. People do, however, behave in consistent ways, that behavior reaches some standard of performance, and the performance is often what we expected. Consequently, human behavior can be reliable, accurate, and valid within some level of precision. If we claim to be measuring that behavior, then our measurements should show the same characteristics.

Reliability is the starting place. If the analyst cannot repeat a set of measurements, either the measurement device or the behavior is unreliable, and in either case we do not know what is being measured. Game over. If the two measurements approximate one another, we will say they are reliable within some level of precision. The approximation device is usually the correlation between the two sets of measures—called a reliability coefficient. Reliability coefficients in the .70 range are routinely reported in the literature. That value means that a little less than half of the variability between the two sets of scores is shared. This is not a very high level of precision, but it is what we work with.

If the measure we are working with correlates with another accepted measure ostensibly of the same behavior, we can claim some level of accuracy. We call that concurrent validity. From here it gets a little murky. We attribute face validity to any measurement that looks like it should be measuring what we hope it is measuring to both the analyst and the respondent. Content validity is the level of test specification that is achieved in the measurement—the measurement covers all known aspects of the intended construct. Content validity can be only as good as what current knowledge of and consensus about the construct will allow. Predictive validity appears when the measurement predicts what it intends to predict. This form is also known as instrumental validity—the measurement does what it says it will do. Finally, a measurement achieves construct validity when the preponderance of the evidence supports its use as a measurement of the construct. This is the process known as consilience, and it is a form of probabilistic epistemology. It essentially says that if it works in all these cases, there must be something more than chance involved. It cannot, however, unproblematically specify what that "something" is.

There are three other forms of validity that enter into the research process. These are the levels of validity that are reached by the research protocol—the entire set of procedures in data generation. They make the most sense for experimental protocols but are used elsewhere as well. The first of these is internal validity. Internal validity is achieved when the protocol meets the standards of causal inference. External validity extends that success to other situations. Ecological validity means that the causal relationship will hold in the ecology of natural performance. We will deal with these issues in Chapter 12.

In the final analysis, we can provide good evidence of reliability but only indirect and often unsatisfactory evidence of validity. Because of this conundrum, we worship at the altar, but don't spend a lot time on the issues.

Interpretive (and Hybrid) Evidence

Interpretive and hybrid arguments depend on the skillful insights of the researcher or scholar responsible for the interpretation. The principle is that the researcher is the instrument. That instrument is not some Cartesian individual—a right mind thinking well—but rather a culturally located and informed agent engaged in somewhat conventionalized practices that are intended to provide the resources of the interpretation. In the end, the argument will stand

or fall on the ability of the researcher to produce a coherent and resonant narrative. In the in-between, there will be plenty of evidence for that success or failure. We start this discussion with the observation and participation of ethnography and then move to the engagement of texts.

Observation and Participation

All ethnography is based on a systematic observation of the practices and performances of some congregation of individuals who share common (but often conflicted) understandings of those practices and performances. The gold standard of that observation is, perhaps surprisingly, participation.

A definitional note here: The term *congregation* and its forms refer to a group of individuals that are semiotically connected in meaningful practices and performances that the members cocreate. They are bound together in commonly understood beliefs, norms, and actions. A congregation would contrast with an aggregation.

The congregate group may be composed of a few to a very large and unknown number of individuals. The observational work that can be done depends on the number of different memberships that run through the group, the access that is granted to those memberships, and the ability of the ethnographer to grasp the significance of the differences that define the memberships.

For example, let's say I wanted to study the use of e-mail and networking sites across a large university in some Midwestern state. Would I expect the use of these communication technologies to be the same in the department of communication as in the department of mechanical engineering? Would I expect these departments to grant me the same access to their e-mail and sites? Finally, would I expect to understand the differences that might appear across these departments? I don't know the answers to those questions, but I better not start with any of those expectations.

Once admitted to any congregate group, then, the analyst is faced with the tasks of identifying the significant memberships, gaining access to those memberships, and developing a deep understanding of the practices that are emblematic of each of those memberships. Ethnographers, particularly academic ethnographers constrained by time, typically have to settle for a less-than-perfect accomplishment.

One measure of success of the observation, however, is the complexity of participation that can be achieved within the group memberships by the ethnographer. Observation and participation function as a duality in full-fledged ethnography. Observation is the foundation of competent participation, and participation opens the field of understanding that allows competent observation.

In the range of the ethnographic form, ethnographic reconnaissance is short on participation as this form has the analyst in the field for a few days or even hours. Consequently, its observations are much closer to the one-way mirror. At the other end, auto-ethnography (sometimes disparagingly called reminiscent ethnography, attic ethnography, or even narcissistic ethnography) is long on participation but suspect in any claim of systematic observation.[5]

---

[5]The emancipatory impulse of critical ethnography speaks for the group rather than about the group. Performance ethnography is an alternative form of participant observation.

Participant-observation ethnography alternates sessions of field participation with the writing of observations to create matched sets of participation and observation. Chapter 15 takes up that relationship is some detail. Here, it is enough to say that, once past that initial phase of hanging around, participation happens at the site when the ethnographer has achieved sufficient competence to actually do something that is both meaningful and instrumental to the membership, but observation happens in the writing of the field notes back at the office or its equivalent. This observation is obviously not photographic; it is a written analysis of experience. Participation takes its first step toward public knowledge in this systematic exploration of one's experience through the process of writing.

The set of field notes produced in an ethnography is a text, and it needs to be approached within the demands of good textual analysis. That text would be augmented by the artifacts, photographs, documents, and the like collected at the site, and these too are texts that must be analyzed. That analysis generally starts with the facts of the text.

Facts of the Text

Studies based on hermeneutic empiricism create a strong foundation for their contribution by starting with a systematic inventory of the facts of the text. The facts of the text are the properties of the text that might be the object of content analysis. These are called facts because they generally are not in dispute. A critic cannot argue a case against the existence of a researcher's field note or whether there are DVD chapters, scenes, or shots in a visual narrative or sentences (paragraphs, sections, whatever the unit) or this word and not that word in a written text. Those are the facts of the case. The first level of evidence resides within them. Whether in ethnography or textual analysis, there has to be factual evidence of the resources available in the action or texts for the interpretation proposed.

An Intimate Familiarity

As the researcher is the instrument, we expect the analyst to gain an intimate familiarity with those facts. In ethnography, this expectation might appear to be met in an extended schedule of multiple engagements in the action and the timely, dedicated production of field notes. But in a lengthy ethnographic effort, that familiarity fades, and the field notes and collected texts become the text of engagement, carrying the additional burden of textual analysis.

In textual analysis for whatever is the text, the expectation of intimate familiarity is met in the process called close reading. For most scholars, that means multiple "readings" of the whole text designed to get a deep understanding of what it achieves—its coherences, references, discontinuities, glosses, and the like. We have not systematized what counts as evidence of this effort as of this writing. Mostly, as we will see in Chapter 13, we just declare it as accomplished. I would argue for a deliberately systematic and publicly accessible approach in which the proximal insights gained in the close reading are archived in some written form. That written form is the second level of evidence.

Coding

For many literary scholars, the analysis ends with the close reading, and the understanding achieved is presented as the critical analysis. I would argue for a firmer empirical foundation, and that involves systematic, perhaps computer-assisted, but certainly grounded coding. Grounded coding is the process of working through a text, unit by unit, and attaching interpretations as to the topic, significance, rhetorical force, cultural location, and the like of each unit. Not only does coding rub one's nose in the text so to speak, but the codes (and the coding software) allow the texts to be manipulated and investigated in ways nearly impossible through the simple strength of one's memory. Coding is powerful evidence of effective engagement.

Warrants

The final step is to reveal the warrants. A warrant is a justification for some action. When a judge issues a search warrant, it means that there is legal justification for the search. A logical warrant means there is a basis for constructing a valid argument. In argumentation, a warrant means there is some foundation in common understanding (those "of course it's true" assumptions) for connecting evidence to claim. In interpretive research, a warrant is the basis for constructing a competent text.

Like a variable, a warrant exists in the evidence of its existence (where there's smoke, there's fire). For example, consider a film in which a White male, a White female, and a Black male are all thrashing about in the water with both rescue ship and shark in sight. What happens? Our White male and Black male are rescuing children from invading Japanese. Who dies?

We start with the facts of the case: There are a fair number of U.S. films spanning decades of time in which a White character and a character of color are in mortal danger. The situation is resolved with the death of one and the survival of the other. And it is mostly true in these cinematic choices that the White character survives and the person of color dies. (In fact, in at least one film, the Black character makes a joke about his impending death.) Exceptions are noted as exceptional.

If we presume that these narratives are competent—that they are believable—what is the cultural warrant that justifies our belief in them? In current cultural theory, we would probably frame that warrant in terms of the construction of Whiteness. In that theory, the warrant might be called the White privilege of survival. Films that resolve the threat of danger by killing off the Black character depend on this warrant for the emotional satisfaction of the solution, affirm the right of this privilege, and instruct us all in the way that it works.[6]

Once we arrive at this understanding, we might want to investigate how the warrant works in other circumstances where the outcomes seem otherwise equally likely for characters of color. We might investigate scenes of competition, romance, and the fates. We might want to see if the principle is apparent naïvely, ironically, or oppositionally in films produced by and for persons of color. Note that in the examination of warrants we have created a program of research, which will now reach well beyond our original interest.

---

[6]The insight of "depends, affirms, and instructs" comes from my colleague, Professor Ronald B. Scott.

Interpretive Trustworthiness: Coherence, Resonance, and Vraisemblance

Coherence, resonance, and vraisemblance are the interpretive equivalents to metric's reliability and validity. Interpretive empiricism is not based on any set of assumptions that places the meaningfulness of human action within the realm of the material. Human action is firmly semiotic and socially constructed. Any suggestion of reliability or validity as external to the narrative would be sadly in error. But there are standards of acceptability and of good work.

Coherence requires that the narrative hangs together and has a recognizable arc of performance. This concept might correspond to the idea of internal validity. Resonance means that the narrative has to make sense within its context and would be recognizable within the domain of the performance that it describes. In some texts, this requirement leads to a recommendation that the narrative be "checked by the members of the community of members." That position requires beliefs in the uncomplicated representation and referentiality of language and discourse. These are modernist beliefs that would pose difficulties for a postmodernist. Nonetheless, resonance might be considered as the correspondence to external validity—the narrative is more than just itself. Vraisemblance is the recognition of the narrative as plausible, without fatal defect. The reader has a sense of what it means to be a member, and enactment is considered possible. Vraisemblance might correspond to ecological validity.

There are no conventionalized measures of these standards. They are particular to the context of the study. And conventionalization is an unlikely event at any rate, given the antiestablishment impulses of interpretation. That absence does not mean that we cannot ask for the evidence of coherence, resonance, and vraisemblance as readers of the work or that we do not have an obligation to provide that evidence as analysts. We can ask and have those responsibilities in all cases.

MOVING ON

This chapter is the first of this "toolbox" section that is intended to provide you with the activities that are found in all research so that we can carry on our conversation in more specialized topics without too many disconnections. In this chapter, we have been discussing the practical foundations of these common activities. Each of us comes to this practice out of some axiomatic framework that provides the warrants for our actions. Textbooks (like all narratives) depend on those warrants, affirm their necessity, and instruct the reader on their right practice. This textbook is mediated communication just like that television show.[7] It speaks from inside the epistemological edifice that these warrants can construct. It is not really a house of cards, but it is one of human endeavor that has to be subject to critical evaluation. The next chapter takes up the specifics of research questions and hypotheses and starts our investigation of the literature.

---

[7]Of course, my intention is not to entertain (nailed it!) or to sell you something else.

REFLECTIONS

What Are Some Points to Remember?

- All research and scholarship is based on some assumptions about the properties, processes, consequences, and character of what we study. In media research, those focus on the media themselves and their texts and audiences. The key is not the suppression of those assumptions but the recognition of their presence, influence, and consequences.
- Just as we have assumptions about the world, so too we have affinities for particular theoretical approaches. A belief in a theory entails the likelihood of a methodology. Theory determines what counts as evidence. It is not by chance that cognitive theorists gravitate toward measurement scales and social action theorists move toward ethnography.
- A methodology is a means of producing evidence. The form of that evidence depends on the theory in use.
- Different methodologies have different criteria for judging the quality of that evidence. Importing criteria from one methodology to another is a political action, not an epistemological one.

Why Does It Matter?

What the analyst does and how well it is done establishes the quality of the contribution that can be made. The better we understand what we are doing and why we are doing it, the better we can achieve the quality we seek.

What Else Could We Talk About?

Media scholars have a conflicted relationship with media. Media researchers and scholars reflect elitist and class-based attitudes toward content and audiences and often hold industries as vast conspiracies. Media are seen as an overlay or a distraction from a true culture or an authentic society. Media are rarely seen as a force for good in society or even as a neutral element in our cultural environment. Scholars focus on one kind of content and attribute high levels of effectiveness to it. Many see audience members as victims or unwitting dupes. Our research mirrors those beliefs.

What Else Might Be Interesting to Read?

Jensen, K. B., & Rosengren, K. E. (1990). Five traditions in search of the audience. *European Journal of Communication, 5,* 411–420.

Mastro, D., Lapinski, M. K., Kopacz, M. A., & Behm-Morawitz, E. (2009). The influence of exposure to depictions of race and crime in TV news on viewer's social judgments. *Journal of Broadcasting & Electronic Media, 53,* 615–635.