



Australian  
National  
University

# Digital literacy training

---

## **NVivo14 Intermediate**

More Coding Tools  
2024



# Table of Contents

<b>NVivo Intermediate: More Coding Tools .....</b>	<b>1</b>
Opening NVivo and a Sample Project .....	1
Compatibility with other versions of NVivo .....	1
Mac vs PC Versions of NVivo .....	1
<b>Other Methods for Creating Codes &amp; Cases.....</b>	<b>2</b>
Autocode.....	2
Autocoding Datasets .....	2
Autocoding Word Documents .....	3
Autocoding Cases .....	3
Autocoding Codes.....	3
Autocoding When Importing Survey Spreadsheets.....	4
Autocoding Themes and Sentiment.....	4
<b>Creating &amp; Coding to Relationships .....</b>	<b>5</b>
<b>Deciding on a Coding Scheme: Word Frequency Queries .....</b>	<b>6</b>
Word Frequency Results.....	7
Word Clouds .....	7
Tree Map.....	8
Cluster Analysis .....	8
Excluding Irrelevant Words from Searches.....	8
Looking at the Results in Context .....	8
<b>Initial Mass Coding: Using Text Search Queries.....</b>	<b>9</b>
Text Query Results Preview .....	10
Word Trees.....	10
Saving the Query Preview Results as a Code.....	11
Expanding the Coding Context of References .....	11
Automatically Expanding the Coding of References .....	11
Querying Within Project Items .....	12
<b>Checking Inter-Coder Reliability.....</b>	<b>12</b>
Exporting the Codebook .....	12
Creating a Set Folder .....	13
Creating a Project Copy.....	13
Coding Comparison Query .....	14
<b>Exporting.....</b>	<b>15</b>
Exporting Results of Queries .....	15
Exporting Codes.....	15
Exporting the List View of the Codes Area .....	15
Exporting Classification Sheets.....	15



---

## NVivo Intermediate: More Coding Tools

---

*Lumivero's* NVivo software can do a phenomenal number of things to help you analyse your qualitative data. It can also help you analyse a large variety of data types: text, audio, video and pictures. This course will show you how to use the **essential tools** available in NVivo for qualitative data analysis in an academic context. For further information about NVivo's full capabilities, please refer to the **NVivo 14 help** website:

PC: [help-nv.qsrinternational.com/14/win/Content/welcome.htm](http://help-nv.qsrinternational.com/14/win/Content/welcome.htm)

Mac: [help-nv.qsrinternational.com/14/mac/Content/welcome.htm](http://help-nv.qsrinternational.com/14/mac/Content/welcome.htm)

This session will cover how to:

- use **Word Frequency Queries** to decide on a coding scheme
- initially mass code your text data using **Text Search Queries**
- **Autocode** to create contextual **Codes** and **Cases**
- check the reliability of your coding scheme using **Coding Comparison Queries**

### Opening NVivo and a Sample Project

When you **open NVivo** for the first time, it will ask you to create a **User ID**: enter your name and initials. At the top of the **Start Screen** open the **Multi-Method Sample Project** as it has the broadest range of data types to explore. It is a two-year study (2008-2009) documenting community perceptions of development and land-use change on coastal communities in the Down East area of Carteret County, North Carolina, USA.

### Compatibility with other versions of NVivo

**If you are collaborating with others on the same NVivo Project, you must use the same version of NVivo** (e.g., NVivo 14). Projects created in older versions of NVivo can be opened in newer versions. However, they will need to be **converted** to the newer format in order to do so. Once this new converted file is saved, it cannot be opened in older versions of NVivo again.

### Mac vs PC Versions of NVivo

**Mac versions of NVivo do not have the same capabilities as Windows versions** (e.g., you cannot Create Cases at import on Macs) and there are features in Windows versions that are not supported in Mac. **You can work on the same project on different operating systems (PC and Mac)** although it's not recommended. You will need to convert your Windows Project to a Mac format and vice-versa when transferring Project data between computers. Please see NVivo 14 Introduction handout for further information.

**These course notes apply specifically to the PC version of NVivo**, so if you are working on a Mac and have trouble finding what you need, please refer to the online NVivo 14 Help for Mac:

[help-nv.qsrinternational.com/14/mac/Content/welcome.htm](http://help-nv.qsrinternational.com/14/mac/Content/welcome.htm)

## Other Methods for Creating Codes & Cases

### Autocode

NVivo 14 has more **Autocoding** functions than previous versions. You can use it to create:

- **Cases:** e.g., identify speakers in a transcript and code their responses to their Cases.
- **Contextual Codes:** e.g., a Code for each question in an interview/survey.
- **Identify Themes:** identify frequently mentioned ideas or concepts to form initial Codes.
- **Identify Sentiment:** identify the emotional context of the text: Very Positive, Moderately Positive, Moderately Negative and Very Negative.

### Autocoding Datasets

This **works best** if you clean your data first and have a spreadsheet file where:

- the Headings are in the first row,
- there is no superfluous information (e.g., annotations, calculations),
- each row represents responses from one Case only, and
- each column represents a Code (such as a survey question).

Then you can automatically code datasets by either:

- **rows** to create a Case for each respondent and code the responses in row to that Case, or
- **columns** to create a Code for each question in a survey and code the responses in the column to that Code.

Example: Autocode the columns in the *Survey Responses Data* file as question Codes.

In the **List View** of the **Data** area, select on the dataset(s) you want to Autocode (e.g., the *Survey Responses Data* in the *Survey* folder). On the **Home** tab click **Autocode** and select **Use the style or structure**. Click **Code columns** then click **Next**. Move the appropriate column(s) you wish to Autocode into the **Selected Columns** box by selecting them and clicking the right arrow button (e.g., all of the available columns) then click **Next**. Select a location for the **Code** (e.g., under **New Code** in the location **Codes** or append to **Existing** ones you have them already), and give the new Code a name (e.g., *Survey Questions*). Then click **Finish**. This will create a new **Parent Code** called *Survey Questions* with multiple **Child Codes** in the **Codes** folder.

Example: Autocode the rows in the *Survey Responses Data* file as Case project items.

Similar to the procedure above, only when the Wizard opens, select **Code rows**. **Choose the column that contains Codes or case names** from the drop-down folder (e.g., *Respondent*). You also have the option of filtering out some cases (e.g., cases below a certain age in the age column) or grouping them (grouping Cases aged 18-30, 31-49, etc. if they aren't already), by clicking the **Filter and Group** button next to this menu. This may be useful if your dataset isn't already "cleaned", however it is **best practice to clean the data before importing it** in a more suitable program, such as Excel or SPSS.

Your data is unlikely to warrant a hierarchical structure, so click **Next**. Select the columns you want to code to Cases in the **Available columns** box by holding the CTRL key and clicking on them (e.g., all question columns), then move them into the **Selected columns** box by clicking on the right arrow, then click **Next**. Usually, you won't already have Cases for this new import, so from the **Under** drop-down menu, select **New Code or Case** (if you are adding content to existing Cases, select **Existing Code of Case**). The Location box should have **Cases** already selected. Given them a group **Name** (e.g., *Interviewees*), then click **Finish**. The new Cases will appear in the higher-order **Cases** folder, in the **Cases** Area.

## Autocoding Word Documents

**Autocode** can also be used in **Word** documents, if you have used consistent **Heading Styles** to define content relating to Cases, questions asked, general themes etc. In the Interview Transcripts of the Sample Project, they have used *Heading 1* for each interview question asked and *Heading 2* for each speaker's name (see excerpt below). Everything the speakers say in the transcript is in *Normal* text.

### Q.1. Connection to Down East

**Henry**

*Tell me about your personal and family history in Down East. How long have you or your family been living Down East full time or part time?*

**Barbara**

My family moved here when I was two years old in 1969. My parents still live here. They live down in Gloucester. But I was raised in Beaufort, in town, and went to Beaufort Elementary and middle school and high school, then moved away for college. So I've lived here most of my life although I've moved away.

**Henry**

*And you've lived Down East how long?*

**Barbara**

Since '96. My husband and I bought this little cottage in '96.

## Autocoding Cases

It's probably easier to explain how Autocoding works using the interview transcript above as an example. If you want to create a **Case** for each speaker, you will need to code by **Heading Style 2**. NVivo will automatically create a new Code for each unique title in that style (e.g., one Case called "Henry", another called "Barbara"). It will then automatically identify all of the sections of text belonging to *Henry* and code it to the **Henry Case**, and all of the text belonging to *Barbara* to the **Barbara Case**. It does this by identifying all of text in between instances of *Heading Style 2* ONLY (regardless of any other *Heading Style* used in between) and coding it to the Case Name written in the preceding *Heading Style 2*. If NVivo finds the identical text in *Heading Style 2* again (e.g., indicating more text relating to "Barbara"), it just adds that section of text to the existing Case it created earlier.

Example: Autocode to Case in the Interview Word Documents.

Ensure all the documents you wish to Autocode are identically formatted using consistent heading names and styles. If you find any inconsistencies, it's easier to correct them now before coding. Select all the documents you want to code (e.g., all the text documents in the *Interviews* folder) in the **Data** area using the CTRL key to click on multiple files. On the **Home** tab, click on **Autocode** and select **Use the style or structure**, then click **Next**. Select **Paragraph styles** and click Next. Select all the appropriate paragraph style(s) from the list of **Available paragraph styles** (e.g., *Heading 2* to create a Case for each speaker) and move it/them into the **Selected paragraph styles box** using the double right arrow button. Click **Next** and then in the **Under** drop-down menu select **New Code or Case** (unless you are adding information to existing ones) and in the **Location** area, click on the **Select** button to change the default **Location** from **Codes** to **Cases**. Give the group of Cases a **Name** (e.g., *Interviewees*), then click **Finish**. You'll find the new Case in the Cases folder in the Cases Area.

## Autocoding Codes

If you want to create a contextual **Code** for responses to each question asked in this interview transcript, you will code by *Heading Style 1*. Again, it will follow the same procedure as indicated above. **Please Note** that if there is a typo in your *Headings*, NVivo will create a **duplicate Code** for that Code/Case, and this can simply be merged into the primary Code (without the typo) as described in the *Introductory Course Notes*.

Example: Autocode to Question Codes in the Interview Word Documents.

Follow the same procedure as for Autocoding Cases but leave the default **Location** set to **Codes**. Give the group of Codes a **Name** (e.g., *Questions*), then click **Finish**. You'll notice in the Codes folder there are 2 Codes for Q.6. This is because the questions were not identical, so you will need to merge these two together as described in the NVivo Introduction handout.

**Note:** You can also create a hierarchical coding structure for Codes if you select multiple Heading Styles at the same time (e.g., *Heading Style 1* for the parent Code, and *Heading Style 2* for the child Codes). Be sure to enter them into the paragraph styles box in the correct parent/child order. You can also **Autocode** documents by **Speaker** if you have identified the person speaking in the transcript at the start of each paragraph of the transcript (not using *Heading Styles*). This is a more laborious process however, as you need to type in the names of every speaker in the documents first so NVivo can identify them. So, check for typos in people's names beforehand.

## Autocoding When Importing Survey Spreadsheets

If you have a survey spreadsheet from an online survey application, such as Survey Monkey, you will already have all the information you need about your Cases and Questions neatly organised. Luckily, you can create **Cases, Question Codes, Case Classification Attributes** AND populate them with their relevant details upon import, in one go. Again, the process is easiest if you follow the rules for spreadsheet set-up outlined above under "Autocoding Datasets".

On the **Import** tab click on the icon for **Survey** and select the file format your survey is in (e.g., Excel) and then select the file in the next window. You can also right-click in blank space in the **Detail View** of the **Files** area and select **Import Survey** and choose the format from the menu. You will then see an explanation of the Survey Import Wizard process, then click **Next**. The Wizard will guide you through the process, providing previews of the result of your choice throughout. At Step 2 tell NVivo how many of the first rows of your spreadsheet contain headings only (best practice is only 1, but you can have 2 if you select it from the drop-down menu). If you have any dates in the file, select the date format used from the drop-down menu, then click **Next**. Choose where you want the **Cases** to be located (the default is usually fine), select the column containing **unique ID for each of your cases** from the drop-down menu, and choose the **Case Classification** you want to use for the Cases in the spreadsheet (usually **Create new classification**, but you can add to an existing one if need be). Click **Next** and check that NVivo has correctly identified the purpose of each column: **Attributes** should be **Closed Ended**, **open ended questions** should be **Open Ended** and if there are any superfluous columns in the spreadsheet you can choose not to import them at this stage too. Then click **Next** and choose NOT to Autocode Themes or Sentiment for the open-ended questions by unticking the boxes. Then click **Finish**.

## Autocoding Themes and Sentiment

As mentioned previously, NVivo 14 has two new Autocoding functions based on the language used in your qualitative data. One, **Thematic Autocoding**, tries to identify frequently mentioned ideas or concepts and use these to form initial **Thematic Codes** for you. The other, **Sentiment Autocoding**, attempts to identify language that expresses positive or negative feelings, emotions or opinions in order to code the **Sentiment** expressed at four strength levels: *Very Positive*, *Moderately Positive*, *Moderately Negative* and *Very Negative*. These **Codes** are stored in a dedicated **Sentiment** folder in the Codes area. Please note: you cannot add to these options or change them. It's therefore **best to create your own sentiment Codes** in the Code area.

**I do not recommend using Sentiment or Thematic Autocoding for academic research.**

The language used in academic data tends to be complex and nuanced, and the reliability and validity of the coding done by a subject matter expert, such as yourself, will always be far superior. These functions also tend to over-code too much, identifying a large number of Codes that only contain one reference, and "false positive" results that you will have to uncode manually. There is also a great risk of "false negatives": sentiment and themes that were not captured in the Autocode at all. If you do choose to use it, you will always have the responsibility to read through the results to verify the validity of the coding done and make the appropriate changes.



The procedure and display of results for **Autocoding Themes** and **Sentiment** are almost identical, but Autocoding Themes is slightly more complex, so we will cover this only in the example below.

Example: Autocoding the Themes in the Interviews.

In the **Files** area, select all the files in the *Interviews* folder and on the **Home** menu tab, click on **Autocode**. Select **Identify themes** then click **Next**. NVivo will provide a list of suggested themes in a hierarchy which you can explore and choose to accept or reject specific suggestions by ticking or unticking the boxes associated with them (e.g., untick *thing*, *gonna* and *place*)., then click **Next**. Select how much of the surrounding content should be coded to the initial **Codes** (e.g., *Code sentences*), and click **Next**. Give the new folder for the Autocoded Themes a **Name** (e.g., *Autocoded Themes*) then click **Finish**. You will find the new folder in the **Codes** area and two new **Detail View** windows will appear. The first is called **Autocoded Themes Results** with a date and time, which has results listed per interview. One sub-tab has a matrix with *Interviewees* in rows and *Themes* in columns, and the other, a chart of these results. The second is called **Compared by number of coding references** and provides a summary of all the themes coded in the interviews. The first sub-tab is a **Hierarchy Chart** displaying the themes and subthemes as separate colour-coded blocks, with the size of each rectangle representing the weighted percentage of occurrences (number of characters coded divided by total number of characters in all interviews). The **Summary** sub-tab provides a table of coding details: number of references and items (files).

## Creating & Coding to Relationships

**Relationships** are technically a type of Code which can be used to:

- note general relationships between any two **Project Items** (e.g., the *Person Case Margaret "is a member of"* the *Place Case Cartaret Catch*),
- note relationships or associations between **Cases** (e.g., *Daniel "is married to" Maria*).
- to collect evidence for a relationship between **Codes** (e.g., coding data to support a theory that *Community Change "contributes to" Environmental Change*).

Relationship types (e.g., *"is a member of"* and *"is married to"*) need to be created first in the **Codes** Area. Then these can be applied to create specific Relationships between 2 Project items later. These relationships are stored in the **Relationships** folder in the **Codes** area.

Example: Create a relationship called *"is friends with"* and apply to *Daniel* and *Paul*.

To create a **Relationship Type**, right click on the **Relationship Types** folder in the **Codes** area and select **New Relationship Type**. Type in a name (e.g. *"is friends with"*) and [optional] description. Then select the appropriate direction (e.g. *two-way*).

To create a **Relationship** between Project items (e.g., *Daniel "is friends with" Paul*) on the **Create** menu tab click on the **Relationship** icon (or right-click on the **Relationships** folder in the **Codes** Area then select **New Relationship**). Next to the **From** box, click the **Select** button and choose a Project item (e.g., the *Daniel* Case) and click **OK**. Then choose a **Relationship Type** from the drop-down menu (e.g., *"is friends with"*) and selecting another Project item in the **To** box (e.g., the *Paul* Case). Then click **OK**, and **OK** again.

Example: Coding content as evidence to support the *"impacts"* Relationship between *Real Estate Development* and *Water Quality*.

Select the text you want to code, right-click on the selection and choose **Code Selection**. Expand the **Relationships** list by clicking on the triangle next to it and select the Relationship you want to code to (*Real estate development (impacts) Water quality*), then click the big blue **Code selection to...** button. Open the Relationship in the **Relationships** folder by double-clicking on it to see all the relevant information coded to it.

**Note:** Relationships between Cases can be used in **NVivo 14** to perform an approximation of Social Network Analysis (SNA) which will be covered in the Advanced notes. However, if you want to perform a SNA with a large amount of data in a statistically appropriate way, it is best to use a program more suited for the purpose. The CSIRO has a free open-source tool called **Constellation**: <https://www.constellation-app.com/index.html> and another one called **StellarGraph**: <https://www.stellargraph.io/> which you may find helpful.

## Deciding on a Coding Scheme: Word Frequency Queries

In the initial stages when you are still deciding on a coding scheme, a **Word Frequency Query** could be useful. This will produce a list of all the words that occur most often in the text documents you have (including memos, transcripts etc. unless you specify otherwise). You can search the entire content of your qualitative data, or you can narrow it down to specific groups of people or individual questions, etc.

Example: Word Frequency Query for Interview Question 4.

On the **Explore** tab, click on **Word Frequency**. This will open a new **Detail View** with the Word Frequency options. Choose where to search in the **Search in** area. For this example, click **Selected Items** and then click the + next to the *Codes* folder to see the other folders within it. Then click on the name of the *Autocoded Responses* folder (not the checkbox - if you do all the Codes within will be selected) and then on the + next to the *Autocoded Interview Questions* folder and tick the box for *Question 4*. Then decide if you want to **limit the number of results** (usually the default of the 1000 most frequent is fine) or change the largest **letter-length** of the words you want in the results (the default of 3 letters is fine for English, so you don't get words like "to", "a", "an" or "I"). In the **Grouping Area**, decide on the grouping criteria by sliding the indicator (e.g., I recommend looking at results from both **With stemmed words** and **With synonyms** Queries - see table below for descriptions of all 5).

**Note:** A Text/Word Search Query does not search text within PDFs created by scanning paper documents as pictures. Consider using optical character recognition (OCR) to convert the scanned images to text before importing. The results also do not include words in NVivo's **Stop Words List** (more on this later).

### The five Text Match levels:

Grouping Level	Returns	Example (sport)
Exact matches	Exact matches only	sport
With stemmed words	+ Words with same stem	sport, sporting
With synonyms	+ Synonyms	sport, sporting, play, fun
With specializations	+ Specializations	sport, sporting, play, fun, running, basketball
With generalizations	+ Generalizations	sport, sporting, play, fun, running, basketball, recreation, business

**It is good practice to save a detailed record of any analyses you conduct in NVivo or otherwise.** This serves as both a reminder as to what you did and how you did it, and also provides a short-cut to run the analysis again if you need to. To save the Word Frequency Query details, click **Save Criteria** and give the Query a name, then select **OK**. Then click **Run Query**. If you need to run the same Query again later, you will find it in the **Queries** folder in the **Search Area**. Just right-click on the Query and select **Run Query**.

**HANDY HINT:** If you need to **run multiple similar Queries** (e.g., Word Frequency for each interview question with the same grouping level etc.), **copy** your first Query (e.g., *Frequency Q1*), then **paste** it into the blank space in the **Queries** List View. This will create a new saved **Query** with identical parameters, called *Frequency Q1 (2)* and you will need to rename it to reflect the changes you will make (e.g., *Frequency Q2*). Right-click on this copy and select **Open Query**. Then change the parameter(s) you wish (e.g., **Selected Items** to Q2), and click **Run Query** to make the changes “stick” in this newly created and saved Query.

**Please Note** that if you ever change a parameter in a saved Query, then run it, that parameter is permanently changed.

## Word Frequency Results

There are several tabs on the right-hand side of the **Detail View** which present the results in different ways. From these you can decide on an appropriate coding scheme and even begin to create Codes. The **Summary** tab shows you a list of all the most frequent words, their frequency *Count*, other summary data and if you have chosen a **Grouping** level higher than *Exact matches*, you will also see the other words it has included as belonging to each Word Count (e.g., the word count for Change also includes that for *Changed, Changes, Changing*). There are also tabs for graphical representations of your Query: **Word Cloud, Tree Map** and **Cluster Analysis** (the latter only appears if there is enough data to run it).

**Note:** You will often see some very strange words included in groups (e.g., “get” will be grouped with “mothers” and “children”). This is because it uses an ordinary American Thesaurus to perform the grouping, and in some areas of America, children are also called your “get”.

In addition, the “ranking” of how **frequently a word or group of words occur, is automatically weighted by percentage**: number of characters in the word/words found as a proportion of the total number of characters in all of the documents analysed. The **Count** shown on the **Summary** is a more accurate metric at the **Exact** and **Stemmed Words** Levels. The size of the font in the Word Cloud and the size of rectangles in the Tree Map, will always represent this weighted percentage. As such, the **Word Clouds, Tree Maps and Results from grouping at a higher level than Exact, are not definitive – only indicative of the language used in your data.**

## Word Clouds

The **Word Cloud** tab displays up to 100 of the most frequent words in your text data, varying in font size: the largest are the *most* frequently occurring, and the smallest the *least* frequent (according to the weighted percentage, not count).



## Tree Map

The **Hierarchy Map** tab displays up to 100 of the most frequent words in rectangles of varying sizes: the largest rectangles are for the most frequently occurring words, and the smallest for the least (according to the weighted percentage, not count).

## Cluster Analysis

The **Cluster Analysis** tab displays up to 100 of the most frequent words in a tree-diagram describing the co-occurrence between these words. The closer the words are on the branches of the tree map, the more frequently they co-occur. This can help by providing more context. Please note that a tree map can only be produced if there is sufficient data.



## Excluding Irrelevant Words from Searches

In the results you might see words that you do not want to include in the Query search in future, such as "any". Right-Click on the word in the results and select **Add to Stop Words List**. If you have included more than exact matches in the text match level section (i.e., >level 1), you must remove any words in the box which you still wish to be included in future searches. When you are ready, click **OK**. You can **edit** the **Stop Words List** in **Project Properties** at any time: on the **File** tab under **Info** select **Project Properties** and on the **General** tab click the **Stop Word List** button and edit the words to exclude. Then click **OK** then **Apply** and **OK** again.

**Note:** Every time you run a Query from now on, these words **will not be included in Word Frequency or Text Search Query results (for this Project only) – even if you specifically search for them**. So, it is a good idea to check the **Stop Words List** before you run any **Text Searches** or **Word Frequency** Queries. There may be some words on the list that you **don't want NVivo to ignore** in these Queries. You will have to remove these words from the Stop Words List, before conducting your Queries.

## Looking at the Results in Context

If you think that one of the words in the Word Frequency results might be a good Code to have in your Coding Scheme, you can **view each word in the results in context** by automatically running a Text Search Query for that word or group of words. Double-click on the word in the Summary table/Word Cloud/Tree Map, or right-click and Choose **Run Text Search Query**. This runs a Text Search for the word (for exact matches) or words (for grouped matches) and opens a new tab in the **Detail View** containing all the relevant references containing the word(s).

**Note:** You can also **create initial Codes** from Word Frequency Queries by right-clicking on the word in the results and selecting **Create as Code**. However, it is probably best to run the **Text Search Query** first, before deciding the results would make a good initial Code at all (procedure described in the section below). Of course, the initial Codes will have to be "cleaned up" afterwards to both remove irrelevant and add relevant information.

## Initial Mass Coding: Using Text Search Queries

**Text Search Queries** can be used to help start the coding process for a large amount of textual data by finding instances of text matching specific **search terms** based on your pre-defined coding scheme (**internal text data files ONLY**). Information coded to Codes this way can always be modified later through uncoding if necessary.

**Note:** Consider checking the spelling before running Queries. You need to open a document and click the blue **Click to edit** text at the top of it. Then you will find the **Spelling** button on the far right of the **Edit** tab's menu in the **Document Tools** group of tabs which appear.

Example: Create an initial, automatically coded Code for Water Quality.

On the **Explore** tab, click on **Text Search** and this will open up a new **Detail View** window with the Query options at the top. Choose where to **Search in** (e.g., Selected Folders) and tick the boxes for the folders you want to search in (e.g., in *Interviews* and *Surveys*), then click OK. In the **Search for** area, enter the word(s)/phrase you want to look for. You can create more complex Queries using the special features displayed when you click the **Special** button:

- **Wildcards:** If you want to search for words that have alternative spellings (e.g., colour/color), or are commonly confused with similar words (compliment/complement), Wildcards can help to find every possibility. The Wildcard **\*** is used to stand in for one or more missing characters (e.g., g\*t will find get, great, gradient etc.). The Wildcard **?** is used to stand in for a single missing character (e.g., g?t will find get, gut, got etc.). Limited to **Exact Match** grouping level.
- **Fuzzy Search:** Also useful to pick up misspellings and find words with similar stems. The symbol **~** behind a word will look for similar words (e.g., *analyze~* will find *analyzes*, *analyse*, etc.). Limited to **Exact Match** grouping level.

**Note:** If you use the above functions you will be limited to **Exact match only**.

- **Boolean Operators:** When you want to find a specific concept you can use a combination of words and/or excluding a word(s) in your search.
  - **AND (or &)** will search for BOTH terms (water AND quality).
  - **OR (or |)** will search for EITHER of the terms (water OR lake OR sea OR river).
  - **NOT (- or !)** will search for the first term and EXCLUDE any results that also contain the second (water NOT tap, water -tap).
  - **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 (+Water river).
  - **Double Quotes " "** should be used to enclose a group of words to search for a specific phrase. If you search for a phrase, the results will not include words with the same stem. For example, if you search for "alternative energy", you will not be able find "alternative renewable energies".
  - **Round Brackets ( )** can be used to group clauses to form sub Queries.
- **Near:** If you want to find words that co-occur but you want to ensure they are mentioned in the same sentence rather than several sentences away, you can specify the distance either side of the required 2 words (e.g., "*water quality*"~20) will find instances of water and quality within 20 words of each other only.

**Note:** Avoid mixing-and-matching your search parameters (e.g., pollu\* OR "water quality") because it will not product the results you expect (but it won't tell you it didn't work either).

For this example, enter "water quality"~20. **Spread Coding** to a **Narrow Context**, so you can see enough information when you review the success of initial coding. Decide on the **matching criteria** (e.g., *With stemmed words*). Click **Save Criteria** if you want to save the Query parameters. Then click on the drop-down menu for **Run Query** and select **Run Query**.

## Text Query Results Preview

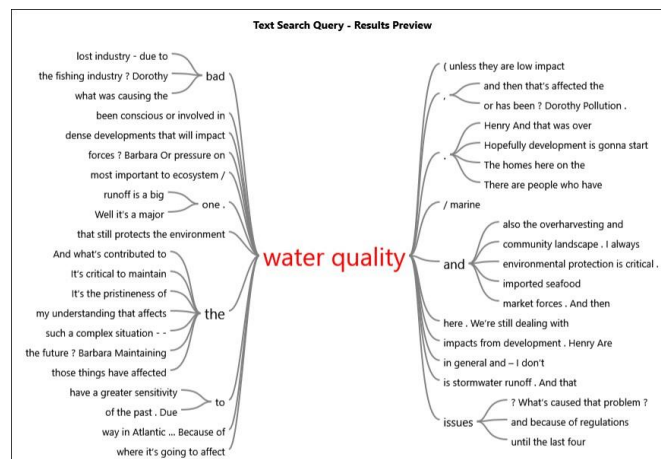
In the preview of the Query results, there will be tabs on the right of the **Detail View** for:

- **Summary:** A list of all the Data Files that contain the word/phrase results.
- **Reference:** The results are opened as a Code preview and the word or phrase is shown in context (which can be broadened or narrowed as outlined below).
- **Text, Picture, Audio, Video, or Dataset:** Displays the results found in each type of Data File (only relevant tabs are available).
- **Word Tree:** A graphical representation of the results in context.

## Word Trees

A **Word Tree** tab is only produced when you DO NOT choose to **spread coding** and before you save the results. The branches of the **Word Tree** represent the various contexts in which the word or phrase occurs. It can be useful for finding recurring themes in phrases that surround the word.

You can click one branch to highlight the other end of the branch to see the actual context. When on the **Word Tree** tab, a main **Word Tree** tab appears at the top of the window. This allows you to change the **Root Term** (e.g. to *the quality of the water*), the **Branch Order** (e.g., by **Number of Matches**) and **Context** level (the number of words surrounding the root term).



You can search for an exact quote from the **Word Tree** in all of your data by right-clicking on the appropriate branch, then selecting **Run Text Search Query**. This has the benefit of finding the quote quickly and getting a link to open the **Data File** from which it comes so that you can see it in more context.

## Saving the Query Preview Results as a Code

In general, if you run any Query and decide retrospectively that the results in the preview would be valuable to save as initial coding for a Code, click on the **Save Results** button at the top-right (or right-click in the Query results in the Detail View and select **Store Query Results**). From the **Option** list, choose to **Create Results as a New Code or Case** and change the **Location** to the **Codes** folder (with all your other Codes - and you will not be able to move it from Query Results later) and give it a **Name** and **Description**. Then click **OK**. You could also choose to **Merge Results into Existing Code or Case** if you are “chunking” (or merging the results of multiple searches) to minimise typos or logic errors.

**Note:** NVivo will just code the word(s) found in that Query unless you expand the coding to include more words (see how to do this below).

## Expanding the Coding Context of References

When you run a **Query**, by default the context for the references in the results is set to **Narrow**, and thus displays 5 words on either side of the result **greyed-out**. When you save the results as a Code, only the black text is coded to the Code, and you will need to add to this. In order to see more context to help with coding more data, you can expand the **Coding Context** for one reference at a time, or all references together.

To expand the context of one reference, within the **Query Results** or **Code** select and right-click on the chosen reference in the **Detail View**, then select **Coding Context** and choose the level you want:

- **Narrow** [default] to see 5 words (text), 5 percent (pictures), or hear/see 5 seconds (audio/video) more on either side of the coding reference, or (spreadsheets) see all the text in the cell.
- **Broad** to see the surrounding paragraph (text), cell or row (spreadsheets), or hear/see 20 seconds (audio/video) or 20 percent (pictures) more on either side of the coding reference.
- **Custom** to define specific reach settings.
- **Entire File**.

Then click **OK**. More context will appear greyed-out on either side of the coded text. If you want to expand the coding context for the entire results (or Code), select all the text at once either using the mouse or holding down the **Ctrl** and **A** keys, then right-click on the selection and follow the instructions above.

## Automatically Expanding the Coding of References

Sometimes it is easier to expand the coding itself as well, and then remove the extra content that does not belong in the Code. To do this, when you are in the **Code** you can select all the text-based coding references you wish to expand and right-click on the selection in the **Detail View** and choose **Spread Coding**. Then choose the level you want (same levels as listed for Coding Context).

**Note:** You can no longer select the whole document and expand the Coding if there are Audio, PDF or Video references amongst them. You need to select the text in the document **ONLY** and expand that.

## Querying Within Project Items

NVivo also has a handy shortcut which allows you to Query within specific Project Items that are open (e.g., Codes, Cases, Documents, etc.). For instance, when you have a **Code** open, on the **Code Tools** tab there is an icon to **Query This Code** with the following options:

- **Word Frequency Query in This Code** (automatically runs a standard Query whose parameters can be changed and run again).
- **Text Search Query in This Code** (automatically selects that specific Code to search in, and you can change the parameters before its run).
- **Codes Coding This Code** (runs a **Group Query** to identify all the other Codes coding the same Code content).
- **Cases Coding This Code** (runs a **Group Query** to identify all the Cases associated with the Code content).

## Checking Inter-Coder Reliability

An important measure of the reliability (consistency) of your coding scheme is to see how well two Coders agree with each other regarding the application of the coding scheme. **Checking inter-coder reliability is important to do at the beginning of the coding process**, so that if any changes need to be made to your coding scheme you don't have to recode the whole Project. It involves asking a second person to code a small sample of your data also, and then seeing how much overlap there is with your own coding. If overlap is high, then reliability of your coding scheme is high. Then the readers of your research can be more confident that if they were also to apply your Coding Scheme to your data, they would come up with similar results.

Checking reliability involves finding at least one other person familiar with your research area and providing them with your NVivo **Codebook** or your detailed coding scheme in another format.

## Exporting the Codebook

As mentioned in the Introduction to NVivo course, it is good practice to add a detailed description for each Thematic Code you create in your Coding Scheme. This serves as your Codebook outlining the definitions for each Theme, rules for applying it to the data (coding) and, if possible, examples of text that would be coded in such a way. If you have not already gone through this process in NVivo (or elsewhere), you can add a description for each Code by right-clicking on a **Code** in the **Codes Area** and selecting **Code Properties**. Enter your coding rules in the **Description** box and click **OK**.

To export your NVivo Codebook, go to the **Share** tab and click on **Export** then **Export Codebook**. Select the relevant folder containing your Thematic Codes by ticking the appropriate box (usually just the Codes folder). You can also choose to include the number of files and references coded to each Code, by ticking the appropriate box at the bottom, but this would only be needed at the end of the Project when coding is finished. Then click **OK**. A **Word** file is produced and saved showing a formatted table like the one below for the *Sample Project*.



## Environmental Change Down East

### Nodes

Name	Description
Attitude	A person's general feeling of <del>favorableness</del> or <del>unfavorableness</del> . Synonym: affect (Ajzen and Fishbein) Evaluation of people, objects, and ideas (Aronson). A set of beliefs.
Mixed	A person's statement that they are indecisive as to whether their attitude is positive or negative. If a person states two different views, such as that they are positive about one aspect and negative about another, this is not mixed, but rather multiple, feelings.
Negative	A negative opinion or a disagreement where an opinion is not shared with someone or with a group.
Neutral	
Positive	specifying positive opinion <del>or</del> agreement with other stakeholders or with an idea. A shared opinion
Balance	
Community change	Change to the people, culture, or built structures of the communities Down East

## Creating a Set Folder

Select a small random sample of your Data for your second Coder to code. It is helpful for them if you create shortcuts to only these files in a **Sets "meta-folder"**. To create a new **Set** folder, go to the **Create** menu tab and select **Static Set** (or in **Sets** Area right-click on the higher-order **Static Sets** folder and Select **New Static Set**). Then type in a name (i.e., *Interviews for coding comparison*) and description (optional). Then click **OK**. Now you have an empty folder in the **Static Sets** folder to populate with links to any type of Project items (e.g., Files, Cases, Externals, Memos). To add new links, click on the relevant **Set** folder and right-click in the blank-space of the **List View** (or click on **Static Set** again on the **Create** menu tab). Select **Add Static Set Members** and click on the + next to the **Files** folder, click on the name of the folder you want to add items from (e.g., *Interviews*), then tick the boxes for the files you wish to add to the **Set**. Then click **OK**. To remove items from a set, select the items you want to remove and on the **Home** tab click **Delete** (or right-click on the item and select **Delete**). Now all the second Coder needs to do is look in the folder to see which files they need to code.

To make the process even easier for the second Coder (particularly if you have a large and complex coding scheme), you can also **limit the number of Codes** they need to code the data to and put links to these in another **Static Set** folder as a reminder. Prioritise those that are particularly difficult to define or potentially ambiguous, as they are more prone to unreliable coding. Remind your second Coder that they **only need to code to the Child Codes** and not their parents as well, because you can always turn aggregation on if need be. **If you cannot find a 2<sup>nd</sup> person**, code this sample of Data yourself a second time after a few weeks using a different User Profile. Then conduct an inter-Coder reliability test to see how consistently you coded both times.

## Creating a Project Copy

It is best to give your second Coder a copy of your Project to code and **NOT your Master Copy**. It is also **best to perform the Coding reliability check in this copy of the Project**, rather than importing that person's coding into your Master Project (details on how this is done in the next section under **Importing Content from other Projects**). Otherwise, their coding will be added to yours when estimating statistics such as **number of references** (e.g., if you both code the same section of text to the same Code, this counts as 2 references).

To create a copy, go to the **Share** menu tab, click on **Copy Project** and select the appropriate format in the **Copy To** area (e.g., *NVivo - (Windows) R14.23.2 project*) and click the **Browse** button to choose where to save it. Then click **Save**.

The second coder will need to open this copy in NVivo using their own unique **User Profile**, so that NVivo can identify who coded what. (**Note:** if the coder uses your computer to code the Project, it will probably open under your User Profile by default. To change User, go to the **File** menu and select **Options**. On the **General** tab in the **User Profile** area, change the **Name** and **Initials** for the new User. It's also a good idea now to tick the **Prompt for user on launch** so that you don't forget to change it back to yourself). Then click **Apply** and **OK**.

## Coding Comparison Query

You can conduct a **Coding Comparison Query** (on both text and region data types) to get various measurements of inter-Coder reliability, which estimate the **amount of agreement and disagreement between the Coders** for each Code in each document separately (e.g., *Habitat* coding in the *Barbara* interview). These are the notable estimates to check to ascertain reliability:

- **Percentage Agreement** refers to the total number of characters/pixels/timespans in the document on which Coders agree either to code or not to code, divided by the total number in the document. If it's 100%, it's probably because neither of them coded the document to that Code. If it's very high (above 80% for example) it could just be an artefact of the large size of the document. You should always check the following for more clarity on reliability:
  - **A and B** is the percentage of content coded by both Coders and is a less ambiguous and clearer indicator of actual **agreement** in coding. If this is high (above 70% for example), the coding for the Code in that instance is good.
  - **A and Not B** is the percentage coded by User A only, **B and Not A** is the percentage coded by User B only, and both these together are an unambiguous measure of **disagreement**.
- **Cohen's Kappa Coefficient** is a statistic which compares the observed level of agreement on *what should be coded* to that which would have been *expected to be coded* by chance alone. This ranges from -1 (no agreement) to +1 (perfect agreement). It ignores the amount of the file on which both Coders agreed should not be coded, thus Kappa values can be low even when percentage is high. There are several guidelines for determining what is an acceptable coefficient for reliability. Fleiss et al (2003) suggests <0.4 is poor, 0.41-0.75 is fair to good, and 0.76-1.00 is very good. If Kappa is 1, it's probably because both of the Coders did not code anything in the document to that Code.

As you will receive **estimates for one Code across multiple documents**, you will want to report the **average or ranges** of percentages/kappa statistics for each Code.

Example: conduct a Coding Comparison Query between Effie and Henry for a limited number of Cases and Codes.

On the **Explore** tab click the **Coding Comparison** icon and select the users you wish to compare in the **Compare coding between** area (e.g., *Effie Reeves* and *Henry Patterson*). In the **At** box click **Select** and choose the Codes you wish to include in the analysis (e.g., choose **Codes and Cases in Selected Static Sets**, then click the **Select** button and choose the **Codes for coding comparison** folder). Choose the specific Data Files to use in the **Scope** box (e.g. choose **Files, Externals & Memos in Selected Static Sets**, then click the **Select** button and choose the **Interviews for coding comparison** folder). Select the measure(s) you want to use: **Kappa** and/or **Percentage Agreement**, then Select the Comparison Types (tick the **Region Coding** also box if required). Click **Run**.

You can **double-click on a row** in the results where there is some disagreement (e.g., Kappa >0 and below <0.3; less than 80%) to compare the coding conducted and see where they disagreed for that **File** and **Code**. There will be a coloured coding stripe for *Effie's* coding and another colour for *Henry's* with their initials attached. The pink coding stripe represents the aggregated coding for that section (Effie and/or Henry's coding).

## Exporting

### Exporting Results of Queries

If you want to export the results of Queries, it's easiest to right-click on the results themselves (e.g., a Word Cloud) and select **Export**. Choose the file format for the exported file in the **Save As Type** list. In the **File Name** box, change the name if you like. Click **Browse** to select where you want to save it and click **OK**.

### Exporting Codes

In the **List View**, select the Code you want to export. On the **Share** tab, click **Export** and from the **Export** list select the appropriate option:

- **Entire Content (.html)** will create a webpage with your coded data, which is useful for sharing your project findings (including audio and video clips) with colleagues who don't have NVivo. The Code is exported in two parts: an executable \*.html file (e.g., *Balance.html*) and a folder containing files that make the webpage work (e.g., a *Balance* folder). When you open the \*.html file webpage has a tab for the Code's *Summary*, as well as separate tabs for all the coded *Text*, *Audio*, *Video* files etc..
- **Reference View** will export the Reference View to a text document (text only).
- **Summary View** will export the Summary View of the Code to a spreadsheet file.

**Note:** If you want to move a Code HTML file to another location, make sure you copy the folder containing the web page files to the new location as well, or the links will be broken and the webpage won't work. For Codes that code audio and video Data Files, the coded sections of media files will be included upon export only if you coded directly on the audio or video timeline.

Click the **Browse** button under **Save As** and change the name/location of the exported files if you need. If you choose **Entire Content** or **Reference View** there are further options to choose from the **Include** section of the menu (e.g., include *Description*, *Annotations* etc.). Select the Properties and Related Content that you want to include then click **OK**.

### Exporting the List View of the Codes Area

You can export all the information shown in the **List View** of the Codes Area into an Excel file. On the **Share** tab, click on **Export List** (or right-click in the List View of the **Code** area and select **Export List**). The Excel file includes all the Code *Names* and the number of *Files* and *References* associated with them.

### Exporting Classification Sheets

If you want to export a **Classification Sheet** containing your demographic or bibliographic information for example, you can export it as either an Excel, SPSS or Tab Separated data file. You must open the Classification Sheet you want to export (e.g., the *Person* classification in the **Case Classifications** folder). On the **Share** tab click **Export** (or right-click on the **Classification Sheet** and select **Export Classification Sheet**). Select the appropriate export options: (e.g., export to *Microsoft Excel*). Choose a location to save the file in the **Save As** area by clicking on the **Browse** button and selecting the appropriate place. There are other aspects you can change such as *Attribute Value Format* (e.g., change *Unassigned* to *(None)*) and the formatting of *Dates*, *Times* and *Numbers*. Then click **OK**.