



Australian
National
University

An Introduction to Qualitative Data Analysis for NVivo

Digital Literacy Team

Presenter: Nyree Mason

Nyree.Mason@anu.edu.au

digital.literacy@anu.edu.au

Purpose of this Presentation

NVivo is a complex program which can be confusing for people with little or no experience in conducting qualitative research.

This presentation aims to provide you with an introduction to the process of qualitative data analysis with specific reference to working in NVivo.

You will learn about the processes of:

- Creating Coding Schemes
- Ensuring Reliability and Validity
- Coding Data

As well as some NVivo-specific terminology.



Purpose and General Research Methodology

QUALITATIVE DATA ANALYSIS (QDA)

The Purpose of QDA

To investigate phenomena by analysing non-numerical data (verbal, written, observational) for **meaning**, in order to gain **insight** and **understanding**.

Also useful in mixed qualitative/quantitative methodology to help explain the reasons behind quantifiable differences and relationships found statistically, and to generate experimental hypotheses.

Quantitative <u>What?</u>	Qualitative <u>Why?</u>
Measurable phenomenon	Unquantifiable phenomenon
Numerical data	Verbal/written/observational/etc. data
Analysed by numerical comparisons/ relationships	Analysed by themes (quantifiable to an extent)
Evaluated against statistical probability and the validity & reliability of data	Evaluated against face validity & reliability of coding scheme
Results inferred to a population	Results describe the sample

Qualitative Data Analysis

Qualitative data is analysed by creating a structured, valid and consistent (reliable) **Coding Scheme**. This allows us to:

- **Categorise,**
- **Synthesise, &**
- **Summarise.**

Qualitative Data Analysis

Coding is the process of **identifying** sections of data where themes occur and labelling them.

NVivo is designed to “file” all the sections of text/audio/video/pictures that you **code** into your **Coding Scheme** “folders” (called **NODES**).



Coding schemes

Top-Down, Bottom-Up or Both?

Coding Schemes

The **Codes** in your scheme are determined by the **scope** of your research and your specific **research questions** (*if you have them*).

They are generated in a **top-down, bottom-up** or **iterative process** (involving both).

Defining an Initial Coding Scheme

Top-Down

- Based on specific **research questions** or **previous research findings**

A Priori/
(Confirmatory)

Bottom-Up

- Based on **what you find** in a sample of the data as you read it (*Grounded Theory*) and/or an NVivo **Word Frequency Query**

A Posteriori/
(Exploratory)

Which Method?

I personally recommend an **iterative** approach:

1. **Define** the Scope of your Research
2. **Create initial Codes** based on specific research questions
3. **Adjust/Add** to these initial Codes when you find unexpected & relevant information in your data

Which Method?

Top-Down-Only methods can lead to **confirmation bias**
→ *Be flexible, you can't predict everything.*

Bottom-Up-Only methods can lead to **implicit confirmation bias** plus make the process of QDA seem **endless & inefficient**

→ *At least define the scope of your research AND identify your implicit expectations.*

*Read to confirm & falsify your assumptions
(à la Karl Popper).*

Defining an Initial Coding Scheme

Considerations:

- **Relevant** vs **Interesting**?
- The **KISS** principle (Keep It Simple Stupid)
- **Positive** vs **Negative** contexts and attitudes
- **Free** vs **Tree** (Hierarchical) codes
- **General** vs **Detailed** (and “Outliers”)

Structure: Tree vs Free Coding

Codes in NVivo can be:

- **Stand-Alone Nodes** (“Free”)
- **Hierarchical Nodes** (“Tree”):
 - **Parent** Nodes (e.g., Assessment)
 - **Child** Nodes (e.g., MCQ, SAQ, Essay)

*Don't code at BOTH parent & child: NVivo allows for “aggregation”
Avoid “grandchildren”. NO duplicates.*



The Problem of Subjectivity

RELIABILITY AND VALIDITY

The problem of subjectivity

Devise a coding scheme which is as **objective as possible** and satisfies the requirements of:

- **Reliability** – **others** would code the data in the same way & **you** would code it the same way if you re-analysed it;
- **Validity** – **you** are identifying the themes you believe you are, and **others** in the same research area would agree.

Defining Codes in a **CODE BOOK**

Define **specifically & unambiguously** what you **mean** by each code & how it will be **applied** to your data. If possible give examples (*both positive & negative*).

*This information can be entered in the **Description** of each **Node** in NVivo to create a **Code Book**, explaining your coding scheme to people who read your research findings (and those who help with reliability checks).*

Ensuring Reliability & Validity

Assess the **Reliability** and **Validity** of your coding scheme early in the coding process using:

- **Coding Comparison Query** in **NVivo** comparing your coding to that of others OR to that you did previously (**reliability**)
- **Discussions** with colleagues/subject experts regarding the **validity** of the coding definitions in your code book.

Inter-Coder Reliability in NVivo

NVivo has a **Coding Comparison Query** which compares the amount of overlap in the coding conducted by two or more different people. The **amount of overlap should be high** if you have a reliable coding scheme.

This should be **checked early** in the coding process, so if there are problems, they can be rectified early (you don't want the onerous task of recoding **ALL** of your data!).

Ensuring Objectivity

Try to **eliminate bias** as much as possible:

- Read with an **open mind**
- Read to both **confirm AND falsify** any preconceptions/theories/hypotheses you have (and all humans have them so it's best to make them explicit. That way you know what you need to look out for and avoid implicit confirmation bias).



Applying Codes
Coding Context

HOW TO CODE

Applying a Coding Scheme

In simple terms, the process of coding is like highlighting sections of text/pictures in an article and then applying one or more **descriptive labels (Codes)** to it.

NODES

Each section is stored in a unique Code folder called a **Node** in NVivo. Technically, you “code your data to Nodes”.

E.g., The Node called “**Habitat**” stores all the sections (references) you have selected and coded as being related to the theme of “Habitat”. Open the Node, and you see all the references in one document.

Handy Hint: Contextual Nodes

It's also a good idea to have some Nodes for Contextual Information, useful for analysis:

- A **Node for each question** in a survey or interview. *Then you can analyse responses to individual questions more easily.*
- Nodes for **Attitudes and non-verbal cues** useful for analysing themes in context (*positive/negative, sad/angry etc.*)

Consider Coding Context

How much information you include in your coding reference (e.g., partial/whole sentences, paragraphs) is very important for **revision** and **analysis**.

- **Include** enough information to justify how **appropriate** your coding was during revision.
- **Exclude** superfluous information if you want to avoid **spurious correlations** in analysis.

Example Node: “*TV is not educational*”

When you open the Node, what do you see?

“television” ?

vs.

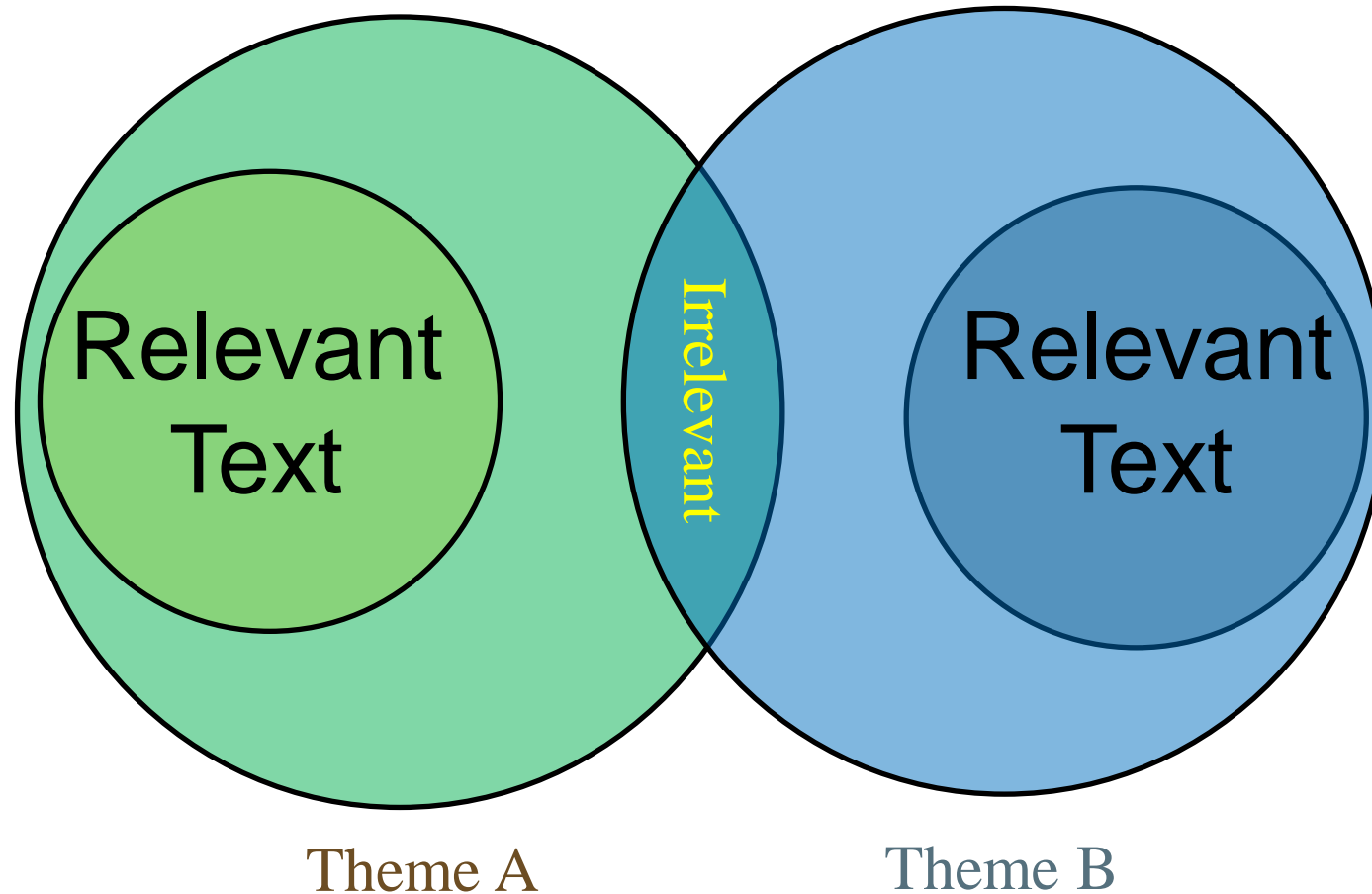
“I find television very educating.” ✘

vs.

“I find television very educating. Every time somebody turns on the set, I go into the other room and read a book.” ✓

Example Coding Query in NVivo:

What do Theme A and Theme B have in common?



NVivo Introduction

The NVivo Introduction Workshop will teach you how to:

- Set-up an NVivo Project efficiently
- Create Nodes
- Code text, audio, video & image data
- Transcribe audio & video within NVivo
- Use additional organisational and note-taking features

NVivo Advanced

The NVivo Advanced Workshop will teach you how to analyse your data once it has been coded, using:

- Word Frequency Queries
- Text Search Queries
- Basic & Matrix Coding Queries
- Cluster Analysis

Plus how to use concept maps and special graphical analysis features.



SEE YOU THEN...