Data Coding Workshop

RTVF 180 Fall 2010
San Jose State U
Dr. T.M. Coopman

This Workshop

This is the online component of a research skill workshop. Combined with the in class component, this workshop provides an overview and information references that supply critical knowledge to help you succeed in this course.

This workshop covers how to code data for your research projects for this course.

Coding

"The core purpose of coding is to mark the units of a text as they meaningfully relate to categories." - Lindloff

WHAT!?

Broadly, coding data is simply a way of organizing what you have observed/found in a way that helps you to answer your research question.

Lindloff:

Categorization refers to the process of characterizing the meaning of a unit of data with respect to certain generic properties. That is, to characterize a certain unit of data as belonging to, representing or being an example of some more general phenomenon.

Category is a covering term for an array of general phenomena such as concepts, constructs, themes, and other types of "bins" in which to put items that are similar.

Codes are linkages between the data and the categories posited by the researcher. Codes are shorthand devices to label, separate, compile, and organize data.

A well-designed coding scheme is a **key** element in any research. It is how you organize you data so you know what it is, where it is from, identify patterns, and (hopefully) what they mean.

Coding

There are several types of coding:

Open coding is the initial unrestricted coding of data. Chunks of text are marked that suggest a category. This opens the inquiry.

In vivo coding are the terms used by the social actors themselves. So if the a group uses the term "facilitator" instead of "leader" this choice has meaning to the actors and should be term used.

Axial coding is using codes that make connections between categories and result in the creation of new categories or a theme that spans categories. Axial coding provides context, the action or strategies in which it is carried out, and the consequences.

Integration of coding schemes changes the nature of categories from collections of incidents into theoretical constructs. This is where your data comes together to tell a story that will answer your question(s).

Do not worry too much about these terms. As a practical manner, these are things you need to be thinking of when you construct your coding scheme and start to code.

Where to Start

Coding schemes emerge from several different influences.

1. The literature around your topic area represents a longterm conversation on the topic and will offer some consensus on terms, definitions, and coding. Always start with the lit.

2. Your own experience will provide potential coding information.

3. The act of gathering data will suggest coding categories and items to code for.

1. The Lit

1.The literature around your topic area represents a longterm conversation on the topic and will offer some consensus on terms, definitions, and coding. Always start with the lit.

The most common error made by beginning researchers is to get ahead of the literature - that is, to start trying to create coding categories **before** reading up on the topic area. This wastes your time and causes extra work.

You really cannot even begin to think about beginning your research design process until you have reviewed the related literature (journal articles).

The lit will allow you to enter this research conversation and take advantage of a vast amount of knowledge and experience.

1. Lit

The lit will give you insight into what to look for and how to describe it.

Consider the example from the last workshop: *"What is the impact of iTunes and similar online services on the television industry?"*

You first have to consider what types of impacts you want to looking at - economic, content, audiences?

1. Lit

It is often useful for beginning researchers to copy or adapt the research method and coding scheme from an existing study (ok if you cite it and use your **own** data).

The advantages: tested methods created by professionals that allow you to focus on other elements of your study.

The disadvantages: it may not be a good fit with your data and you may miss something novel.

2. Personal Experience

2. Your own experience will provide potential coding information.

Since researchers often research scenes that they are familiar with, your insider knowledge may give you insights into elements that other researchers without that knowledge may have missed.

For example, as a person who uses iTunes or other services to watch TV and a RTVF student you will have special knowledge as both a user and a producer of media.

2. Personal Experience

Therefore, you may not only aware of the viewing habits of friends and family, you may also know the technical and creative demands of producing media for different formats.

3. Emergent coding

3. The act of gathering data will suggest coding categories and items to code for.

Often, you are not aware of potential behaviors or coding categories until you get into the middle of actually collecting data.

This is why it is a good idea to "pilot" your study by doing some preliminary observations before starting a more thorough examination.

For example, you may talk with friends and family and discover that they do use traditional TV in specific (or even interesting) situations.

Coding Categories

Consider the earlier example: *"What is the impact of iTunes and similar online services on the television industry?"*

Let's say you decide that you are more interested in how people use media. So you tweak it to: *"What is the impact of iTunes and similar online services on the television viewing habits of students?"*

First you need to decide how you are going to answer this question. Therefore you need to break it down in a way you can record and analyze it.

Coding Categories

For example, based on the lit, your own experiences, and some early observations you decide you want to do some short interviews and have students fill out a survey.

What do you want to know? How they use media, when, and why?

You may want to know the program watched, the physical environment, if it was live or recorded, and why the choice was made. Was it convenience? You even need to decide what qualifies as a TV program. Just serial comedies or dramas? Does a movie on TV count? What about DVDs? Or watching on a iPod or iPhone?

Coding

Coding all data. especially qualitative data, can be subjective. In all research, categories and schemes could be changed, and thus potentially providing different results.

The goal is to select and create a classification system that is internally and externally consistent and is firmly based in the lit and your own experiences and observations.

A further challenge is when two or people are coding the same material.

Coding

Team Coding (coder reliability)

If you work with others, you need to be sure you are coding consistently. This means that if you both observe the same behavior, you code it the same way.

The best method is to collaboratively create the coding scheme and then code some data together. Check for inconsistencies and discuss why you made different choices.

Alter your process until you can code consistently.

Coding

Let's concentrate on what motivates someone to watch TV on a particular device. You want your report to be interesting and compelling, so description and context are important.

We could call this particular coding category anything that makes sense, but lets keep it simple and call "Viewer Motivation." In order to ensure you code for the same thing every time, you need to define each coding category. Viewer Motivation is defined as "the reasoning or rationale for choosing to view specific programs on a particular device."

Note that individuals in your study need identifiers (real or assigned names). Each coding item would have its own section, this is simply one such section. So, Viewer Motivation could go into your coding sheet like this:

Individual	Johnny B.	Alice K.
Viewer Motivation	"I like Hulu because a lot of it is free and I can watch it anytime. I don't want to pay for cable and I want to be able to watch it whenever."	"Watching when something comes on the first time is the best. I like taking the time off to sit and watch it, I even like the commercials because they can be funny."
Device Choice	Laptop Has no TV set	Cable TV for most regular programs Has DVDs of TV series that she watches on laptop.

Note that in this case your data is organized by individual. We also added a category for their devices choices. Setting things up like this allows you to organize and get a clear view of your data.

Exercise

The best way to understand coding is to try doing it. I have an exercise for you to work through. Download the coding worksheets on the workshop page and follow along.

This coding example is from my Ph.D. dissertation. It happens to be coding a text. But the same principles apply to other types of coding.

Please note that you DO NOT code journal articles for your research projects. Articles are not data (unless you are studying them).

Exercise

This is coding from one section of my dissertation "Dissentworks: Emergent Collective Action at the Turn of the Millennium (2008)." The broad categories I used emerged from my reading of the websites I visited and the Mission Statements of the Independent Media Center (IMC), also known as Indymedia, collectives that operated the websites. I use the **Grounded Theory** approach to coding where the coding is open and emerges from the data.

I categorized emerging measures using four broad categories I labeled: format/position, struggle, identity, and structure and production.

The format/position examined how statements were labeled, their uniqueness to the particular IMC, use of the network's defining philosophy "don't hate the media, be the media" in some form, and degree to which the IMC defined its focus to a specific geographic area.

Struggle addressed actual mobilization or protest activity, statements invoking the struggle against globalization/capitalism and dominant media, and declarations of solidarity with traditionally oppressed groups.

Identity addressed the IMC's position in relation to its context, identifying the degree to which the mission statement invoked the network or its history, the local IMC's history, local or regional history where the IMC was located, and social/resistance movements.

Structure and production examined creating media, open publishing/journalism, the website's structure and the collective's editorial policies, media training, and organizational policies and themes such as use of consensus or volunteer run.

Exercise

The categories of format/position, struggle, identity, and structure and production are broad and are designed to help we with my final analysis.

Once I got into actually collecting data, I refined my categories based on specific examples I discovered. Realize that you may code for a specific item that you ultimately do not wind-up using or is not helpful in your final analysis. Re-collecting data is a pain. So it is better to collect too much than too little while you are at it.

I organized my coding like this:

Format: this is mostly technical information on each mission statement.
Word count: how long in words. (format)
How it is labeled: as "mission statement" or "about us." (format)
Is it an original statement, a commonly used format or "boiler plate" or a variation on the network's statement. (format)
Is it a mission statement or closer to the "Principles of Unity" (another founding document. (format)

Themes: these are themes that emerged from reading the data.
Verbatim excerpts from the Indymedia mission statement
"Don't hate the media, be the media" (all or part of the networks slogan)
Struggle against Globalization/capitalism (general discussion)
Struggle against Dominant Media (specific mentioning media or media systems)
Invokes the Indymedia network (position in, solidarity with)
Invokes oppressed groups (specific or general – voiceless, GLBT, etc.)
Invokes Org issues (consensus decision making, non-hierarchical, horizontal, volunteer, non-commercial, etc.)

Events/mobilization: these are actual events that occurred, general use of protests, and other references.
Protests/mobilization (use of, being a part of)
Invokes network history (WTO, Seattle, Genoa protests)
Invokes local history/story of IMC (how it was founded etc.)
Invokes local/regional history (general - puts the local IMC in the context of its community)
Invokes locality/service area (community it serves)
Invokes social/resistance movements (general or specific - these are other movements that promote causes, such as the environmental movement or the anti-capitalist movement)

Media Production: the actual practice of empowering people to create media
Making media/open publishing/journalism (discuss philosophies and practices)
Website structure/policies (how people can participate, what is allowed/forbidden
Media Training: training people in media creation and journalism.

Exercise

This is an actual coding sheet with the different items abbreviated. Note that there is a space for whether or not the item is present and the important description and/or notes. This sheet is set up as a **key**, with the definitions of each category in the notes section. This is useful to ensure consistency.

IMC	Y	N	Notes
Be the media			"Don't hate the media, be the media" (all or part of the networks slogan)
Globalization/ Capitalism			Struggle against Globalization/capitalism (general discussion)
Dominant media			Struggle against Dominant Media (specific mentioning media or media systems)
Invokes network			Invokes the Indymedia network (position in, solidarity with)
Oppressed			Invokes oppressed groups (specific or general – voiceless, GLBT, etc.)
Org Issues			Invokes Org issues (consensus decision making, non-hierarchical, horizontal, volunteer, non-commercial, etc.)
protest/mobilization			Protests/mobilization (use of, being a part of)
network history			Invokes network history (WTO, Seattle, Genoa protests)
Local/IMC history			Invokes local history/story of IMC (how it was founded etc.)
Local/regional history			Invokes local/regional history (general - puts the local IMC in the context of its community)
Locality/service area			Invokes the community it serves by name
Movements			Invokes social/resistance movements (general or specific - these are other movements that promote causes, such as the environmental movement or the anti-capitalist movement)
Making Media			Making media/open publishing/journalism (discuss philosophies and practices)
Site structures/ policies			Website structure/policies (how people can participate, what is allowed/forbidden
Training			Media Training: training people in media creation and journalism.

Here is an example of how a coding sheet is used to code a text.

IMC	Y	N	Barcelona <http://barcelona.indymedia.org/>
Be the media	x		"become the media"
Globalization/ Capitalism		x	
Dominant media	x		corporate media's distortions and unwillingness to cover efforts to develop an egalitarian and sustainable society. We are dedicated to addressing issues that the mainstream media neglects and we do not conceal our politics behind a false objectivity.
Invokes network	x		The Independent Media Center is a global network of collectively run media outlets for the creation of radical, accurate, and passionate tellings of truth.
Oppressed		x	
Org Issues		x	
protest/ mobilization		x	
network history		x	
Local/IMC history		x	
Local/regional history		x	
Locality/ service area		x	in the barcelona Area
Movements	x		promoting social and economic justice
Making Media	x		media production and distribution as tools
Site structures/ policies		x	
Training		x	

Barcelona <http:// barcelona.indymedia.org/>

Drafted 7/25/01

[**The Independent Media Center is a global network of collectively run media outlets for the creation of radical, accurate, and passionate tellings of truth.]** We are motivated by a love and inspiration for people who work for a better world despite [**corporate media's distortions and unwillingness to cover efforts to develop an egalitarian and sustainable society.]**

The barcelona Independent Media Center is a grassroots organization committed to using [**media production and distribution as tools]** [**for promoting social and economic justice]** [**in the barcelona Area]** [**We are dedicated to addressing issues that the mainstream media neglects and we do not conceal our politics behind a false objectivity.]** We will empower people to [**"become the media**"] by providing democratic access to available technologies and information.

IMC	Y	N	Barcelona <http://barcelona.indymedia.org/>
Be the media	x		"become the media"
Globalization/ Capitalism		x	
Dominant media	x		corporate media's distortions and unwillingness to cover efforts to develop an egalitarian and sustainable society. We are dedicated to addressing issues that the mainstream media neglects and we do not conceal our politics behind a false objectivity.
Invokes network	x		The Independent Media Center is a global network of collectively run media outlets for the creation of radical, accurate, and passionate tellings of truth.
Oppressed		x	
Org Issues		x	
protest/ mobilization		x	
network history		x	
Local/IMC history		x	
Local/regional history		x	
Locality/ service area		x	in the barcelona Area
Movements	x		promoting social and economic justice
Making Media	x		media production and distribution as tools
Site structures/ policies		x	
Training		x	

Let's unpack some of this.

Barcelona <http://barcelona.indymedia.org/>

Drafted 7/25/01

[**The Independent Media Center is a global network of collectively run media outlets for the creation of radical, accurate, and passionate tellings of truth.**] We are motivated by a love and inspiration for people who work for a better world despite [**corporate media's distortions and unwillingness to cover efforts to develop an egalitarian and sustainable society.**]

The barcelona Independent Media Center is a grassroots organization committed to using [**media production and distribution as tools**] [**for promoting social and economic justice**] [**in the barcelona Area**]. [**We are dedicated to addressing issues that the mainstream media neglects and we do not conceal our politics behind a false objectivity.**] We will empower people to [**"become the media**"] by providing democratic access to available technologies and information.

Movements: Invokes social/resistance movements (general or specific - these are other movements that promote causes, such as the environmental movement or the anti-capitalist movement).

Based on my knowledge, I have decided that words like "social and economic justice" are broad terms that reflect the general goals of most progressive social movements.

Globalization/Capitalism: Struggle against Globalization/capitalism (general discussion).

Based on my knowledge, I have decided that just mentioning corporations does not invoke struggle against the broader issues of globalization/capitalism. If, for example, the words "transnational" or "global" were used it would then also fit in this category.

Exercise

Now try some coding.

1. On the downloaded word document you have both a coding form and a text to code. This is identical to the form used here, but the example is different.

2. Look at the definitions of the terms on the earlier slides and try and decide what parts of the text go in which category and item.

3. Because of the nature of compound sentences, there may be some overlap between items. Excessive overlap is a sign that categories/items need to be combined or refined.

4. It is also useful to write your own rationale for selecting a category/item if it is unclear.

5. At the end, review your choices. Do they make sense? Could you have chosen differently?

There are no "correct" answers - the point is to try and figure it out.

Coding

Coding is simple in its concept but difficult in its execution.

Coding systems are refined over time.

Do not get overwhelmed by the idea! This is challenging. The point is to learn about the method by going through the process.

You will make mistakes, but this is how you learn.

Tasks

To earn credit you need to post a substancive 100+ word comment on this workshop's wiki page and complete the following tasks and bring them to class.

Since you will be conducting a participant observation in a class at SJSU, think about how you might go about coding your study.

1. Complete the assignment as described in the workshop.

2. Bring it to class to compare with other students,

3. Work with other students in your team to create a coding sheet you all agree on.