# Autocoding

Useful for using existing structures in your source material (like columns/rows in a spreadsheet or heading styles in a Word document) to automatically create **Cases** or **c**ontextual **Codes** (e.g., a Code for each question in a survey or interview). You can also ask NVivo to Autocode your text data to create initial **Themes** and **Sentiment Codes**. I do not recommend using the latter two, as they may not do an accurate job of coding the complex qualitative data you find in an academic context. If used, the results MUST be thoroughly checked for reliability and validity, including ensuring nothing was missed in the coding process (difficult to check), and this will likely be more time-consuming than coding to them yourself. There is no substitute for a human subject expert like yourself.

# Cases

These are the *units of measurement* in your research. They can be individual people/reports/articles (or other special instances you wish to be able to measure) (e.g., "if you want to be able to say "*X number of people mentioned Y*", then people are your cases). You can have more than one type of case in your data set.

These are technically another form of **Codes** to which you code all the qualitative data related to the individual cases in your project (e.g, everything person X said will be coded to **Case X**). **Please note** that <u>this coding counts</u> in total numbers of **References** and **Codes** in Files.

# **Classifications and Attributes**

Classifications are used by NVivo to categorise your **Files** and your **Cases**. **Attributes** are the relevant categories/variables that relate to each **Classification**.

### File Classifications:

The ways in which you categorise the different types of **Source** (and reference) material you have. For example, interviews, surveys, webpages etc.. <u>Please note</u>: they are designed to be used for file management purposes, rather than analysis.

# File Classification Attributes

The relevant classification variables related to the type of **Source** material (e.g., information about when and where interviews were conducted and by whom). Within each attribute is a defined set of possible **Values** (e.g., within the attribute *Interviewer* you would have the different names of interviewers to choose from as values).

### **Case Classification**

The ways in which you categorise the different types of **Cases** (units of measurement) you have in your data. For example, people, reports, articles etc.. <u>Please note:</u> these are designed to be used for analysis purposes and become the "independent" or "breakdown" variables for your analyses.

#### Case Classification Attributes

The relevant classification variables related to the **Case(s)** you are studying (e.g., the predefined *Person* **Classification** has *Age, Sex* and *Occupation* attributes). Again, you need to define a set of possible **Values** for each **Attribute**.

## **Cluster Analysis**

This provides a tree-diagram and similarity metrics describing the relationships between the Codes, Cases or Files you select based on their similarity, either in terms of Words used, Files Coded or Case Attribute Values). Codes on closer branches of the tree, are more closely related on the relevant measure than those further apart. This can be useful for identifying superfluous Codes as well if you need to simplify your coding scheme.

## Codes

This is the name NVivo uses for a **collection** of all the sections of qualitative data (**references**) that have been coded as belonging to a particular theme. For example, all the segments of text/audio/video/picture in the project which refers to assessment will be **coded to a Code** called *"Assessment"*. Codes were known as **Nodes** in older versions of NVivo.

#### Free Codes:

Stand-alone topics/themes (e.g., *Recommendations*).

#### **Tree Codes:**

Hierarchical themes. The higher-order Code is termed the "**parent**" and lower-order Codes "**children**" (e.g., *Assessment* could be a parent Code and *Multiple Choice, Short-Answer, Essay* etc. would be its children).

### **Percent Coverage**

A curious statistic which tells you the percentage of the source that the reference coding represents (e.g., the number of characters coded represents X percentage of characters in the entire source file). It means little in complex qualitative data analysis.

# Coding

The process of identifying important **themes/topics/ideas/concepts/keywords** in your qualitative data, selecting the relevant text/audio/video/picture and applying **descriptive labels (or codes)** to them. This allows you to collate all the information related to each theme, quantify, organise and analyse it in order to summarise and synthesise the data as a whole. **Codes** must **be well defined** so they can be consistently applied to the data (for **reliability**) and their meaning easily understood (for **validity**).

### **Coding Context**

How much information you include in your **Codes**: individual words, partial/whole sentences, etc.. You should provide only <u>relevant</u> content in your coded data in order to understand why it was coded in that and to prevent spurious correlations between **Codes**.

#### Coding On

Often one reference will be a relevant for more than one Code. This is NVivo's term for coding one section of data to multiple **Codes**, and this function eliminates the need for duplicate Codes.

#### **Contextual Codes**

Codes to assist with analysing specific **subsections** of your data (e.g., individual survey questions, report sections etc.). **Context Codes** can also include things like **attitudinal constructs**, such as positive, negative, mixed attitudes. If you are working with Word documents or Spreadsheets, you can use the **Autocoding** function to create **contextual Codes** and **Cases**.

#### Files

These are the **documents and other files** containing the data you wish to analyse: text, audio, video and picture files as well as spreadsheets. You can also store other documents such as reference material, in the **Sources Area** of NVivo.

#### **Internal Files**

All the data and notes which you want NVivo to analyse should be stored here within the NVivo Project itself.

#### Externals

Any data that you do not need to analyse within NVivo you can store as <u>links</u> to data files, websites etc which are <u>outside</u> the Project (e.g., things that you cannot import into NVivo, PowerPoint slides, or really big files like **audio** and **video** that are better to store outside to **save computing power**).

#### Links

#### Annotations

Notes attached to specific sections of your source material (e.g., a portion of text, picture or audio timeline). Annotated content will be highlighted in <u>blue</u> and the linked text will be displayed in the **Annotations** tab at the bottom of the window. This text can also be included in any text searches and/or queries.

#### See Also Links

These link specific sections of your source material (e.g., a portion of text, picture or audio timeline) to another **Project Item** (e.g., File, Code, Case) using a hyperlink. Linked content will be highlighted in <u>pink</u> and the linked item will be displayed in the **See Also Links** tab at the bottom of the window.

#### Memos

**Memos** can be linked to specific **Files** or **Codes or Unlinked.** They can include text, tables as well as annotations. In addition, you can **code** the content of memos and the text can be included in any text searches and/or queries.

# Maps

These are diagrams that can help you to visually represent your Project structure or the end result of your analysis. There are 3 types:

#### Mind Maps

Visualise your coding scheme and create Codes automatically based on the map structure.

#### **Project Maps**

Show you how your Project items are related (e.g., what your coding scheme hierarchy is at any given time).

#### **Concept Maps**

A free-form mapping tool, where you have full control over the elements and connections between them.

### **NCapture**

This is an add-on for **Internet Explorer** and **Google Chrome** web browsers to assist with importing webpages and social media into NVivo

### Queries

This is the term NVivo uses for its **analytic tools**.

### **Coding Query**

This will produce a summary of all text/pictures/audio/video sections **that have been coded** in a specified way (e.g., to all the Codes X, Y and Z).

### Coding Comparison Query

This is used to assess the **Reliability** of your coding scheme by comparing the coding conducted between 2 people (or the same person at two different times, using two different User Names).

### **Compound Query**

This will produce a summary of the results that satisfy a **combination of text and coding queries**.

### Matrix Coding Query

This produces a cross tabulation of Codes and/or Attributes on the x-axis and y-axis. By default it shows you the number of references corresponding to each cell, but this can be changed to represent cases or files.

#### **Text Search Query**

This will produce a summary of all the words that fulfill a **specified text search term**. These search terms can include **Wildcards**, **Fuzzy Searches**, **Boolean Operators and Near** terms.

### Word Frequency Query

This will produce a list of 1000 words that occur **the most often** within the <u>text</u> of your documents (including memos, transcripts etc.). You can choose the number of exact matches, or different levels to group similar words (stemmed, synonyms, generalisations and specialisations).

### References

This is the term used by NVivo for the **discrete sections of qualitative data** that have been **coded** a specified way. One **Code** can have **multiple references** (i.e. sections of text) related to one **Case** (i.e. one person); so please take care to ensure that you <u>do not confuse</u> the number of **References** with the number of **Cases** that have said **Theme Y**.

<u>Please note:</u> the following examples counterintuitively count as 2 references in NVivo:

- Coding one continuous reference across 2 pdf pages.
- If 2 Project Users code the same section of text, to the same Code.
- If you code 2 contiguous sentences and don't include the space between them

# Relationships

These are links that you can use in NVivo to keep track of any relationships (or connections) which exist between Codes, Cases and/or other project items (e.g., a *University* may be <u>a member of</u> the *Group of 8,* so you can create a *Membership* relationship to link the 2 items) These must be setup in the **Classifications** area first before been applied to project items. Relationships between two Cases, can be recorded and displayed on a **Social Network Analysis** Sociogram.

# Reliability

Reliability refers to the **consistency with which codes have been applied** to your data If multiple coders all agree with how the coding should be applied (i.e. they all code the same sections of text in the same way), then you have a reliable coding scheme. Reliability can be assessed using a coding comparison query in NVivo.

# **Text Search Terms**

### **Boolean Operators**

- AND (or &) will search for BOTH terms (professional AND development).
- **OR (**or **| )** will search for EITHER of the terms (*professional* OR *professionalism* OR *ethics* OR *ethical*).
- **NOT (- or !)** will search for the first term and EXCLUDE any results that also contain the second (*professionalism* NOT *staff*, *professionalism* -*staff*).
- **REQUIRED (+)**, will search for items containing the required term and the second term is optionally included in results if and only if it occurs together with the first term (+*professional development*).
- **Double Quotes "** "can be used to enclose a group of words to search for a specific phrase ("*professional development*") with no exceptions.
- Round Brackets () can be used to group clauses to form sub queries.

### **Fuzzy Search**

The symbol ~ will look for all words that are similar in structure (e.g., analyze~ will result in *analyzes, analyse,* etc.).

#### Near Search

Helps you to find the co-occurrence of words within a specified distance (<u>e.g., "professional AND</u> <u>development"~15</u> will find instances of professional <u>and</u> development within 15 words of each other only).

#### Wildcards

The ? wildcard stands-in for any single letter (e.g., g?t will result in *get, gut, got* etc.).The \* wildcard stands-in for multiple letters (e.g., g\*t will result in *get, great, gradient* etc.).

## Validity

Validity Refers to the legitimacy of your coding scheme. In qualitative data analysis, you can only really assess the face validity (or surface validity), based on logic and consensus. Essentially, are you measuring what you think you're measuring?

# Areas in the NVivo 14 Navigation Pane

## Data Area

- Files: Import and view internal source files that are stored within the Project
- <u>File Classifications</u>: Create and store File Classifications, Attributes and Values
- <u>Externals</u>: Create links to and view **external sources** like file paths and website addresses stored outside of the Project

## Coding Area

- <u>Codes</u>: Create, modify and view Codes
- <u>Sentiment</u>: Store the results of Nvivo's Autocoding for Sentiment
- <u>Relationships</u>: Apply and store **Relationships**
- <u>Relationship Types</u>: Create Relationship types to apply to Project items

### Cases Area

- <u>Cases</u>: Create, modify and view **Case codes**
- <u>Case Classifications</u>: Create and store Case Classifications, Attributes and Values

#### Notes Area

- Store and view linked Annotations, See Also Links and Memo Links in the Project
- <u>Framework Matrices</u>: Create and modify your own notes in a tabular form with Cases as rows and Codes as columns.

#### Sets Area

- Static Sets: Create and view Sets folders that contains links to specific project items
- <u>Dynamic Sets</u>: Create and view **Search Folders** that contain specific search queries for Project items (e.g., all interviews conducted by Fred) that run every time you open them

### Queries Area

- <u>Query Criteria</u>: View and re-run the criteria/parameters of saved **Queries**
- <u>Query Results</u>: View saved **Results** from Queries that were run
- <u>Coding Matrices</u>: View Matrix Coding Query results here (optional)

### **Visualizations** Area

• Save, modify and view Mind Maps, Concept Maps and Project Maps

### **Reports Area**

• Run Formatted or Text Reports summaries of NVivo's contents