This Workshop

This is a research skill workshop. This workshop provides an overview and information references that supply critical knowledge to help you succeed in this course.

This workshop covers coding in more detail including how to identify useful data and combine data from different methods. The exercise in this workshop is combined with the method workshops to provide an opportunity to code data you have collected.
Review

First, let’s review some of the basics from the first coding workshop.

Remember, categories and coding change over time based on the data you collect. Some categories and codes need to be expanded, while others might be consolidated or eliminated.

Coding is not just some random exercise, but must have utility in helping you organize and analyze your data.

Coding saves time because it allows you to get at rich informative data with having to go back to your raw data and it allows you to see patterns across your methods to help judge reliability.
Triangulation

You will collect and code a lot of data for your project.

The challenge is what data is important or meaningful?

We tend to get swept up in the weird or different, but the focus needs to be on what data is confirmed across multiple methods.

This is called triangulation.

Data that is not confirmed in at least two methods has low reliability and should set aside.

In order to speak confidently about your findings, those findings need to have high validity and reliability.

This is achieved when the same data appear in at least two of three methods and perhaps in the lit as well. This data is the sweet spot that guides your analysis.
Reliability and Validity

Your goal is reliable and valid results. Broadly and for the purposes of this project Reliability is that the data your analysis relies on is confirmed across two or more methods and perhaps in previous research (the lit). Internal Validity is based on your design and its ability to address your question based on the social scientific process laid out in this course. External Validity is your examination of your results and addressing whether or not their might be other explanations.
Organizing your Process

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 more people are coding the same material.

Team Coding (coder inter-reliability)

If you work with others, you need to be sure you are coding consistently. This means that if you all 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.
Definitions

There are a wide variety of ways you can organize, define, and code your data. It all depends on what you are trying to discover.

For example, some of my research looked at activist group’s mission statements and within that I wanted to know how they discussed (and how often) protest (mobilization) versus organizing (meetings, everyday work). These are two similar, yet importantly different, activities.

It is important to both define terms concretely and ensure terms do not overlap and to keep going back to your definitions to make sure you do not confuse them during the coding process.
Codes & Categories

Coding categories should be broad areas of behaviors under which specific behaviors are arranged. For example, if your research question is "How do people determine what modes of communication are appropriate for a given situation." That is, why do people decide to text versus a voice call or a face-to-face conversation.

DO NOT use questions as coding categories. Answers may fit into several different categories and codes because real people speak in complex ways.

After reading up on the literature and piloting your methods you decide one category of behavior falls under convenience defined as "the least mount of effort to accomplish a communication goal."

What types of specific communication behaviors (what people actually do) that would fall under "convenience?" After reviewing what people actually told you, several things emerge.

People want to be able to control interaction. This falls into two main areas. The first is the amount of time for the communication act. The second is the depth or the scope of the interaction. That is, keeping it to a single subject or limit the complexity of discussion on a single subject.
Putting Data in Codes

Caveat: Keep in mind that any coding scheme needs to make sense to you and have utility in that it helps you see patterns and analyze your data. This workshop example has its limits because it is hypothetical and you do not have the background and context you would normally have after reviewing the lit and working and thinking about the topic.

Specific data goes underneath specific codes. Let’s look at pieces of data from different method and how they might fit into the same code.

Convenience/interaction/time

Interview: Tom stated "If all I need to say is 'meet us at 4th Street at 7' I want to hit it and be done and not get bogged down in conversation. That is why I text stuff like that."

Focus Group: Jenny stated "If something is complex and confusing it is faster to just call and talk to someone instead exchanging a bunch of texts."

Survey: 62% of respondents listed convenience as the primary criteria for deciding what mode to communicate in.

For the category of convenience you have two codes, Interaction/time and Interaction/depth. If you think there may be more items/behaviors under convenience you could make interaction a sub-category and time and depth codes beneath it.

This is an example of how your coding scheme may change based on the types and amount of data you collect. The idea is that you want a system that is complex enough that you can see patterns in the data but not so complex that gets too busy and hard to work with or understand.

So, do not be afraid to alter your scheme based on the impact of your data - this approach is known as "grounded theory."

Codes should be the same for all data regardless of method.
Making Sense of Data

Now you have three different bits of data, two great quotes, a stat, and a code that connects them. Keep in mind there would need to be a lot more data to support any contention that would approach reliability.

When writing up your data you need use examples and show how data is consistent across methods and therefore reliable. You also need to bring in your lit to provide more context. Here is how a paragraph might look that combines all the different data, lit, and analysis.

Deciding the best way to contact peers is a complex process that includes several elements. For example, students place a premium on convenience when communicating with peers. A majority (62%) of those surveyed cited convenience as the primary criteria for deciding the communication mode to use. This supports Ulrich’s (2013) findings on student mobile device use in the UK. However, participants interpreted convenience in a variety of ways. For Tom, it was limiting the task to a minimum time to achieve his goal. He stated, "If all I need to say is ‘meet us at 4th Street at 7’ I want to hit it and be done and not get bogged down in conversation. That is why I text stuff like that." Where, as Jenny described, "If something is complex and confusing it is faster to just call and talk to someone instead exchanging a bunch of texts." For her, speed was about clarity or efficiency within the interaction as opposed to Tom’s strategy in limiting/eliminating interaction. His message was simple, her’s complex. So, while people value convenience they enact it in different ways depending on the best mode to meet their complex communication goals around the interaction itself and the effort taken to complete it.
Deciding Relevancy

These examples are not the only way or even the best way to code this material.

Not every survey, focus group, or interview will yield data for every category or code. Some codes may wind up with little or no data, other codes may have to be split up because there is too much or too diverse data.

Some data will lack utility. That is, will not help you answer your RQ or help you discuss your topic. Research can be like panning for gold, you have to sift a lot of material to get a few good nuggets.

Remember, the goal is to organize your data in such a way that you can see patterns and shed some light on human communication processes. In the end, you use your data to tell a story of what happened and why. The utility of your data depends on how it helps to tell that story.