Qualitative Research Methods
Assistant Prof. Aradhna Malik
Vinod Gupta School of Management
Indian Institute of Technology - Kharagpur

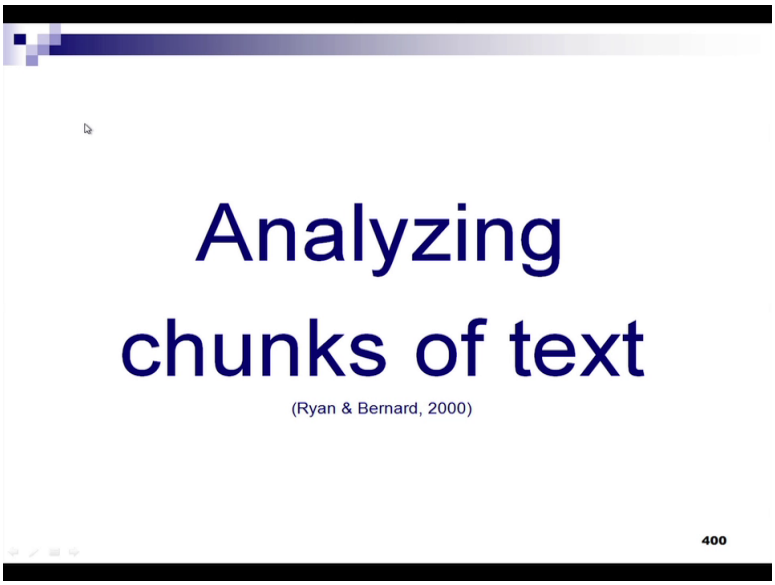
Lecture 37
Data Management and Analysis Methods (Contd.)

Welcome back to the NOC course on qualitative research methods my name is Aradhna Malik and I am helping you with this course, we have discussed various aspects of qualitative research, we have talked about what qualitative researcher is, we have talked what strategies of inquiry we talked about, we started talking about data collection and analysis methods and we were specifically discussing the data collection and analysis methods for textual data.

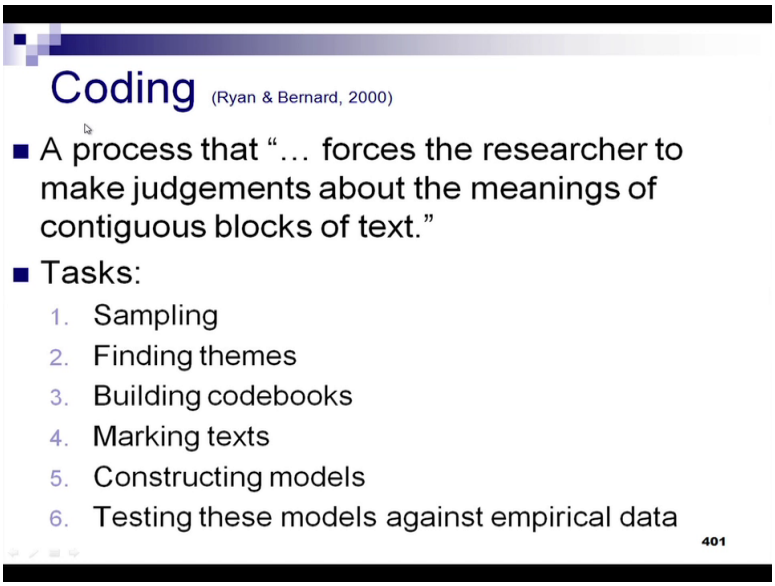
So last time we talked about the data for words, you know analyzing words and today we are going to talk about analysis of chunks of words, so we will move take the same discussion a little forward and we will discuss how we can analyze words within their own context, we will talk about paragraph, we will talk about narratives and we will talk about how these words can be analyzed within the context.

So these words together create a context and this context are represented in and through the language that we used to describe that context by way of words, that we use in that particular language, so these are chunks of word, groups of words that appear together, okay.

(Refer Slide Time: 01:35)



(Refer Slide Time: 01:37)



Now what we do here is we code the data, coding is a process that "forces the researcher to make judgments about the meaning of contiguous blocks of text." So we are forced, we read something and we are forced to see the chunks of word as building up a context, we are forced to make a judgment about what we see emerging out of the relationships of this that these words with each other and that is what coding does.

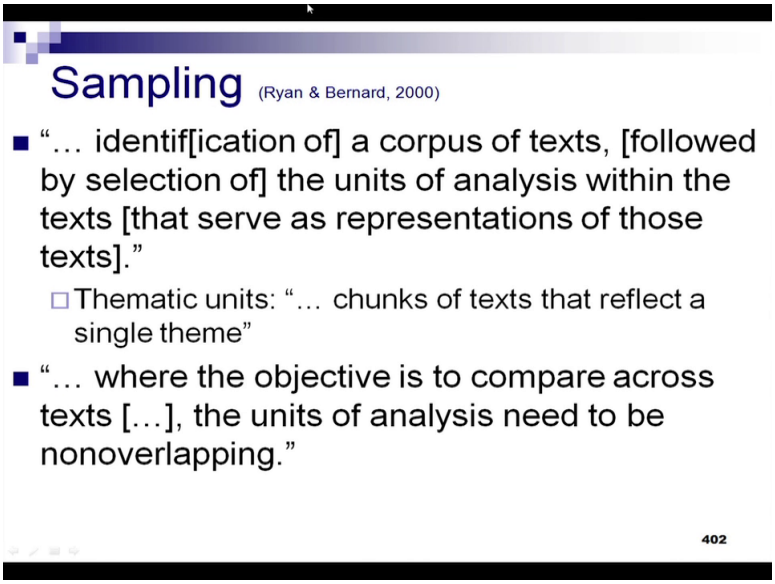
Coding asks us to focus our thoughts, to focus our judgment, to focus our interpretation of what we read as you know this collection, of collections of words, and you know come up with a single or with a very focused, very clear understanding, of what it is that we are talking about what it is that this scribes this collection of words and the tasks in this are, how do we

do coding, we discussed coding a little bit in grounded theory earlier. So now we will discuss coding in greater detail.

What happens in coding, the first step in coding is sampling, we need to find the most appropriate sample, the best representative of this whole ongoing you know volume of text that we have, so sampling is of utmost importance. The second step here is finding themes we take a sample and we find what emerges, we take these blocks of text and we try to find out what is emerging out of these texts.

The third step here is building code books, we try and understand or we make a list of what you know these words indicate in terms of code books and we had talked about code books even the previous lectures, so I am not going to go into detail about that. Then we mark the text with references from code books and we use these marks or this marked text, we use the information that we garnered from these marked text to build models, to construct models that relates to the theory that we are talking about. And then we test these models against empirical data that is the whole process that we will be discussing okay.

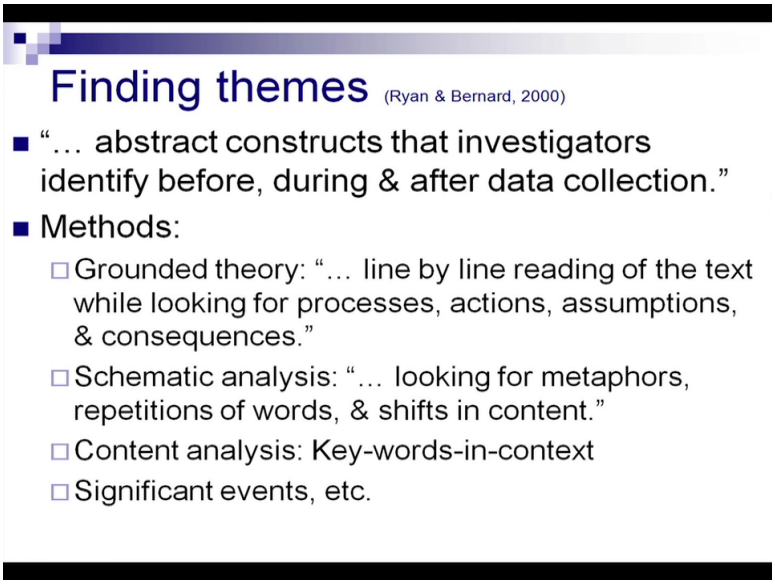
(Refer Slide Time: 04:14)



Sampling: sampling is the "identification of corpus of text followed by selection of the units of analysis within the text that serve as representation of those text." So we have a corpus, we have a large collection of texts and we select the units of analysis within those large collections of text and we take those markers of those units or we classify that portion of the text that represents the text, that we are talking about as best as possible. So we draw it from there.

Thematic units" are "chunks of texts that reflect a single theme" simple to understand "where the objective is to compare across texts, the units of analysis need to be nonoverlapping." We cannot have two units of texts that are overlapping that indicate something similar, they need to be distinctly different, so where our object is to compare something, compare a phenomenon across text, the unit of analysis needs to be completely different.

(Refer Slide Time: 05:34)



How do we find themes, themes are "abstract constructs that investigators identify before, during and after data collection." Why before? Because we needed direction for a data collection. Why during? Because we need to stay on track and we observe, we have discussed how qualitative researchers are very closely connected with environment that they study, as opposed to researchers who engage in quantitative research or you know the objective research as we say so.

We are very closely connected to the objects of our research and we understand and we acknowledge that we as researchers, our presence in the environment of the phenomenon that we are studying, is very likely to influence the phenomenon that is being studied and that in turn will affect what we get out of that study, that in turn will affect our observations, will affect how the environment behaves, as a result of us being there, even though we are not participating, but just by being in that environment, environment is affected and the environment behaves in a certain way.

Then we take, we take this themes, so these are abstract constructs that we identify before to guide a research during the data collection to make sure that we are in tune, we are connected to the to the environment or to the phenomenon that we are studying and to our goals from that study and after data collections, after we have collected at data, we try and we verify that the appropriateness of the data that we have collected in terms of the research question that we had asked.

Sometimes we get sidetracked while collecting data, we say okay, this also looks interesting, this also looks interesting, so we start collecting a whole lot of data that initially was not planned, however you know we need to stay on track, so we can check and we go back and forth, with you know we will to the research question that we have asked, we go back and forth to the direction that we have taken to the theme, that we had hoped, we would collect data on, so that is something that we do.

The second thing that is likely to happen or that does happen in such cases is that during data collection or after the data has been collected, we realize that that we could modify or we realize the possibility of modification of the initial direction, that we had taken, so it's okay we started out with this assumption, we started out with an idea in our minds.

But in the process of collection of data and after having collected this data we realize that this direction can be modified, to include the data that we have collected and come up with something much more significant okay. So we need to continuously revise, modify, update our themes, before during and after data collection.

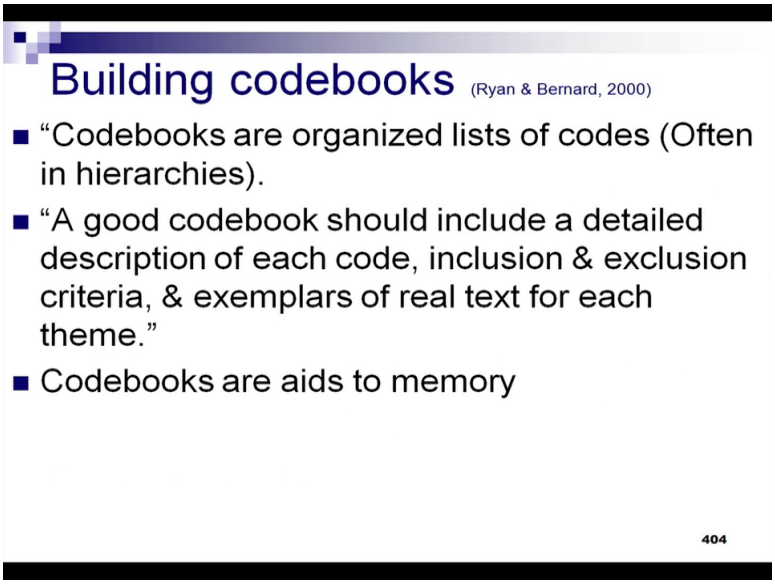
Method for these are grounded theory which is, in grounded theory the finding of themes happens through "line by line reading of the text while looking for processes, actions, assumptions and consequences." So we go line by line, we have discuss them grounded theory I'm not going to go into it. Schematic analysis we will talk about it, we "look for metaphors; we look for repetitions of words and shifts in content." We look for abstract representations of the context that we are referring to okay.

In content analysis look for key-words-in-context, we will talk about this as well and we also use significant events, we use information that we gather during significant events, as to find themes, so significant events actually act as markers for themes, so our themes can revolve

around this significant events, for example if we are trying to study the impact of global warming on availability of groundwater to farmers, that one thing.

When you studying that, maybe in some years there was more, the was higher monsoon, now that is an outcome of global warming, then the amount of rainfall crossed all you know crossed all records and or maybe in some year we are talking about, we talking about the impacts, but then in some you maybe there was the draught. So these be significant events help us find out themes help us anchor our quest for data. Okay.

(Refer Slide Time: 10:50)

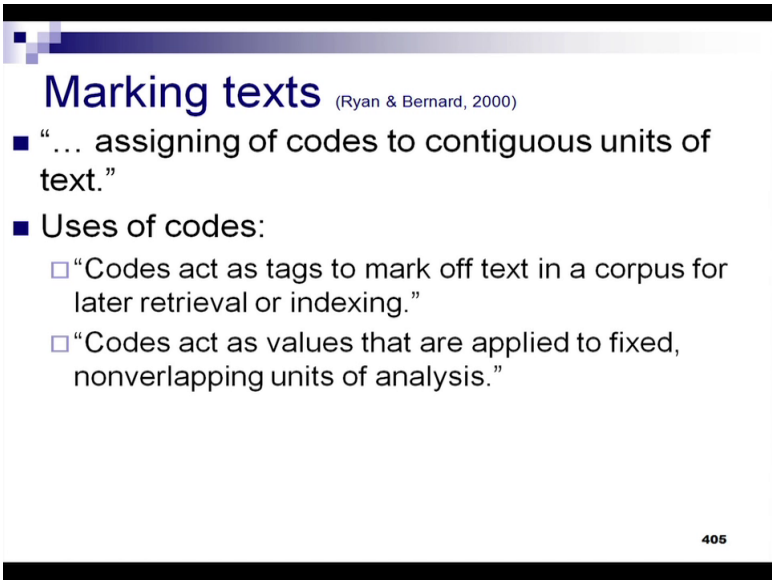


Then Building Codebooks, "Codebooks are organized list of codes which are often in hierarchies." "A good codebook should include a detailed description of each code, inclusion and exclusion, criteria and exemplars of real text for each theme." Codebooks are aids to memory. So what we do is, we build these codebooks, with what the theme is and the characteristics of this theme.

So and themes are arranged in a specific hierarchical list, what is more important? What comes next? What comes next? Etcetera. A good codebooks should include a detailed description of each code, each code is describe what it includes, what is what describes it, what defines the code, it should also includes the inclusion and exclusion criteria, so what on what basis will we say that something belongs to the code or this code has to be assigned to a particular aspect or events.

On what basis will we say that this code has to be you know this item does not belong to this code, so inclusion and exclusion criteria very important and exemplars of real text for each theme to help the reader understand, you have the interpreter understand how to use the codebook effectively, codebooks are service aids to memory. Okay.

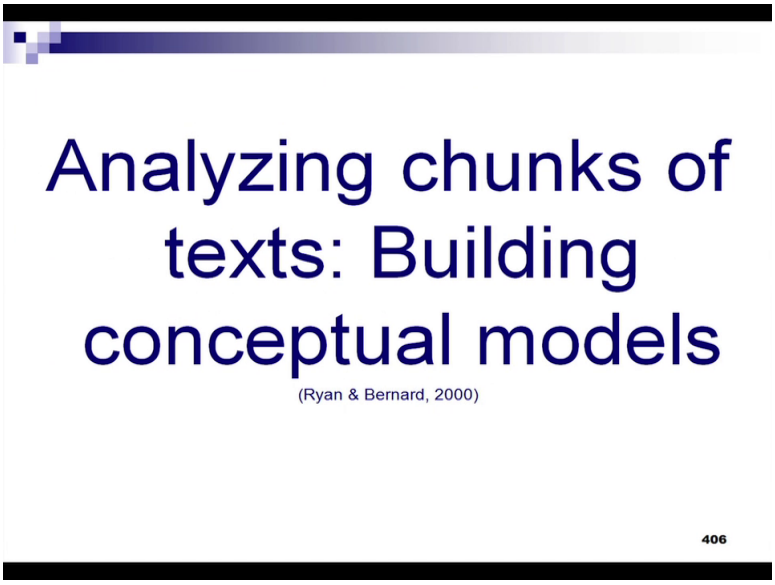
(Refer Slide Time: 12:21)



Marking texts or when we talk about marking texts we "assign ot codes to contiguous units of text." So be built our codebook when we take those codes and then we go back to the text and we try and connect whatever we have built in the code book to the text itself. The use of codes: "codes act as tags to mark of text in a corpus for a later retrieval or indexing."

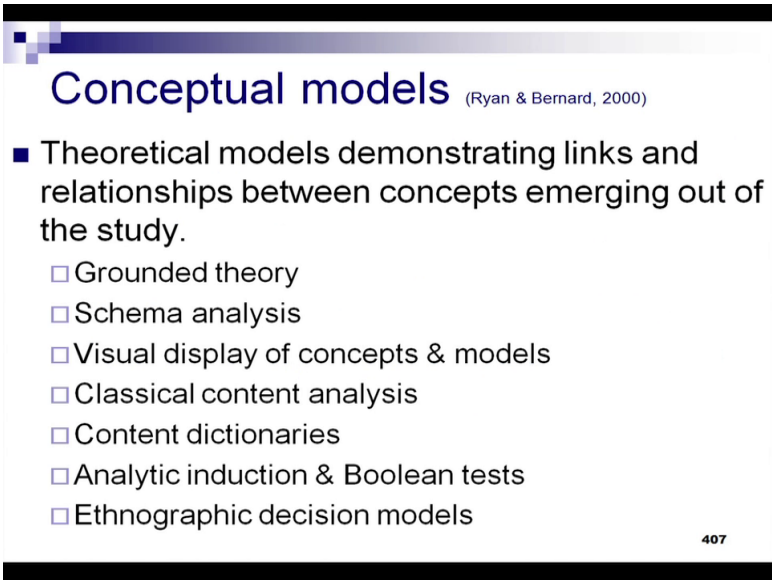
"Codes also act as values that are applied to fix nonoverlapping unit of analysis." So they act as tags and they act as values that are applied to fixed non-overlapping units of analysis, they are judgments that we used to categories and represent our data.

(Refer Slide Time: 13:19)



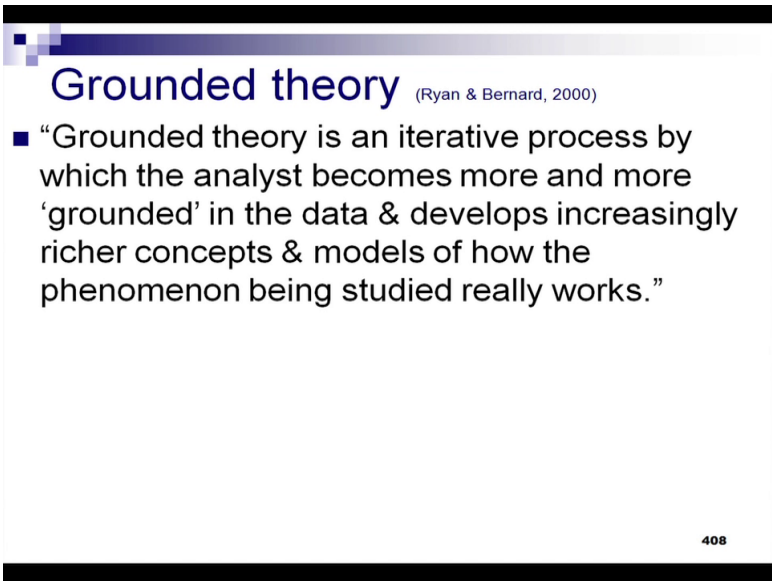
Analyzing chunks of texts, now we move on to building conceptual models, we coded our text, we know where each thing belongs, we know what themes are emerging, now the time comes to take all that we have collected and build model out of it. how do we do this.

(Refer Slide Time: 13:38)



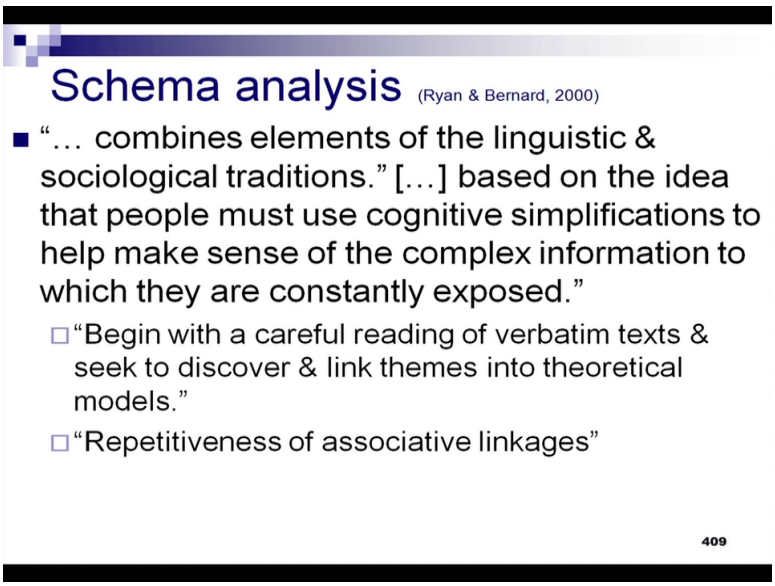
Conceptual models are theoretical models demonstrating links and relationships between concepts emerging out of the study. Various ways in which we can build is conceptual models are grounded theory; we studied this so we won't go into this. Schema analysis, visual display of concepts and models, classical content analysis, content dictionaries, analytic induction and Boolean tests and ethnographic decision model. So we will talk about it this one by one.

(Refer Slide Time: 14:08)



Grounded theory: "grounded theory is an iterative process by which the analyst becomes more and more 'grounded' in the data and develops increasingly richer concepts and models of how the phenomenon being studied really works." Okay so the analyst becomes more and more rooted, the analyst becomes more and more connected to the data, becomes more and more link to the data and develops increasingly richer concepts and models of how the phenomenon really works, schema analysis, we have talked about grounded theory, so we are not going to discuss that today.

(Refer Slide Time: 14:49)



Schema analysis, now "it combines elements of the linguistic and sociological tradition." Based on the idea that people must use cognitive simplifications to help make sense of the complex information to which they are constantly exposed." They "begin with careful

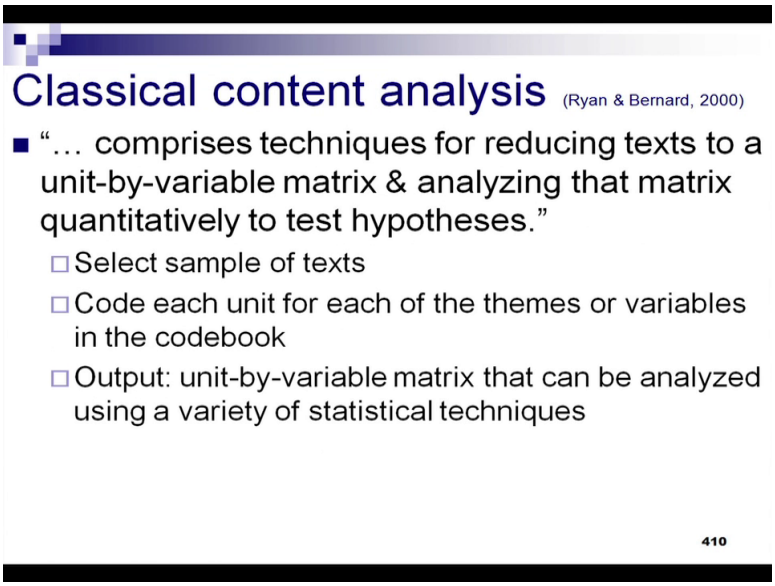
reading of verbatim texts and seek to discover and link themes into theoretical models." And they focus on the "repetitiveness of associative linkages."

Schema analysis really take see the these elements of linguistic and sociological tradition and it is based on the idea that people must be able to simplify whatever they see in their environment in to schema or schemata, schemata is a plural, schema is singular. So into different schemata that represent you know schema is a picture that forms in our head, it's our understanding, it's your understanding, it's my understanding of a particular phenomenon.

How do I categorize it based on my understanding of the phenomenon, in an through my past experiences, connected to my understanding, my interpretation, my study, my you know the way I process the data, so all of this and study creation of a schema, now it begins with careful reading of verbatim texts and seek to discover and link themes into theoretical models.

So we see all this texts, we have created all these themes and then the next question is how do these themes connect to each other, so we try and take whatever we have understood from these themes, about these themes, from the text and we try and connect it into a theoretical model. It requires or it rests on the repetitiveness of associative linkages okay, so next.

(Refer Slide Time: 16:53)

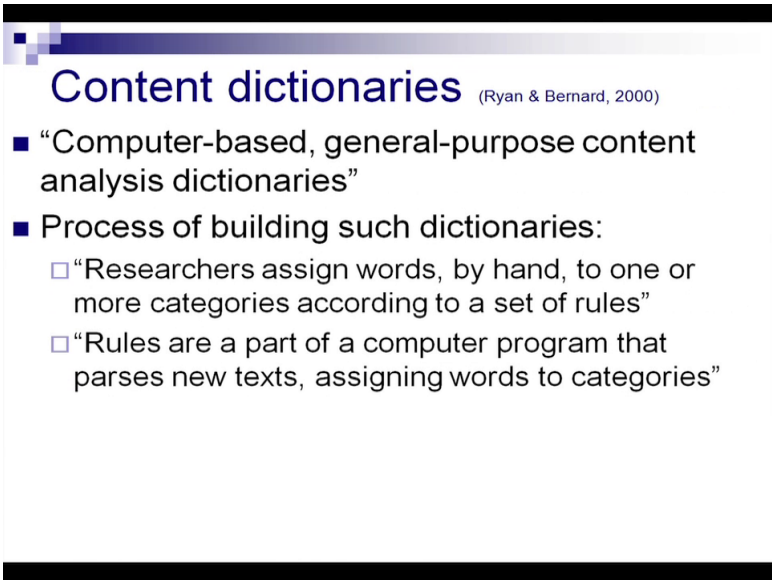


Classical content analysis, this describes this "comprises techniques for reducing text to a unit by variable math matrix and analyzing that matrix quantitatively to test hypotheses." It comprises of, it includes techniques that reduce the text to a unit by variable and we just

talked about this and then we analyze that Matrix quantitatively and then test hypotheses which are our beliefs about a particular phenomenon.

So to do this we first select the sample of texts, then we code is unit for each of the themes of variables in the codebook and the output of this is a unit by variable matrix that can be analyzed using a variety of statistical techniques. Okay.

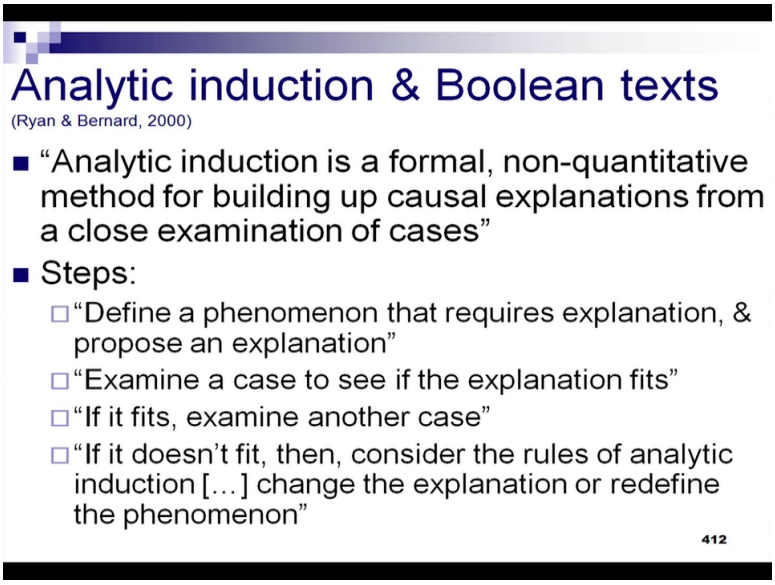
(Refer Slide Time: 17:42)



Content dictionaries are "computer based, general purpose content analysis lists or dictionaries." Okay, the process of building such dictionaries is that "researchers assign words, by hand, to one or more categories according to a set of rules", we have a set of rules, we go through the texts, we say okay, this category of words means this, this category of words, you know we take a collection of words, we form categories and name those categories.

And these names are given according to a set of rules, these "rules are a part of a computer program that parses, that parses means breaks up new texts, assigning words to categories" and we will talk more about this when we discuss software, the role of software in qualitative research sorry the role of computer technology in qualitative research.

(Refer Slide Time: 18:40)

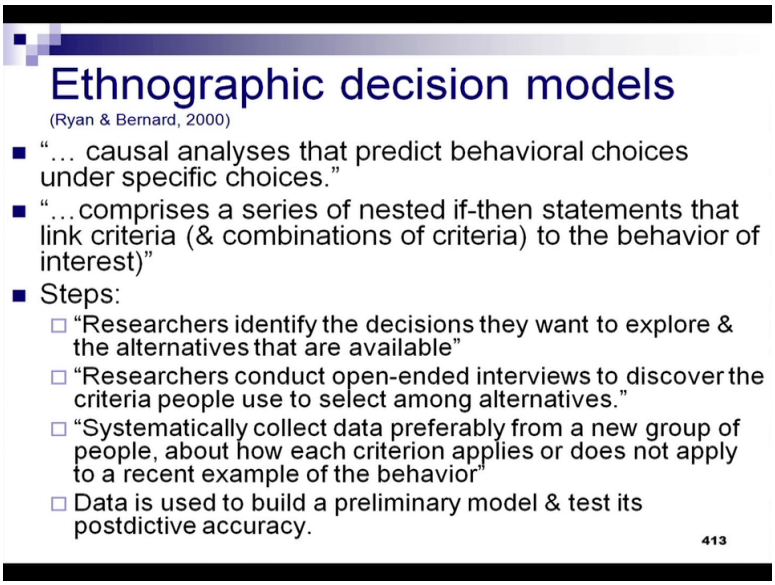


Analytic induction and Boolean texts: "analytic induction is a formal, non quantitative method for building a causal explanations from a close examination of cases" and the steps here is, so we, it is a non quantitative method but is a formal methods for building a causal explanations from a close examination of texts, if then loop is a Boolean method, so the steps in this are the first: "define a phenomenon that requires explanation and propose an explanation."

Then "examine a case to see the explanation fits", "if it fits then we examine another case", then we see whether it fits, if it fits then we examine another case, in the event that the case does not fit the definition for the explanation that we have given to the phenomenon that then we consider the rules of analytic induction and change the explanation or redefine the phenomenon to include whatever doesn't fit okay.

So we go back and forth, so the two outcomes we have are it fits or it doesn't fits, there is no middle ground that's what you know, that's where the Boolean reference comes from, its either plus or minus, one or two okay sorry zero or 1. So we it either fits or it does not fits and if it doesn't fit then we do something else for it, okay, so we change the rules of analytic induction.

(Refer Slide Time: 20:23)



The last aspect that you are going to discuss in this is or the last method of analysis of large chunks of data that we are going to talk about now is ethnographic decision models. Ethnographic decision models are "causal analyses that predict behavioral choices under specific choices."This "comprises a series of nested if-then statements that link criteria and combinations of criteria to the behavior of interest."

If X happens when Y will happen, if the sun shines, the water in lakes will dry and water dries up the marine life will die, if the marine life dies then you know then that is the way it is it happens, all right, it is a series of nested if then statements that link criteria to the behavior of interest. The steps here are number one: "researchers identify the decision they want to explore and the alternatives that are available." So they find out what they want to do, what they want to do, and the alternatives that are available in order for them to do what they want to do.

The second step is "researcher conduct open-ended interviews to discover the criteria people used to select among alternatives." So they find out what kind of criteria we need to use to select or what kind of criteria people used to select among these alternatives. The third one is this "systematically collect data preferably from a new group of people, about how each criterion applies or does not apply to a recent example of the behavior."

So there is cross checking, maybe this group of people says that it fit or it applies, but another group of people says does not really you know that they don't really see that, so that's what

they do. Data is used to build a preliminary model and test it postdictive accuracy, and that is where, you know that is how then we take the status that has been verified, that has been checked with the day with the people.

The alternatives that are available that has been you know we find out from people, how they are you know the criteria they are using to select among alternatives and then we build the preliminary model and then which test is accuracy in light of the work we have done and that is all we have time for in this lecture, we will meet again with some more inputs on qualitative research, thank you very much for listening.